



UNIVERSIDADE ESTADUAL DE CAMPINAS  
INSTITUTO DE ESTUDOS DA LINGUAGEM

EMERSON SANTOS DE SOUZA

A PLASTICIDADE DIALETAL DE MIGRANTES BAIANOS  
NA REGIÃO METROPOLITANA DE SÃO PAULO

CAMPINAS  
2023

EMERSON SANTOS DE SOUZA

A PLASTICIDADE DIALETAL DE MIGRANTES BAIANOS  
NA REGIÃO METROPOLITANA DE SÃO PAULO

Tese apresentada ao Programa de Pós-Graduação em Linguística do Instituto de Estudos da Linguagem da Universidade Estadual de Campinas como requisito parcial para a obtenção do título de Doutor em Linguística.

Orientadora: Profa. Dra. Livia Oushiro

Este exemplar corresponde à versão final da Tese defendida pelo aluno Emerson Santos de Souza e orientada pela Profa. Dra. Livia Oushiro.

CAMPINAS

2023

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

So89p Souza, Emerson Santos de, 1985-  
A plasticidade dialetal de migrantes baianos na Região Metropolitana de São Paulo / Emerson Santos de Souza. – Campinas, SP : [s.n.], 2023.

Orientador: Livia Oushiro.  
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Estudos da Linguagem.

1. Sociolinguística. 2. Dialetologia. 3. Contato dialetal. 4. Migrantes. 5. Língua Portuguesa - Regionalismos - Bahia. 6. Língua Portuguesa - Regionalismos - São Paulo (Estado). I. Oushiro, Livia, 1980-. II. Universidade Estadual de Campinas. Instituto de Estudos da Linguagem. III. Título.

Informações Complementares

**Título em outro idioma:**

Dialectal plasticity in the speech of Bahian migrants in the São Paulo Metropolitan Region

**Palavras-chave em inglês:**

Sociolinguistics

Dialetology

Dialectal Contact

Migrants

Portuguese language - Provincialisms - Bahia (Brazil : State)

Portuguese language - Provincialisms - São Paulo (Brazil : State)

**Área de concentração:** Linguística

**Titulação:** Doutor em Linguística

**Banca examinadora:**

Livia Oushiro [Orientador]

Anna Christina Bentes da Silva

Ronald Beline Mendes

Josane Moreira de Oliveira

Elisa Battisti

**Data de defesa:** 27-11-2023

**Programa de Pós-Graduação:** Linguística

**Identificação e informações acadêmicas do(a) aluno(a)**

- ORCID do autor: <https://orcid.org/0000-0003-4123-0788>

- Currículo Lattes do autor: <https://lattes.cnpq.br/6006857841226887>



**BANCA EXAMINADORA:**

**Livia Oushiro**

**Anna Christina Bentes da Silva**

**Ronald Beline Mendes**

**Josane Moreira de Oliveira**

**Elisa Battisti**

**IEL/UNICAMP  
2023**

**Ata da defesa, assinada pelos membros da Comissão Examinadora, consta no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria de Pós Graduação do IEL.**

*Para Arleide Souza, minha mãe.*

# Agradecimentos

O período do meu doutoramento foi repleto de desafios. A ausência de bolsa, devido aos cortes de investimentos na Educação, a pandemia da Covid-19, as demandas do trabalho e as crises de ansiedade quase me fizeram desistir de um dos meus sonhos: tornar-me Doutor. Nessa travessia, que durou mais de seis anos, a caminhada só foi possível, porque tive uma rede de apoio de Anjos visíveis e invisíveis. Para eles, reporto-me agora em forma de agradecimento.

Agradeço a Deus e aos Anjos de luz por terem me dado ânimo e sabedoria para concluir uma das etapas da minha vida acadêmica.

Aos meus pais, Arleide e Edvaldo (Mainha e Painho), não tenho palavras para expressar a minha gratidão pelo amor inabalável que vocês me deram ao longo dessa trajetória. Desde quando iniciei o Doutorado, vocês me apoiaram, incentivando-me a nunca esmorecer e a acreditar em mim mesmo. Vocês são o meu alicerce.

À minha orientadora, Livia Oushiro, por nunca ter soltado a minha mão nos momentos mais difíceis. Livia, tenha certeza de que você me ensinou muito mais que fazer pesquisa em Sociolinguística. Meus futuros orientandos serão vistos com o mesmo olhar humano que você teve para comigo. Você é um dos Seres de luz que Deus colocou em minha vida.

Ao meu esposo, Léo, por todos os atos de serviço. Amor, você não apenas me ouviu atentamente a cada vez que precisei desabafar, mas também me fez enxergar que eu era capaz de vencer qualquer obstáculo. O meu sucesso neste trabalho não teria sido possível sem você ao meu lado.

Ao meu filho pet, Dudu, pela presença. Mesmo sem entender o que eu fazia sentado à escrivaninha boa parte do dia, ele estava lá, deitado com sua cabeça sobre o meu pé. E, nos momentos de desespero, cansaço e estresse, ele me acalmava com os “lambeijos”.

A Anderson, meu irmão; Ravana, minha cunhada; Miguel, meu sobrinho; Goreti e Eduardo, meus sogros; Lucas, meu cunhado, agradeço por todo o apoio.

À Josane Oliveira, minha mãe acadêmica, por me acompanhar desde o período da Graduação e ter me incentivado a cursar o Doutorado na UNICAMP.

À Joana Figueredo, “minha marida”, pela parceria e cumplicidade desde a época do mestrado. À Williane Corôa e Antônio Codina por sempre me ouvirem e festejarem a cada etapa do Doutorado concluída. O meu doutoramento só foi possível, porque vocês estiveram ao

meu lado.

Agradeço aos colegas/amigos do Laboratório Variação, Identidade, Estilo e Mudança (VARIEM), Julia Adams, Natasha Mourão, Flávia Sousa, Andreza Parra, Marina Russo, Fernanda Guirelli, Gabriel Catani, Danielle Bento, Sarah Poli, Gustavo Silveira e Leonardo Ferraz por terem contribuído diretamente na minha pesquisa.

Aos meus colegas/amigos do Instituto de Estudos da Linguagem (IEL/UNICAMP) Júlio César Cavalcante, Lucas D'Alva, Sônia Coutinho, Victor Veríssimo, Laís Medeiros, Shirley Guedes, Flavio Benayon e Liliane Anjos pela troca de saberes.

Aos professores Anna Bentes e Ronald Mendes pelas valiosas contribuições no Exame de Qualificação da Tese.

Aos professores Charlotte Galves e Plínio Barbosa pelo carinho, atenção e respeito.

Aos meus colegas e alunos do Centro de Educação Municipal Oliveira Brito (CE-MOB) pela compreensão e solidariedade durante o processo de escrita desta Tese.

Aos meus queridos amigos, Allexandra Souza, Maria do Carmo, Ana Glória, Isabelle Costa, Eliene Alves, Daniel Ribeiro, André Nascimento, Lud Barros, Tereza Borges, Candice Luz, Rogério Reis, Mirley Melo, Luzia Boccanera, Vinicius Soglia, Victor Costa, Monique Souza, Léo Tenório, Laís Maciel, Marília Amorim, Rômulo Freire, Layla Cupertino, Phillippe Cupertino, Hilana Rios e André Mascarenhas pela rede de apoio.

“Aqueles que passam por nós não vão sós, não nos deixam sós. Deixam um pouco de si, levam um pouco de nós.”

**Antoine de Saint-Exupéry,**  
O Pequeno Príncipe

# Resumo

Esta Tese aborda as adaptações linguísticas resultantes do contato de migrantes baianos com paulistas na Região Metropolitana de São Paulo (RMSP). Os objetivos deste estudo foram: (i) analisar se variáveis sociolinguísticas de diferentes naturezas perpassam os mesmos processos de adaptação durante o contato dialetal; e (ii) avaliar como se dá o encaixamento social e linguístico de suas variantes. Para atingir esses objetivos, recrutaram-se 50 migrantes baianos residentes na RMSP, por meio do método de rede social (Bortoni-Ricardo, 2011 [1985]), com os quais se realizaram entrevistas sociolinguísticas a fim de coletar dados das variáveis lexicais Tangerina/Mexerica, Aipim/Mandioca, Quentinha/Marmita, Trabalho/Serviço; da variável fonético-fonológica (-r) em coda silábica, como em “porta” e “mar”, e da variável morfossintática Negação sentencial, como em Neg1, “*Não* vou”; Neg2, “*Não* vou *não*” e Neg3, “*Vou não*”. Para obter os resultados, realizaram-se, na plataforma R (R Core, 2023), análises de regressão logística de efeitos mistos com a inclusão do Participante e do Item lexical como variáveis aleatórias. A expectativa era a de que os migrantes usariam as variantes típicas da comunidade anfitriã (Mexerica, Mandioca, Marmita, Serviço, /r/ Retroflexo/Tepe e a baixa frequência de uso de Neg2/Neg3) em relação às variantes típicas da comunidade de origem (Tangerina, Aipim, Quentinha, Trabalho, /r/ Velar/Glotal e a alta frequência de Neg2/Neg3, respectivamente). Para descrever os padrões de encaixamento social e linguístico das adaptações dialetais dos migrantes baianos, testou-se a correlação desse conjunto de variáveis dependentes com os preditores Sexo/Gênero (homem; mulher), Idade (variável contínua), Nível de escolaridade (Ensino Fundamental; Ensino Médio), Idade de migração (variável contínua), Tempo de residência (variável contínua), Rede de contato com baianos (variável contínua) e Grau de autoidentificação paulista declarada (variável contínua) e com os preditores linguísticos que têm sido analisados nos estudos de cada uma das variáveis dependentes examinadas. Os resultados revelam que os processos de adaptação dialetal não estão relacionados apenas à natureza das variáveis sociolinguísticas, mas aos casos de regularidade e irregularidade linguística entre os dialetos em contato. Nos casos de regularidade, isto é, quando os dialetos em contato compartilham a mesma variável sociolinguística, mas divergem quanto às taxas de uso de suas variantes, identificaram-se os processos de *potencialização* e *contenção*. Entende-se por potencialização o processo de adaptação linguística em que as taxas das variantes de menor uso

na gramática nativa dos migrantes são elevadas devido a sua alta frequência na comunidade anfitriã. Denomina-se contenção o processo de adaptação em que os migrantes reduzem as taxas das variantes de maior uso em sua gramática nativa por influência de suas baixas taxas de uso na comunidade anfitriã. Já nos casos de irregularidade, i.e., quando os dialetos em contato se diferenciam quanto à presença ou à ausência de certas variantes, notaram-se os processos de *aquisição*, *ampliação de contexto* e *ampliação semântica*. Entende-se aqui por aquisição o processo de adaptação dialetal em que os migrantes adquirem variantes da comunidade anfitriã que são totalmente atípicas em sua gramática nativa. A ampliação de contexto é o processo em que os migrantes expandem os contextos de uso de variantes que têm contextos restritos em sua gramática nativa. Entende-se por ampliação semântica o processo em que os migrantes incorporam a variantes da sua gramática nativa sentidos utilizados na comunidade anfitriã. Além desses resultados, observou-se que: (i) migrantes de uma mesma comunidade de origem residentes em pontos diferentes de uma mesma região, como capital e interior, podem apresentar padrões distintos quanto ao uso de uma mesma variável sociolinguística; (ii) migrantes de origens distintas, mas que residem numa mesma comunidade anfitriã podem apresentar padrões diferentes para uma mesma variável sociolinguística. Os resultados do presente estudo permitem afirmar que os termos “acomodação”, utilizado por [Trudgill \(1986\)](#), e “aquisição de segundo dialeto”, usado por [Chambers \(1992\)](#) e [Siegel \(2010\)](#), não abarcam todos os processos de adaptação linguística identificados na fala de migrantes em situação de contato dialetal. Considerando a amplitude de tais processos, propõe-se o termo “plasticidade dialetal”.

Palavras-chave: plasticidade dialetal; contato dialetal; fala de migrantes; português baiano; português paulista.

# Abstract

This thesis addresses the linguistic adaptations resulting from the contact between Bahian migrants and Paulistas in São Paulo Metropolitan Region. The objectives of this study were: (i) to analyze whether sociolinguistic variables of different natures undergo the same adaptation processes during dialectal contact, and (ii) to evaluate how the social and linguistic embedding of their variants takes place. To achieve these objectives, 50 Bahian migrants residing in São Paulo were recruited using method of the social network (Bortoni-Ricardo, 2011 [1985]), and sociolinguistic interviews were conducted to collect data on lexical variables *Tangerina/Mexerica*, *Aipim/Mandioca*, *Quentinha/Marmita*, *Trabalho/Serviço*; the phonetic-phonological variable (-r) in coda, as in “porta” and “mar”, and the morphosyntactic variable (Neg), as in Neg1, “*Não vou*”; Neg2, “*Não vou não*”; and Neg3, “*Vou não*”. To obtain the results, logistic regression analyses with mixed effects were performed in the R platform (R Core, 2023), including Participant and Lexical Item as random variables. The expectation was that migrants would use the variants typical of the host community (*Mexerica*, *Mandioca*, *Marmita*, *Serviço*, *Retroflex/Tepe /r/*, and low frequency of *Neg2/Neg3*) in relation to the variants typical of the home community (*Tangerina*, *Aipim*, *Quentinha*, *Trabalho*, *Velar/Glotal /r/*, and relatively higher frequency of *Neg2/Neg3*, respectively). To describe the patterns of social and linguistic embedding of the dialectal adaptations of Bahian migrants, the correlation of this set of dependent variables with the predictors Gender (male; female), Age (continuous variable), Education Level (Elementary School; High School), Age of Migration (continuous variable), Length of Residency (continuous variable), Contacts with Bahian (continuous variable), and Declared Level of Paulista Self-Identification (continuous variable), as well as the linguistic predictors that have been analyzed in the studies of each of the dependent variables examined. The results revealed that the processes of dialectal adaptation are not only are related to the nature of sociolinguistic variables but also to cases of linguistic regularity and irregularity between the dialects in contact. In cases of regularity, where the dialects in contact share the same sociolinguistic variable but differ in the rates of usage of their variants, processes of enhancement and constraint were identified. Enhancement is the process of linguistic adaptation in which the rates of less commonly used variants in the migrants’ native grammar are increased due to their high frequency in the host community. Constraint, on the other hand, is the process of adaptation in which

migrants reduce the rates of more commonly used variants in their native grammar due to the low rates of usage in the host community. In cases of irregularity, where the dialects in contact differ in the presence or absence of certain variants, processes of acquisition, context expansion, and semantic expansion were observed. Acquisition is the dialectal adaptation process in which migrants acquire variants from the host community that are entirely atypical in their native grammar. Context expansion is the process in which migrants expand the contexts of use of variants that have restricted contexts in their native grammar. Semantic expansion is the process in which migrants incorporate meanings from their native grammar into the variants used in the host community. Beyond these results, it was observed that: (i) migrants from the same community of origin residing in different points of the same region, such as capital and interior, may exhibit distinct patterns in the use of the same sociolinguistic variable; (ii) migrants from different origins, but residing in the same host community, may present different patterns for the same sociolinguistic variable. The results of this study allow us to assert that the terms “accommodation”, used by [Trudgill \(1986\)](#), and “second dialect acquisition”, used by [Chambers \(1992\)](#) and [Siegel \(2010\)](#), do not encompass all the linguistic adaptation processes identified in the speech of migrants in dialectal contact. Given the breadth of these processes, the term “dialectal plasticity” is proposed.

Keywords: dialectal plasticity; dialectal contact; migrant speech; baiano portuguese; paulista portuguese.

# Lista de Figuras

1	Proporção da população entre 30 e 60 anos de idade da Região Metropolitana de São Paulo segundo a naturalidade . . . . .	24
2	Rede 1 – AlanaS . . . . .	55
3	Rede 2 – MarieneS . . . . .	56
4	Rede 3 – JoãoS . . . . .	58
5	Frequência da idade dos migrantes baianos residentes na RMSP participantes da amostra . . . . .	59
6	Frequência da idade de migração dos baianos residentes na RMSP participantes da amostra . . . . .	61
7	Frequência da escala de preconceito contra os migrantes baianos residentes na RMSP participantes da amostra . . . . .	62
8	Frequência do grau de autoidentificação nordestina, baiana e paulista declarada dos migrantes baianos residentes na RMSP participantes da amostra . . . . .	63
9	Frequência do tempo de residência dos migrantes baianos residentes na RMSP participantes da amostra . . . . .	64
10	Comparação das proporções dos itens lexicais Tangerina e Mexerica de baianos não migrantes, de migrantes baianos na RMSP e de paulistas não migrantes . . .	78
11	Comparação das proporções dos itens lexicais Aipim e Mandioca de baianos não migrantes, de migrantes baianos na RMSP e de paulistas não migrantes . . .	79
12	Comparação das proporções da realização dos itens lexicais Quentinha e Marmita de baianos na Bahia, de migrantes baianos na RMSP e de paulistas não migrantes . . . . .	81
13	Comparação das proporções de uso dos itens lexicais Trabalho e Serviço de baianos não migrantes, de migrantes baianos na RMSP e de paulistas não migrantes	83
14	Proporção dos Itens lexicais baianos e paulistas utilizados pelos migrantes baianos residentes na RMSP . . . . .	84
15	Correlação dos Itens lexicais paulistas e a Idade de migração dos baianos residentes na RMSP . . . . .	87

16	Proporção dos Itens lexicais paulistas de acordo com a Faixa de idade de migração dos baianos residentes na RMSP . . . . .	88
17	Gráfico da correlação dos Itens lexicais paulistas e o Tempo de residência dos migrantes baianos na RMSP . . . . .	89
18	Proporção dos Itens lexicais paulistas de acordo com a Faixa de tempo de residência dos baianos na RMSP . . . . .	90
19	Comparação das proporções de uso do (-r) em coda silábica de baianos não migrantes, de migrantes baianos na RMSP e de paulistas não migrantes . . . . .	106
20	Proporção do (-r) em coda silábica da fala de migrantes baianos residentes na RMSP depois da exclusão dos dados de Apagamento e da junção das ocorrências do Retroflexo com o Tepe . . . . .	107
21	Interação das variáveis Contexto fônico precedente e Rede de contato com baianos em relação ao uso do /r/ Retroflexo/Tepe de migrantes baianos residentes na RMSP . . . . .	119
22	Interação das variáveis Contexto fônico seguinte e Rede de contato com baianos em relação ao uso do /r/ Retroflexo/Tepe na fala dos migrantes baianos residentes na RMSP . . . . .	121
23	Interação das variáveis Estilo e Posição da sílaba quanto ao uso do /r/ Retroflexo/Tepe dos migrantes baianos residentes na RMSP . . . . .	123
24	Comparação de Estilo quanto ao uso do (-r) em posição de coda dos migrantes baianos residentes na RMSP . . . . .	124
25	Comparação das proporções da Negação sentencial de baianos não migrantes, migrantes baianos residentes na RMSP e paulistas não migrantes . . . . .	139

# Lista de Tabelas

1	Distribuição dos migrantes baianos residentes na RMSP participantes da amostra segundo a proveniência dos integrantes de suas redes sociais . . . . .	65
2	Frequência e proporção do emprego das variáveis lexicais Tangerina/Mexerica, Aipim/Mandioca, Quentinha/Marmita, Trabalho/Serviço nas variedades baiana e paulista . . . . .	76
3	Resultados do modelo de regressão logística de efeitos fixos quanto à aplicação dos Itens lexicais paulistas dos migrantes baianos na RMSP segundo o Sexo/-Gênero . . . . .	85
4	Resultados do modelo de regressão logística de efeitos fixos para a realização dos Itens lexicais paulistas dos migrantes baianos na RMSP segundo a Idade . . . . .	85
5	Resultados do modelo de regressão logística de efeitos fixos para a realização dos Itens lexicais paulistas dos migrantes baianos na RMSP segundo a Rede de contato com baianos . . . . .	85
6	Resultados do modelo de regressão logística de efeitos fixos para a realização dos Itens lexicais paulistas dos migrantes baianos na RMSP segundo o Grau de autoidentificação paulista declarada . . . . .	86
7	Resultados do modelo de regressão logística de efeitos fixos quanto à aplicação dos Itens lexicais paulistas dos migrantes baianos na RMSP segundo o Nível de escolaridade . . . . .	86
8	Resultados do modelo de regressão logística de efeitos fixos para a realização dos Itens lexicais paulistas dos migrantes baianos na RMSP segundo a Idade de migração . . . . .	87
9	Resultados do modelo de regressão logística de efeitos fixos da realização dos Itens lexicais paulistas segundo o Tempo de residência dos migrantes baianos na RMSP . . . . .	89
10	Distribuição das ocorrências do (-r) em coda silábica medial e final da fala dos migrantes baianos residentes na RMSP . . . . .	102

11	Resultados do primeiro modelo de regressão logística de efeitos mistos com inclusão do Participante e do Item lexical como variáveis aleatórias quanto ao uso do /r/ Retroflexo/Tepe na fala dos migrantes baianos residentes na RMSP . . . . .	108
12	Resultados do segundo modelo de regressão logística de efeitos mistos com inclusão do Participante e do Item lexical como variáveis aleatórias quanto ao uso do /r/ Retroflexo/Tepe da fala dos migrantes baianos na RMSP . . . . .	110
13	Resultados do terceiro modelo de regressão logística de efeitos mistos com inclusão do Participante e do Item lexical como variáveis aleatórias quanto ao uso do /r/ Retroflexo/Tepe da fala dos migrantes baianos na RMSP . . . . .	110
14	Taxas de uso das variantes de (Neg) em variedades do PB . . . . .	129
15	Resultados do primeiro modelo de regressão logística de efeitos mistos com inclusão do Participante como variável aleatória quanto ao uso de Neg2/Neg3 na fala dos migrantes baianos residentes na RMSP . . . . .	140
16	Resultados do segundo modelo de regressão logística de efeitos mistos com inclusão do Participante como variável aleatória quanto ao uso de Neg2/Neg3 na fala dos migrantes baianos residentes na RMSP . . . . .	141

# Lista de Quadros

1	Variáveis linguísticas testadas na análise do (-r) em coda . . . . .	103
2	Variáveis linguísticas testadas na análise de (Neg) . . . . .	134
3	Comparação das correlações das variáveis sociolinguísticas Lexias, (-r) em coda e (Neg) identificadas na fala dos migrantes baianos na RMSP com os preditores sociais . . . . .	150
4	Comparação das correlações do uso do (-r) em coda na fala dos migrantes baianos residentes em Bauru (SP) e na RMSP com os preditores sociais . . . . .	151
5	Comparação das correlações do uso de (Neg) na fala dos migrantes baianos e sergipanos residentes na RMSP com os preditores sociais . . . . .	152
6	Comparação dos padrões de Estilo (modelo laboviano) de migrantes em situação de contato dialetal quanto ao emprego de variantes típicas da comunidade de origem (VTCO), da comunidade anfitriã (VTCA) e da comunidade de onde os migrantes regressaram (VTR) . . . . .	154

# Sumário

<b>1</b>	<b>Introdução</b>	<b>20</b>
<b>2</b>	<b>O contato entre dialetos mutuamente inteligíveis</b>	<b>26</b>
2.1	Base epistemológica para estudos sobre o contato entre dialetos mutuamente inteligíveis . . . . .	26
2.2	Outros estudos sobre o contato entre variedades do português brasileiro . . . . .	32
2.3	Síntese . . . . .	48
<b>3</b>	<b>Materiais e métodos</b>	<b>50</b>
3.1	Constituição da amostra . . . . .	50
3.2	Roteiro utilizado nas entrevistas sociolinguísticas . . . . .	52
3.3	Os participantes da amostra . . . . .	53
3.4	Distribuição dos participantes nas categorias sociais investigadas . . . . .	58
3.5	Manipulação e análise de dados . . . . .	66
3.6	Preditores sociais e hipóteses . . . . .	67
3.7	Síntese . . . . .	71
<b>4</b>	<b>Adaptação em nível lexical: Tangerina/Mexerica, Aipim/Mandioca, Quentinha/Marmita e Trabalho/Serviço</b>	<b>73</b>
4.1	O uso variável de Tangerina/Mexerica, Aipim/Mandioca, Quentinha/Marmita e Trabalho/Serviço no português brasileiro: Bahia e São Paulo em destaque . . . . .	74
4.2	Manipulação dos dados . . . . .	77
4.3	Resultados . . . . .	78
4.3.1	Descrição dos achados . . . . .	78
4.3.2	Discussão dos achados . . . . .	91
4.4	Síntese . . . . .	93
<b>5</b>	<b>Adaptação em nível fonético-fonológico: o (-r) em coda silábica</b>	<b>95</b>
5.1	A variação do (-r) em posição de coda silábica em variedades do português brasileiro: Bahia e São Paulo em evidência . . . . .	96
5.2	Manipulação dos dados . . . . .	101

5.3	Resultados . . . . .	105
5.3.1	Descrição dos achados . . . . .	105
5.3.2	Discussão dos achados . . . . .	111
5.4	Síntese . . . . .	125
<b>6</b>	<b>Adaptação em nível morfossintático: a negação sentencial</b>	<b>126</b>
6.1	A variação sintática do negador “não” em variedades do português brasileiro: Bahia e São Paulo em foco . . . . .	127
6.2	Manipulação dos dados . . . . .	133
6.3	Resultados . . . . .	138
6.3.1	Descrição dos achados . . . . .	139
6.3.2	Discussão dos achados . . . . .	142
6.4	Síntese . . . . .	144
<b>7</b>	<b>A plasticidade dialetal</b>	<b>146</b>
7.1	Processos de adaptação linguística de migrantes em situação de contato dialetal	146
7.2	As adaptações linguísticas de migrantes em situação de contato dialetal no con- texto social e estilístico . . . . .	148
7.3	Sugestões de análises . . . . .	156
<b>8</b>	<b>Considerações finais</b>	<b>158</b>
	<b>Referências</b>	<b>162</b>
	<b>Apêndices</b>	<b>172</b>

# 1

## Introdução

Esta Tese trata de adaptações linguísticas resultantes do contato entre dialetos<sup>1</sup> mutuamente inteligíveis (doravante contato dialetal). Quanto a esse assunto, há um consenso entre Chambers (1992) e Siegel (2010) sobre as críticas à aplicação da *Teoria da Acomodação e Comunicação*, da Psicologia Social (Giles; Taylor; Bourhis, 1973), nos estudos sociolinguísticos que abordam as adaptações linguísticas de migrantes em situação de contato dialetal. A argumentação de Chambers (1992) e Siegel (2010) baseia-se essencialmente nas teses de que as adaptações linguísticas identificadas na fala dos migrantes em situação de contato dialetal nem sempre acontecem face a face e nem sempre as atitudes positivas as favorecem, como defendido por Trudgill (1986). Aliás, como se verá mais adiante, os resultados obtidos no presente estudo corroboram essa segunda afirmação. Assim, do ponto de vista teórico, o uso do termo *acomodação* (a longo prazo) para descrever os efeitos do contato dialetal é problemático.

Do ponto de vista técnico, também se observam problemas no emprego desse termo em trabalhos sobre a fala de migrantes. Um deles é a sobreposição da área de conhecimento. Reconhecendo que a Teoria da Acomodação é uma epistemologia da Psicologia Social, a menos que se especifique a área de concentração: como identificar, à primeira vista, a que área pertence um trabalho intitulado “Acomodação dialetal de baianos na Região Metropolitana de São Paulo”? Seria da Psicologia Social? Da Sociolinguística? Embora essas áreas dialoguem, cada uma tem o seu escopo.

Outro problema técnico está na carga negativa popularmente difundida do adjetivo “acomodado(a)”. É bem verdade que esse adjetivo é sinônimo de “adaptado”, “ajustado”, “instalado” (Ferreira, 2001, p. 17), mas também significa “que aceita ou se submete a uma situação indesejada” (Houaiss, 2012, p. 13). Ao que parece, este significado tem ganhado sentidos mais negativos. Quando se diz, por exemplo, que “João está acomodado” pode-se ter a interpretação de que João está adaptado/instalado em algum lugar, mas também de que João é “passivo” di-

---

<sup>1</sup> Com base em Brandão (1991), empregam-se, neste estudo, os termos “dialetos”, “variedades linguísticas” e “falares” como sinônimos.

ante de alguma situação, é “inerte”, “preguiçoso”. Sendo a população baiana historicamente e midiaticamente estigmatizada como “preguiçosa”, “descansada”, a interpretação do título “Acomodação dialetal de baianos residentes na Região Metropolitana de São Paulo” poderia potencializar tal preconceito.

Em oposição ao emprego da Teoria da Acomodação nos estudos sobre a fala de migrantes, [Chambers \(1992\)](#) e [Siegel \(2010\)](#) propõem que as adaptações linguísticas resultantes do contato dialetal são fruto da *Aquisição de segundo dialeto*. No entanto, ao compreender que, no processo de aquisição por meio do contato, o falante adquire novas variantes, como considerar casos nos quais os dialetos em contato frequentemente diferem apenas nas taxas de uso de variantes de uma variável sociolinguística específica?

A título de ilustração, no português brasileiro, o negador “não” pode variar de posição em relação ao verbo, aparecendo anteposto a ele, como em “*Não* vou” (Neg1); antes e depois do verbo, como em “*Não* vou *não*” (Neg2) e posposto ao verbo, como em “Vou *não*” (Neg3). Ao comparar os dados da Negação sentencial (Neg) identificados em Matinha (BA) (Neg1 78%; Neg2 20% e Neg3 2%) por [Santana e Nascimento \(2011\)](#), com os encontrados em São Paulo (SP) (Neg1 94%; Neg2 5,8% e Neg3 0,2%) por [Rocha \(2013\)](#) e com os observados em Porto Alegre (RS) (Neg1 99,4%; Neg2 0,6% e Neg3 0%) por [Goldnadel et al. \(2013\)](#), observa-se que as características dialetais divergem apenas quanto às taxas de uso de (Neg) (ver a Seção 6).

Sendo assim, é coerente considerar aquisição o processo no qual um porto-alegrense incorpore Neg3 ao seu repertório linguístico ao manter de contato com baianos, porque essa variante parece não estar presente em sua gramática nativa. No entanto, não seria correto classificar como aquisição o caso em que um baiano deixa de usar Neg3 por manter contato com os porto-alegrenses. Da mesma forma, não se pode considerar aquisição qualquer resultado de variação das formas da Negação sentencial no contato entre baianos e paulistas, porque ambos os dialetos compartilham de suas variantes, embora em proporções diferentes. Em uma pesquisa sobre a fala de migrantes, não atentar para as particularidades dos dialetos em contato é marginalizar um conjunto de variáveis sociolinguísticas que matizam a diversidade linguística.

Os resultados da pesquisa que aqui se apresenta, obtidos das análises de uma amostra de fala de 50 migrantes baianos residentes na Região Metropolitana de São Paulo (doravante RMSP), recrutados a partir do método de rede social ([Milroy, 1987 \[1980\]](#); [Bortoni-Ricardo, 2011 \[1985\]](#); [Santana, 2018](#)) (ver a Seção 3), mostram que aquisição é um de tantos outros processos de adaptação linguística que podem ser identificados na fala de migrantes em situação de contato dialetal. Além disso, os resultados revelam que tais processos estão ligados também às características linguísticas e sociais dos dialetos em contato.

Conforme reportado na Seção 7, a classificação de tais processos depende da identificação dos casos de regularidade e irregularidade linguística nos dialetos em contato. Por

regularidade entende-se quando os dialetos em contato compartilham uma mesma variável sociolinguística, mas se diferenciam apenas quanto às proporções de uso de suas variantes. Por irregularidade entende-se os casos em que os dialetos em contato divergem quanto à presença de determinada variante na variedade da comunidade anfitriã e a sua ausência na variedade nativa do migrante; à versatilidade de contextos de uso de determinada variante na comunidade anfitriã e sua restrição na gramática nativa do migrante; à diferença de significados de significantes compartilhados pelos dialetos em contato. Em outras palavras, é preciso ponderar (i) a distribuição das variantes x e y da variável z empregada nos dialetos A e B; (ii) e a existência e inexistência de variantes entre os dialetos A e B.

Em relação aos casos de regularidade, tomando como referência as adaptações dos migrantes baianos em favor da variedade paulista, verifica-se que o contato dialetal resultou em *potencialização* ou *contenção* de determinada variante. Quanto aos casos de irregularidade, descobriu-se que houve *aquisição*, *ampliação de significado* e *ampliação de contexto*. Ademais, ao comparar os resultados obtidos das análises da amostra de fala aqui investigada com outros que se debruçaram sobre a fala de migrantes no Estado de São Paulo, viu-se que: (i) migrantes da mesma origem e residentes em pontos diferentes de uma mesma região, como capital e interior, podem apresentar padrões diferentes para uma variável sociolinguística específica e (ii) migrantes de origens distintas, ainda que de áreas geográficas próximas, e residentes numa mesma comunidade anfitriã podem apresentar padrões distintos para uma mesma variável sociolinguística.

Como se pode notar, o uso do termo “acomodação” e “aquisição” para denominar os efeitos linguísticos do contato dialetal, quando não problemático, é reducionista. Assim, reconhecendo que “o léxico categoriza o conhecimento humano na forma de palavras” (Biderman, 1998, p. 164), urge um termo capaz de abarcar os diferentes processos de adaptação linguística da fala de migrantes bem como os seus resultados variáveis em relação aos aspectos sociais e linguísticos da comunidade de origem<sup>2</sup> e da comunidade anfitriã.

O termo *plasticidade* parece englobar todas essas nuances. Segundo Ferreira (1986, p. 1345), plasticidade é “o estado daquilo que é plástico”, isto é, de algo “que tem propriedade de adquirir determinadas formas sensíveis, por efeito de uma ação exterior”. Esse termo tem sido utilizado em diferentes áreas do conhecimento para nomear as adaptações que um mesmo tipo de ser vivo (Biologia), órgão (Neurologia, Fonoaudiologia) ou material (Física) pode assumir mediante às condições externas.

Na Biologia, por exemplo, denomina-se plasticidade fenotípica a capacidade que

---

<sup>2</sup> Atualmente, consideram-se os termos “origem” e “destino” reducionistas nos estudos sobre as migrações, porque não contemplam as sazonais e temporárias. Apesar disso, eles são utilizados neste estudo por ainda não haver um consenso entre os especialistas de como tratá-los em diferentes casos de migração. Para outras informações, recomenda-se a leitura de Menezes (2012, p. 21–40).

certo ser vivo tem de apresentar diferentes características (morfológicas, fisiológicas, comportamentais etc.) em função do meio em que vive (Lima et al., 2017). A título de exemplo, um mesmo tipo de vegetal cultivado em lugares diferentes pode apresentar variações em sua estrutura (a altura do caule, a quantidade de folhas por ramo, a sua morfologia, por exemplo) devido às condições ambientais em que ele está inserido, como a altitude, a iluminação, a umidade, entre outros.

Na Sociolinguística, por sua vez, ao considerar que (i) a língua é passível de variação (Weinreich; Labov; Herzog, 2006 [1968]); (ii) os dialetos mutuamente inteligíveis compartilham a mesma base de uma língua, diferenciando-se apenas em certos aspectos de acordo com a área geográfica em que são falados (Brandão, 1991; Trudgill, 2004; Siegel, 2010; Cardoso et al., 2014); (iii) a língua tende a variar conforme os contextos sociais e linguísticos (Weinreich; Labov; Herzog, 2006 [1968]; Labov, 2008 [1972], 2001); o termo *plasticidade dialetal* parece ser apropriado para os estudos sociolinguísticos sobre a fala de migrantes em situação de contato dialetal.

Esta Tese, portanto, trata da *plasticidade dialetal de migrantes baianos residentes na Região Metropolitana de São Paulo*. As motivações para a realização da presente pesquisa, incluindo a escolha da população observada e o escopo das análises, são atribuídas à parceria com o projeto Processos de acomodação dialetal na fala de nordestinos residentes em São Paulo<sup>3</sup> (Oushiro, 2017) (doravante Projeto Acomodação). Esse projeto tem se dedicado, sistematicamente, ao estudo da fala de migrantes (ver a Seção 2), com o objetivo de ampliar as discussões acerca da variação e da mudança linguística, uma vez que os falantes móveis, durante muito tempo, foram marginalizados nos estudos sobre comunidades de fala (Chambers, 1992; Oushiro, 2020c).

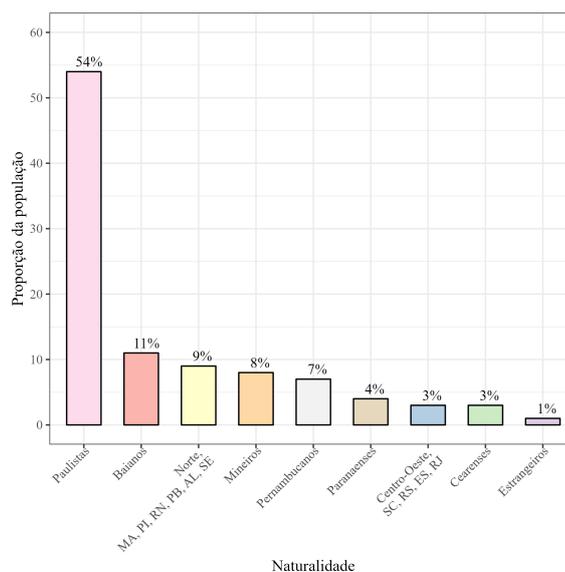
Acerca da população investigada neste estudo, sabe-se que a migração de nordestinos para São Paulo (SP) foi, durante muito tempo, reflexo, sobretudo, da precária condição social do Nordeste e do processo de industrialização dessa capital na primeira metade do século XX. No entanto, esse êxodo não ocorreu num simples estalar de dedos, tampouco de modo aleatório. Com a diminuição da migração estrangeira, o governo redirecionou os subsídios destinados a ela para a migração nacional, que, basicamente, resumia-se ao norte de Minas Gerais e à Região Nordeste (Fontes, 2008). Dessa maneira, vê-se que a primeira leva de migrantes nordestinos para São Paulo (SP) foi agenciada. Tempos depois, de acordo com Fontes (2008), foi quando os próprios migrantes nordestinos, tendo se estabelecido nessa capital, promoveram o movimento migratório de parentes e amigos para essa cidade, formando uma rede de apoio.

Da Região Nordeste, a maior leva de migrantes saiu da Bahia. Segundo Fontes (2008, p. 70), entre as décadas de 1950 e 1960, os órgãos de controle da migração registraram

<sup>3</sup> O projeto teve o financiamento da FAPESP, cujo processo está registrado com o número 2016/04960-7.

cerca de 330 mil trabalhadores baianos na capital paulista. Em dados de 2011, conforme mostra a Figura 1, verifica-se que os baianos representam, depois dos paulistas, a segunda maior população da Região Metropolitana de São Paulo com idade entre 30 e 60 anos, totalizando 11%.

**Figura 1** – Proporção da população entre 30 e 60 anos de idade da Região Metropolitana de São Paulo segundo a naturalidade



Fonte: Elaboração própria a partir de dados do IPEA (2011, p. 5).

O principal objetivo deste estudo foi responder a duas questões de pesquisa propostas na Agenda do Projeto Acomodação (Oushiro, 2019; Oushiro et al., 2023):

(i) As variáveis sociolinguísticas de diferentes naturezas (lexical, fonético-fonológica e morfossintática) perpassam os mesmos processos de adaptação linguística durante o contato do migrante com a comunidade anfitriã?

(ii) Como se dá o encaixamento social e linguístico das adaptações de variáveis sociolinguísticas de distintas naturezas (lexical, fonético-fonológica e morfossintática) identificadas na fala de migrantes em situação de contato dialetal?

Para tanto, analisaram-se, a partir de modelos estatísticos realizados na plataforma R (R Core, 2023), as variáveis lexicais Tangerina/Mexerica, Aipim/Mandioca, Quentinha/Marmita e Trabalho/Serviço; a variável fonético-fonológica (-r) em coda silábica e a variável morfossintática Negação sentencial (Neg). A escolha dessas variáveis levou em consideração as diferenças diaatópicas apresentadas pelo *Atlas Linguístico do Brasil* (Cardoso et al., 2014), a pesquisa de campo e as variáveis que têm sido foco no Projeto Acomodação.

Esta Tese está estruturada da seguinte forma. Na Seção 2, faz-se uma breve revisão de pesquisas que analisaram as adaptações linguísticas de migrantes em situação de contato dialetal, focalizando, principalmente, os estudos sobre o contato entre variedades do português brasileiro. Na Seção 3, descrevem-se: o método de abordagem dos participantes da amostra; o roteiro utilizado nas entrevistas sociolinguísticas; as características sociais da amostra; as variáveis sociais testadas e suas hipóteses; a maneira como os dados foram coletados, manipulados e tratados estatisticamente. Na Seção 4, apresentam-se os resultados das análises dos Itens lexicais nas quais se ponderou o uso de Mexerica, Mandioca, Marmita e Serviço como variantes típicas de São Paulo, em relação a Tangerina, Aipim, Quentinha e Trabalho, consideradas como variantes típicas da Bahia. Na Seção 5, reportam-se os resultados das análises do (-r) em coda, nas quais se consideraram as variantes Retroflexo e Tepe como típicas da comunidade anfitriã, em comparação com as variantes Velar e Glotal como típicas da Bahia. Na Seção 6, explicitam-se os resultados das análises da Negação sentencial, nas quais se observou a baixa frequência de uso de Neg2 e Neg3 como característico do falar paulista, em relação às altas frequências dessas variantes, consideradas como típicas da variedade baiana.

A Seção 7 sistematiza os resultados encontrados na amostra de fala dos baianos residentes na RMSP e propõe uma tipologia de processos de adaptação linguística. Nele, argumenta-se em favor do uso do termo *plasticidade dialetal* em substituição aos termos *acomodação* e *aquisição de segundo dialeto*, porque ele consegue não só abarcar todos os processos de adaptação identificados na fala dos migrantes, mas também focaliza as particularidades linguísticas e sociais dos dialetos em contato.

---

## 2

# O contato entre dialetos mutuamente inteligíveis

Nesta Seção, revisam-se os trabalhos sobre o contato entre dialetos mutuamente inteligíveis que embasaram o estudo das adaptações linguísticas dos migrantes baianos residentes na RMSP. A presente Seção está assim organizada: na Subseção 2.1, resenham-se pesquisas que propuseram generalizações acerca da fala de migrantes em situação de contato dialetal, mais especificamente os estudos de Trudgill (1986), Chambers (1992), Siegel (2010) e Oushiro (2024, no prelo[a],[b]). Na Subseção 2.2, realiza-se um levantamento sucinto de pesquisas que investigaram o contato entre variedades do português brasileiro. Essa abordagem considera o tipo de análise efetuada, seja com base em apenas uma variável sociolinguística ou a partir de múltiplas variáveis. Na Subseção 2.3, sintetizam-se as ideias desta Seção.

### 2.1 Base epistemológica para estudos sobre o contato entre dialetos mutuamente inteligíveis

A episteme sobre os efeitos do contato entre dialetos mutuamente inteligíveis deve-se especialmente ao trabalho seminal de Trudgill (1986), intitulado *Dialects in contact*. Nessa obra, o autor se debruça sobre pesquisas que tiveram como objetivo investigar as adaptações linguísticas resultantes do contato dialetal e apresenta generalizações acerca desse fenômeno linguístico-social, argumentando que as alterações na fala de migrantes em situação de contato estão relacionadas, sobretudo, à saliência social das variáveis e à acomodação linguística.

Sobre a saliência social das variáveis, Trudgill (1986) sugere que, no contato de migrantes com a comunidade anfitriã, as variáveis classificadas como marcadores e estereótipos, formas socialmente eminentes (Labov, 2008 [1972]), são mais suscetíveis a alterações. Isso ocorre porque essas variáveis estão acima da consciência do indivíduo em comparação com os indicadores. Em outras palavras, variáveis cujas variantes são percebidas pelos migrantes como diferenciadoras de categorias sociais, especialmente as locais (rural e urbano) e as regionais

(Nordeste e Sudeste, por exemplo), tendem a ser alvo de adaptações durante o contato com a comunidade anfitriã, talvez por carregarem algum estigma.

De acordo com [Trudgill \(1986\)](#), as alterações linguísticas podem ocorrer por meio do ajustamento das variáveis das quais os migrantes já têm controle, da incorporação de novas variantes e da substituição de formas que não são inteligíveis para a comunidade anfitriã. Com base nas análises que o levaram a fazer tais generalizações, [Trudgill \(1986\)](#) estabelece uma rota fixa de adaptação da fala no contato do inglês norte-americano com o inglês britânico, segundo a qual as alterações linguísticas acontecem primeiro no léxico, depois na morfologia e, por fim, na fonologia. [Trudgill \(1986\)](#) se fundamenta na *Teoria da Acomodação e Comunicação*, postulada pelos psicólogos sociais Howard Giles, Donald Taylor e Richard Bourhis, e propõe que elas são consequências da rotinização da acomodação linguística cotidiana feita pelo falante aos padrões da nova comunidade.

A Teoria da Acomodação e Comunicação surgiu na década de 1970, quando [Giles, Taylor e Bourhis \(1973\)](#) testaram a hipótese de que, numa situação de contato entre pessoas com algum domínio de uma segunda língua, maior esforço na acomodação do sotaque facilitaria a compreensão e aceitação pelos interlocutores. Para investigar essa questão, [Giles, Taylor e Bourhis \(1973\)](#) conduziram um estudo com 80 participantes ingleses-canadenses (43 mulheres e 37 homens), todos estudantes universitários proficientes em francês. Os pesquisadores dividiram os participantes em quatro grupos e os submeteram ao experimento em cinco etapas. Na primeira, os experimentadores pediram aos participantes que classificassem, a partir de questões de escala, o nível de proficiência do francês e do inglês deles mesmos. Depois, para cada grupo, mas individualmente com cada participante, eles reproduziram uma das versões dos áudios gravados para a coleta de dados: francês; francês e inglês misturados; inglês fluente, mas com sotaque franco-canadense; e inglês não fluente. Todos eles foram gravados pela mesma pessoa e com o mesmo conteúdo: a descrição de uma paisagem simples. Em seguida, [Giles, Taylor e Bourhis \(1973\)](#) solicitaram aos participantes que, à medida que fossem ouvindo a gravação, desenhassem a cena que estava sendo descrita. Na sequência, os pesquisadores apresentaram-lhes a imagem padrão da cena descrita e lhes pediram que gravassem a mesma descrição para a pessoa do áudio que tinham escutado. Por fim, eles aplicaram um questionário com escalas, cujas perguntas solicitavam que os participantes avaliassem e demonstrassem atitudes para a interlocução. Os resultados desse experimento, em linhas gerais, mostram que a adaptação ao sotaque do interlocutor tende a acontecer no momento em que o falante deseja uma aprovação social (convergência). Quando não há esse interesse (divergência), a adaptação tende a não ocorrer.

Na Psicologia Social, essa teoria passou a ser utilizada<sup>4</sup> em análises sobre as adaptações inconscientes a curto prazo nas interações face a face, principalmente quanto ao sotaque,

---

<sup>4</sup> Para ver o uso da Teoria da Acomodação em outras áreas, sugere-se a leitura de [Gasiorek \(2016\)](#).

gestos, expressões faciais, entre outros. [Trudgill \(1986\)](#), mesmo tendo reconhecido ruídos de ordem metodológica dessa teoria, passou a empregá-la em seus estudos sociolinguísticos, sugerindo que as alterações linguísticas na fala de migrantes em situação de contato resultam de sua convergência para com a comunidade anfitriã a longo prazo. Ou seja, para [Trudgill \(1986\)](#), a acomodação a longo prazo é efeito de sucessivas acomodações a curto prazo.

[Chambers \(1992\)](#), por outro lado, discorda de [Trudgill \(1986\)](#) e propõe que as alterações notadas na fala de migrantes devido ao contato dialetal são consequências da *aquisição de segundo dialeto*, aventando oito hipóteses para esse processo. Em seu estudo, [Chambers \(1992\)](#) analisou dados oriundos da fala de seis adolescentes canadenses que migraram com suas famílias para a Inglaterra entre 1983 e 1984. Tais ocorrências foram coletadas por esse mesmo pesquisador, em dois momentos, com intervalo de dois anos, a partir de entrevistas sociolinguísticas guiadas por perguntas sobre a caracterização do lugar de origem e da comunidade anfitriã, avaliação de sotaque, cartões com ilustrações e listas de palavras.

Com os resultados de sua pesquisa, [Chambers \(1992\)](#) aventou oito hipóteses empiricamente testáveis sobre a aquisição de segundo dialeto, a saber: (i) palavras diferentes empregadas para nomear o mesmo objeto em variedades distintas são substituídas rapidamente; (ii) a substituição de palavras ocorre rapidamente na primeira fase do contato; (iii) regras fonológicas simples são adquiridas mais rápido do que regras mais complexas; (iv) regras complexas e variantes fonéticas são adquiridas de maneiras diferentes entre quem migra mais cedo e quem migra mais tarde; (v) na primeira fase do contato, há variação de regras categóricas e de regras variáveis; (vi) inovações fonológicas podem se espelhar em outros contextos a partir da Difusão Lexical; (vii) eliminar regras antigas acontece mais rápido do que adquirir regras novas; (viii) palavras com ortografia distintas são rapidamente adquiridas em relação a palavras desconhecidas.

[Siegel \(2010\)](#) concorda com [Chambers \(1992\)](#) acerca da terminologia *aquisição de segundo dialeto* para nomear as adaptações linguísticas resultantes do contato dialetal. Com base em resultados quantitativos de estudos sociolinguísticos (17 no total), [Siegel \(2010\)](#) elenca ruídos no que diz respeito ao uso da Teoria da Acomodação em pesquisas que tratam do contato entre dialetos mutuamente inteligíveis. De acordo com [Siegel \(2010, p. 70–73\)](#): (i) nem sempre aquisição acontece em interações face a face, logo não se pode assumir que a acomodação a curto prazo precede a acomodação a longo prazo; (ii) a aquisição dialetal não depende da acomodação linguística; (iii) nem sempre as atitudes positivas favorecem a aquisição; (iv) nem sempre a aquisição tem a ver com os estereótipos.

Em substituição ao uso da Teoria da Acomodação em estudos sobre contato dialetal, [Siegel \(2010\)](#) propõe que a aquisição de segundo dialeto deve ser analisada tal qual a aquisição de segunda língua, cujo processo pode acontecer de maneira naturalística, isto é, sem a interferência do ensino formal, ou de modo educacional. Tanto [Chambers \(1992\)](#) quanto [Siegel \(2010\)](#)

referem-se à aquisição dialetal essencialmente quanto à substituição lexical e à adição de novos traços fonéticos. Siegel (2010), além de criticar o uso da Teoria da Acomodação nos estudos de contato dialetal, admoesta as metodologias utilizadas nos estudos sobre os quais ele se debruçou e recomenda ter um olhar mais atento para: (i) as diferenças linguísticas dos dialetos contactantes; (ii) o número de ocorrências das variáveis analisadas; (iii) a natureza das variáveis sociolinguísticas investigadas; (iv) a distinção entre variantes que não são fáceis de ser diferenciadas à oitiva; (v) os preditores Idade de migração, Tempo de residência, Atitudes dos falantes, Gênero, Rede social, Motivo da migração e Ocupação.

Levando em conta orientações metodológicas como essas, o Projeto Acomodação (Oushiro, 2017) tem se debruçado sistematicamente sobre a fala de migrantes alagoanos e paraibanos residentes na cidade de São Paulo (SP) e na Região Metropolitana de Campinas (SP). Antes de apresentar seus principais resultados, ressalta-se que, embora o projeto tenha esse nome, ele não desconsidera as críticas feitas por Chambers (1992) e Siegel (2010) para o uso da Teoria da Acomodação em estudos sobre contato entre dialetos mutuamente inteligíveis.

O Projeto Acomodação, realizado em duas fases, teve como objetivo analisar múltiplas variáveis de distintas natureza (a Altura das vogais médias pretônicas /e/ e /o/, medidas de F1 em Hertz; o (-r) em posição de coda silábica; a realização de /t, d/ seguido de [i]; (Neg); e a Concordância nominal (CN)), partindo da expectativa de que os migrantes tendem a se acomodar às variantes típicas da comunidade anfitriã. Na primeira fase, analisou-se, para fins exploratórios, uma amostra de fala que já havia sido previamente coletada para o Projeto Casadinho (Hora; Negrão, 2011) e para a dissertação de mestrado de Gomes da Silva (2014). Essa amostra é composta por 32 migrantes alagoanos e paraibanos residentes na capital paulista, estratificada em Gênero (masculino; feminino), Nível de escolaridade (até o Ensino Médio; Ensino Superior) e Faixa etária (20-34 anos; 35-59 anos e 60+ anos). Já na segunda fase, investigou-se uma amostra de fala de 40 migrantes alagoanos e paraibanos residentes na Região Metropolitana de Campinas (SP), balanceada em relação ao Gênero (masculino; feminino), à Idade de migração (até 19 anos; 20+ anos) e ao Tempo de residência (até 9 anos; 10+ anos). Oushiro (2017) observou os aspectos sociais da amostra e concluiu que os participantes são, em sua maioria, provenientes da zona rural; adultos (com idade ente 20-60 anos); e que estudaram apenas no Ensino Básico. Nessa segunda fase do projeto, diferentemente da primeira, nas entrevistas sociolinguísticas, aplicaram-se dois questionários, um socioeconômico a fim de coletar dados que pudessem indicar a classe social dos migrantes; e outro de rede social, hábitos e graus de identificação, cujas perguntas permitiram obter dados de suas atitudes em relação à comunidade de origem e à comunidade anfitriã, de sua rede social e de práticas sociais, como o tipo de lazer, por exemplo. É pertinente comentar que, para chegar a conclusões mais robustas, Oushiro (2017) contrastou seus resultados com amostras controle de dados de paraibanos (Hora, 1993) e alagoanos (Oliveira, 2017) não migrantes, de paulistas não migrantes (Mendes, 2012) e de

campineiros não migrantes (Mourão, 2018).

Oushiro (2020c), a partir da Amostra 1 do Projeto Acomodação, testou, por meio de análises de regressão linear e logística de efeitos mistos, feitas na plataforma R, a correlação das variáveis sociolinguísticas Altura das vogais médias pretônicas /e/ e /o/, (-r) em posição de coda, /t, d/ diante de [i], (Neg), (CN) com os preditores sociais Sexo/Gênero, Faixa etária, Escolaridade (Fundamental; Médio; Superior), Motivos de migração (família; qualidade de vida; trabalho), Idade de migração (até 17 anos; 18-24 anos; 30+ anos), Tempo em São Paulo (menos de 10 anos; 11-29 anos; 30+ anos) e Proporção de vida em SP (menos de 1/3; entre 1/3 e 2/3; mais de 2/3). Os resultados dessas análises mostram que (i) Sexo/Gênero não se correlacionou com nenhuma variável dependente; (ii) Faixa etária se correlacionou com (t, d), (Neg) e (CN), de maneira que migrantes mais novos favorecem o uso das variantes tipicamente paulistas; (iii) Escolaridade se correlacionou com (-r), (Neg) e (CN), de modo que os migrantes com mais escolaridade favorecem o uso das variantes morfossintáticas típicas de São Paulo e os migrantes com menos escolaridade favorecem a variante fonética-fonológica típica da comunidade anfitriã; (iv) Motivo da migração se correlacionou com (Neg) e (CN), sendo que os que migraram em busca de qualidade de vida favorecem as variantes típicas da comunidade anfitriã; (v) Tempo em São Paulo se correlacionou com (-r) e (Neg), de sorte que os migrantes que estão há mais tempo em São Paulo favorecem o uso do Tepe/Retroflexo e desfavorecem o uso de Neg1; (vi) Proporção de vida em São Paulo se correlacionou apenas com o (-r), cujo resultado mostra que os migrantes que estão há mais de dois terços da vida residindo em São Paulo favorecem o uso do Tepe/Retroflexo.

Numa análise mais refinada desses dados, Oushiro (2020a) mostra que, quando não se inclui o Participante como variável aleatória, os resultados para Sexo/Gênero se diferenciam dos que se viram nas análises de efeitos mistos, indicando haver correlação deste preditor com /o/ pretônico, (Neg) e (CN). Contudo, ao testar a interação dos preditores Escolaridade com Sexo/Gênero, Oushiro (2020a) notou correlação com /t, d/ diante de [i] e com (CN), de modo que as mulheres com nível de escolaridade maior tendem a empregar mais as formas padrão da comunidade anfitriã em comparação com os homens do mesmo nível de escolaridade. Partindo desse resultado, Oushiro (2020a) salienta a relevância de se considerar o papel do indivíduo nas análises sobre a fala de migrantes.

Com os dados da Amostra 2, Oushiro (2020b) testou a correlação de (-r), (t, d), (Neg) e (CN) com os preditores sociais Sexo/Gênero, Idade de migração, Tempo de residência e Grau de autoidentificação declarada com a comunidade de origem e Grau de autoidentificação paulista declarada. Os dados para análise destes dois últimos preditores resultam de duas perguntas do questionário de identidade: “de 0 a 10, quanto você se considera alagoano/parai-bano (a depender da proveniência do entrevistado)?” e “de 0 a 10, quanto você se considera paulista?”. Com essa metodologia, segundo Oushiro (2024, no prelo[b]), não se pretende quan-

tificar a identidade dos migrantes, já que é algo complexo de se aferir, mas é uma tentativa de analisar as atitudes dos migrantes como uma variável contínua. Os resultados das análises de regressão logística de efeitos mistos mostram a correlação significativa: do preditor Sexo/Gênero com (t, d) e a (CN), de maneira que as variantes consideradas típicas da comunidade anfitriã são favorecidas pelas mulheres em comparação com os homens; da Idade de migração com (-r) e (t, d), de modo que o Tepe/Retroflexo e a forma Palatal são favorecidos pelos migrantes que migraram mais cedo em relação aos que migraram mais tarde; do Tempo de residência apenas com o (-r), sendo que o Tepe/Retroflexo são favorecidos por quem está há mais tempo em Campinas (SP) em relação a quem está há menos tempo; do Grau de autoidentificação declarada em relação à comunidade de origem com (-r) e (Neg), de sorte que o Tepe/Retroflexo e Neg1 são favorecidos pelos migrantes que se deram notas mais baixas para o seu pertencimento ao lugar de origem em relação àqueles que se deram notas mais altas; do Grau de autoidentificação declarada em relação à comunidade anfitriã com (-r), de forma que o Tepe/Retroflexo são favorecidos pelos migrantes que se autoavaliaram com notas mais altas para o seu pertencimento à comunidade anfitriã em comparação com aqueles que se deram notas mais baixas.

Em Oushiro (2021), a autora analisou o emprego de (-r) e (-s) em posição de coda numa amostra de fala de 24 alagoanos e paraibanos residentes em Campinas (SP), composta por 9 mulheres e 15 homens (todos com baixo nível de escolaridade e provenientes da zona rural), estratificada em Idade de chegada (até 19 anos; depois de 20 anos) e Tempo de residência (até 9 anos; 10+ anos). Os resultados de Oushiro (2021) foram obtidos a partir de análises de regressão logística de efeitos mistos, nas quais se testou a correlação das variáveis respostas com os preditores estratificadores da amostra e também com Estilo (modelo arbóreo de Labov (2001)), Contexto fonológico precedente (anterior; central; posterior), Contexto fonológico seguinte (coronal; dorsal; labial; pausa) e Posição (medial; final). Nas análises do (-r), Oushiro (2021) mostrou que o Retroflexo é desfavorecido pelas mulheres que migraram mais cedo, em contextos de entrevista e conversação, e quando o /r/ está na posição medial. Por outro lado, é favorecido por quem está há mais tempo residindo na comunidade anfitriã, em contexto de estilo *soapbox*, ou seja, quando o participante desenvolve um assunto do qual ele demonstra ter respaldo para falar, em contexto linguístico em que o /r/ é seguido de consoante coronal ou dorsal. Na análise de (-s) em coda, Oushiro (2021) verificou correlação apenas com o preditor Estilo, de modo que a conversação desfavorece o uso da variante típica da comunidade anfitriã. Oushiro (2021) evidencia o resultado do uso de /r/ nos contextos de estilo e argumenta que o fato de os migrantes utilizarem a variante de sua gramática nativa em contextos não monitorados, como em entrevistas e conversas informais, enquanto usam a variante da comunidade anfitriã em contextos monitorados, pode estar relacionado: (i) às particularidades das normas local e nacional; (ii) às distintas orientações que homens e mulheres têm acerca delas; (iii) ao uso da língua como forma de ratificação das diferenças dialetais; (iv) e à convergência do migrante

para a variedade da comunidade anfitriã.

Em Oushiro (2024, no prelo[a],[b]), a autora coadunou os resultados obtidos das duas amostras do Projeto Acomodação com os dados encontrados em diversas pesquisas realizadas em variedades do português, como os trabalhos de Bortoni-Ricardo (2011 [1985]), Marques (2006), Chacon (2012) e Santana (2018), por exemplo, e faz generalizações sobre as adaptações linguísticas da fala de migrantes, revelando que: (i) os migrantes tendem a escolher as variantes da comunidade de origem em contexto de leitura; (ii) a Idade de migração tende a se correlacionar com as variáveis fonéticas em comparação com as morfossintáticas; (iii) o Tempo de residência tem apresentado resultados pouco consistentes; (iv) as variantes rurais tendem a ser abandonadas pelas mulheres do que pelos homens. Ademais, Oushiro (2024, no prelo[a],[b]) chama atenção para questões metodológicas, sugerindo que as investigações sobre o contato entre dialetos mutuamente inteligíveis precisam analisar: (i) a natureza da variável e sua distribuição social na comunidade de origem; (ii) os significados sociais das variantes na comunidade anfitriã; (iii) as variáveis confundidoras<sup>5</sup> e não ortogonais<sup>6</sup>; e (iv) a força das correlações para não fazer generalizações ruidosas.

## 2.2 Outros estudos sobre o contato entre variedades do português brasileiro

Nas pesquisas sobre as quais Oushiro (2024, no prelo[a],[b]) se debruçou, de modo particular aquelas desenvolvidas com base no contato entre variedades brasileiras, examinaram-se essencialmente variáveis dependentes assinaladas pela Sociolinguística e pela Dialectologia, a partir de dados obtidos de projetos como o NURC (Castilho, 1990) e o ALiB (Cardoso et al., 2014), por exemplo, como diferenciadoras de dialetos regionais e/ou locais nos distintos níveis da língua: lexical, fonético-fonológico, morfossintático etc. Muitos desses trabalhos se concentram na análise de uma única variável sociolinguística, principalmente de naturezas fonético-fonológica e morfossintática. Em contrapartida, poucos são os que observaram o comportamento de múltiplas variáveis. Inexistentes, pelo menos até o momento de produção desta Tese, são as pesquisas, no âmbito da Sociolinguística Variacionista, acerca da variação lexical, à exceção de estudos produzidos na área da Dialectologia Pluridimensional, como se verá mais adiante.

Nesta Subseção, retomam-se pesquisas cujos resultados acenderam luz para conjecturar hipóteses sobre a adaptação linguística dos migrantes baianos residentes na RMSP. Essa seleção não está organizada em ordem cronológica, mas com base no foco de suas análises. Dessa maneira, inicialmente, apresenta-se o estudo de Alves (1979) sobre atitudes linguísticas. Em seguida, os estudos que investigaram apenas variáveis sociolinguísticas de natureza

<sup>5</sup> Variáveis que falseiam as correlações, uma vez que elas influenciam tanto a variável dependente quanto a variável independente (ver Gries (2009), Levshina (2015) e Oushiro (2022)).

<sup>6</sup> Variáveis que não apresentam todas as combinações possíveis entre as células (ver Gries (2009), Levshina (2015) e Oushiro (2022)).

fonético-fonológicas, como as pesquisas de Santana (2018) e de Amorim e Costa (2019) sobre as Vogais médias pretônicas /e/ e /o/; de Martins (2008) e de Santana (2021) acerca do /t, d/ diante de [i]; o trabalho de Chacon (2012) quanto à realização do /s/ seguido de /t/ e /d/; as análises de Possatti (2020) e de Barbosa (2022) sobre o /s/ em posição de coda; de Bieler da Silva (2015), de Hora e Wetzels (2010) e de Oliveira (2020) a respeito do (-r) em posição de coda. Em seguida, resenham-se os trabalhos que investigaram somente variáveis morfossintáticas, como o de Guedes (2019 [2017]) e o de Guirelli (2018) sobre o uso do Artigo antes de possessivos; a pesquisa de Souza (2019) a respeito do Imperativo com morfologia de indicativo. Por fim, apresentam-se as pesquisas Adant (1989), Bortoni-Ricardo (2011 [1985]), Figueiredo (2014), Meurer (2022), Oliveira (2023, inédito) e Santana (2023) que observaram múltiplas variáveis.

Alves (1979) analisou, a partir de dados coletados por meio de um questionário aplicado a 116 migrantes pernambucanos e baianos em São Paulo, no final da década de 1970, as suas atitudes linguísticas acerca da fala de sua comunidade de origem e da comunidade anfitriã. A pesquisadora partiu da hipótese de que esses migrantes, estratificados em níveis sociocultural alto e baixo, apresentariam diferentes atitudes em relação a cada variedade, a saber, os falantes de um nível sociocultural baixo teriam atitudes linguísticas positivas em relação ao dialeto paulista e os de nível sociocultural elevado prestigiariam o seu dialeto nativo, o que de fato se verificou, segundo seus dados.

Inspirada na análise de rede social feita por Milroy (1987 [1980]) e Bortoni-Ricardo (2011 [1985]), cuja metodologia merece destaque já que se diferencia da abordagem tradicional da Sociolinguística, Santana (2018) investigou a acomodação linguística de migrantes sergipanos em São Paulo (SP) quanto à Altura das vogais médias pretônicas /e/ e /o/ a partir das medidas de F1 (em Hz) extraídas com o auxílio do Praat. Em relação à análise de rede social, segundo Bortoni-Ricardo (2011 [1985]), cujo trabalho será resenhado mais adiante, é um método de abordagem usado que consiste em mapear as pessoas com as quais os falantes convivem a fim de verificar se o uso de determinadas variantes se correlaciona com a densidade de comunicação dos participantes com seus pares.

De acordo com Milroy (1987 [1980]), as redes podem ser *multiplexas*, nas quais as relações sociais são fortes e permitem suportar as pressões de padrões externos; ou *uniplexas*, em que os laços são mais frouxos e, por consequência, a fala dos indivíduos é mais suscetível a mudanças. Esses laços podem ser classificados como de primeira ordem (entre aqueles que mantêm contato direto com o ego), ou de segunda ordem (entre aqueles que têm contato indireto) (Milroy, 1987 [1980]). A depender do tipo de laço que os indivíduos tiverem, as redes podem ser denominadas *redes de trocas*, em que há trocas mútuas entre o ego e as pessoas com as quais ele mais interage como, por exemplo, a família e os amigos mais íntimos, ou *redes interativas*, nas quais o ego interage com frequência, mas não confia seus bens simbólicos, como o vínculo entre patrão e empregado (Bortoni-Ricardo, 2011 [1985]).

Assim, para compor a amostra de fala de sua pesquisa, Santana (2018) escolheu dois âncoras<sup>7</sup> que serviram como pontos de partida para fazer o contato com outros informantes das redes observadas (Santana, 2018, p. 32). O critério utilizado por Santana (2018), antes do mapeamento das redes, foi o de que os informantes de uma rede não poderiam manter contato com os da outra. Com isso, a pesquisadora pôde avaliar se havia diferenças ou semelhanças entre os agrupamentos observados quanto à forma de realização das vogais médias pretônicas.

A rede 1 teve como ponto de partida uma mulher, cuja relação com a pesquisadora era de pouca aproximação. A rede 2, por sua vez, foi iniciada a partir de um homem bastante próximo a ela. Ao final da entrevista com essas duas pessoas, Santana (2018) perguntou “quais eram as dez pessoas com as quais elas mais mantinham contato diariamente, incluindo os membros da família”. Das pessoas indicadas por eles, a autora entrevistou apenas as que também eram sergipanas, para as quais igualmente se fez, ao final da entrevista, a mesma pergunta de indicação. Os que não eram sergipanos, apesar de não terem sido entrevistados, foram listados com o objetivo de, posteriormente, estabelecer um índice para aferir a “força” de relação dos migrantes sergipanos entre si e a relação deles com os indivíduos da comunidade anfitriã e/ou de outras localidades. O método aplicado por Santana (2018) resultou em uma amostra de fala de 27 participantes, distribuídos em duas redes sociais, a primeira formada por 16 informantes e a segunda por 11. Em sua pesquisa, Santana (2018), baseando-se em Hoffman e Walker (2010) e Oushiro (2017), elaborou um questionário o qual ela denominou de “Questionário para a obtenção do índice de vínculo com Sergipe” (ver Santana (2018, p. 141)) a fim de coletar dados sobre o *modus vivendi* dos migrantes sergipanos em São Paulo (SP). Em tal questionário, constam perguntas como “Numa escala de 0 a 10, quanto você se considera sergipano?”, “A maioria dos seus amigos é sergipana?” e outras.

Santana (2018) testou a correlação da Altura das vogais médias pretônicas com as variáveis linguísticas Contexto fônico precedente (dental-alveolar; labial; velar; palatal; sibilante; vibrante), Contexto fônico seguinte (dental-alveolar; labial; velar; palatal; sibilante; vibrante), Distância em relação à vogal tônica (1 a 5 sílabas) e Tipo de sílaba pretônica (CV; CCV; CCVs; CVr; CVs); e com as variáveis sociais Rede social (aberta; fechada), Índice de vínculo com Sergipe (variável contínua), Índice de integração à rede (variável contínua), Proporção da vida em São Paulo (variável contínua), Sexo/Gênero (masculino; feminino), Faixa etária (-50 e +50) e Escolaridade (Fundamental I; Fundamental II; Ensino Médio). As análises de regressão linear de efeitos mistos, nas quais se ponderaram, além das variáveis linguísticas e sociais, o Participante e o Item lexical como variáveis aleatórias, foram realizadas na plataforma R.

Os resultados desse estudo indicam haver bastante variação individual e, por causa disso, segundo Santana (2018), não houve diferença significativa entre os valores de F1 da

<sup>7</sup> ‘Pessoas cujo comportamento o pesquisador deseja examinar’ (Mitchell, 1969, p. 13 *apud* Bortoni-Ricardo, 2011 [1985], p. 147).

amostra dos migrantes e da dos sergipanos não migrantes, utilizada como grupo controle. De acordo com Santana (2018), a maioria dos migrantes (19) se adaptou à fala paulistana quanto à realização do [e] e apenas uma minoria (6) quanto à vogal pretônica [o] e os demais não se adaptaram a nenhum desses dois fonemas. Com base nos modelos de regressão linear, Santana (2018) verificou a correlação tanto de /e/ (0,261 *logodds*,  $p = 0,001$ ) quanto de /o/ (0,329 *logodds*,  $p = 0,001$ ) com o F1 da Vogal da sílaba seguinte. No tocante às variáveis sociais, Santana (2018) viu que os valores de F1 de /o/ se correlacionaram apenas com a Proporção de vida em São Paulo (13,841 *logodds*,  $p = 0,049$ ), de maneira que os migrantes sergipanos que apresentam as menores proporções de vida em São Paulo tendem a abaixar menos essa vogal; e que os valores de F1 de /e/ se correlacionaram apenas com a Idade de migração (0,343 *logodds*,  $p = 0,009$ ), de modo que quem migrou mais cedo tendeu a empregar mais a variante paulistana.

Amorim e Costa (2019) examinaram a fala de seis migrantes gaúchos residentes em Teresina (PI) a fim de saber se eles estariam se adaptando<sup>8</sup> ao falar teresinense quanto à abertura das Vogais médias pretônicas /e/ e /o/. A amostra utilizada foi estratificada de acordo com o Tempo de residência (entre 2-4 anos; 5-10 anos; e de +10 anos). O resultado dessa análise revela que os gaúchos residentes em Teresina (PI) tendem a não se adaptar ao uso das variantes da comunidade anfitriã [ɛ] e [ɔ], visto que o percentual de uso foi de 1% e 3%, respectivamente.

Martins (2008) pesquisou a Palatalização das oclusivas dentais /t, d/ diante de [i] na fala de sete migrantes paraibanos radicados no Rio de Janeiro (RJ), partindo da expectativa de que eles, ao manterem contato com os cariocas, se adaptariam linguisticamente quanto ao uso das formas Palatalizadas. A partir de análises estatísticas feitas no Goldvarb X, a autora testou a correlação de /t, d/ com as variáveis linguísticas Sonoridade (surda; sonora), Posição em relação à sílaba tônica (pretônica; tônica; postônica), Posição da sílaba na palavra (inicial; medial; final; clítico), Modo de articulação subsequente (oclusivas; fricativas; nasais; líquidas; glides/vogais; outros; nenhum) e com as variáveis sociais Faixa etária (até 25 anos; de 26–43 anos; +43), Sexo (masculino; feminino) e Tempo de permanência no Rio de Janeiro (até 10 anos; de 11–25 anos; +25). O resultado geral desse estudo mostra que, do total de 1.430 dados de /t, d/ diante de [i], 64% foram das variantes Alvéolo-palatais e 36% de Oclusivas dentais. Os valores das análises multivariadas mostram que a Palatalização tende a ser realizada na fala dos migrantes paraibanos com até 25 anos, que residem no Rio de Janeiro (RJ) há 10 anos (P.R. .93), em sílabas com a consoante surda (P.R. .57), seguida de glides/vogais (P.R. .66) e de fricativas (P.R. .56), em sílabas postônicas (P.R. .61), na posição medial (P.R. .62) e final (P.R. .60) de palavras polissílabas (P.R. .62) e dissílabas (P.R. .55).

Santana (2021), ao perceber que em cidades de Sergipe, comparado a São Paulo (SP), não há a predominância da Palatalização das oclusivas dentais /t, d/ diante de [i], inves-

<sup>8</sup> Os autores empregam o termo “acomodar”.

tigou se os migrantes sergipanos domiciliados em São Paulo (SP) passam a palatalizá-las ao manterem contato com os paulistas. A partir de uma amostra de fala de 27 migrantes sergipanos (Santana, 2018), Santana (2021) testou a correlação dessa variável sociolinguística com as variáveis linguísticas Contexto fonológico precedente (vogal baixa; vogal não baixa; fricativa alveolar; rótico; nasal; início de palavra), Contexto fonológico posterior (vogal não posterior; vogal posterior; oclusiva; fricativa; nasal; líquida; fim de palavra), Sonoridade (surda; sonora), Tonicidade (pretônica inicial; pretônica não inicial; tônica; postônica) e Status da vogal alta (fonológica; fonética). A autora testou a correlação também com as variáveis sociais Rede social (aberta; fechada), Idade de migração (variável contínua), Escolaridade (Fundamental I; Fundamental II; Ensino Médio) e Sexo/Gênero (homens; mulheres). Os resultados desse estudo, obtidos nas análises descritivas e inferenciais feitas na plataforma R, mostram que, de 2.153 ocorrências desse fenômeno variável, 61,36% foram das formas Palatalizadas e 38,64% das variantes Não palatalizadas. Ao comparar com os valores da Palatalização de Aracaju (SE) (21,7%), de Lagarto (SE) (6,6%) e de Itabaiana (SE) (7,7%), a autora assume que o aumento das taxas de Palatalização dos migrantes sergipanos em São Paulo (SP) é um efeito do contato com a comunidade anfitriã. Nas análises multivariadas de efeitos mistos com inclusão do Participante e do Item lexical como efeito aleatório, observou-se que a Palatalização tende a acontecer diante da consoante surda (1,352 *logodds*,  $p = 0,001$ ), em sílabas tônicas (0,930 *logodds*,  $p = 0,024$ ), precedidas por fricativas alveolares (2,952 *logodds*,  $p = 0,001$ ) e nasais (0,767 *logodds*,  $p = 0,018$ ), quando não está no fim de palavras (-2,403 *logodds*,  $p = 0,001$ ) e quando a vogal-gatilho não é fonológica (-1,429 *logodds*,  $p = 0,001$ ). Acerca das variáveis sociais, os valores encontrados mostram que a Palatalização é favorecida por pessoas que migraram mais novas e que possuem o Ensino Fundamental II (0,376 *logodds*,  $p = 0,015$ ). Segundo Santana (2021), esse resultado tem relação com o fato de esses migrantes terem estudado em São Paulo (SP) e não em Sergipe.

Chacon (2012), utilizando um *corpus* com os dados de dez migrantes paulistas residentes em João Pessoa (PB), pesquisou a possível adaptação<sup>9</sup> linguística deles quanto à Palatalização de /s/ diante de /t/ e /d/, como em “pasta” e “desde”. A autora testou a correlação dessa variável sociolinguística com as variáveis sociais Faixa etária (I, 19-25 anos; II, +30 anos), Tempo de exposição (1-3 anos; 3-5 anos; +8 anos), Naturalidade dos pais (Sudeste; Nordeste; Nordeste-Sudeste), com a variável Atitudinal, analisada de maneira qualitativa, e com a variável Contextual de estilo (leitura; entrevista). O resultado geral dessa análise indica que, do total de 730 de ocorrências de /s/ diante de /t/ e /d/, 34,8% foram da variante Palatalizada e 65,2% da Não palatalizada. As análises multivariadas realizadas no Goldvarb X sugeriram que a Palatalização tende a ocorrer mais na fala dos migrantes paulistas com idade entre 19 e 25 anos (P.R

<sup>9</sup> A autora utiliza o termo “acomodação”.

.67), filhos de nordestinos e sudestinos (P.R. .62), que residem há mais de 8 anos (P.R. .70) em João Pessoa (PB) e em contexto de entrevista (P.R. .58). Quanto às atitudes, segundo a autora, as positivas convergem para a adaptação linguística dos migrantes na comunidade anfitriã.

Possatti (2020) investigou a Não palatalização do /s/ em coda, como em “mesmo” e “mas”, de migrantes cariocas residentes em João Pessoa (PB) a partir de uma amostra de fala de 16 participantes, estratificada em Sexo (masculino; feminino) e Idade (18–29 anos; +30 anos). A fim de alcançar padrões acerca dos efeitos desse contato dialetal, o autor testou, com base em análises multivariadas feitas no Goldvarb X, a correlação do /s/ em coda com as variáveis sociais estratificadoras da amostra e também com as variáveis Tempo de exposição (1 a 5 anos; +5 anos) e Motivo da migração (vinda espontânea; vinda obrigatória). Além dessas variáveis, analisaram-se o Contexto fonológico anterior (vogais anteriores; vogal central; vogal posterior) e o Contexto fonológico posterior (pausa; consoante). O resultado global desse estudo revela que os cariocas residentes em João Pessoa (PB) aplicaram 24,8% da variante Não palatalizada e 75,2% da forma Palatalizada. Os resultados das análises multivariadas indicam que a Não palatalização na fala dos migrantes cariocas residentes em João Pessoa (PB) é favorecida pelas mulheres (P.R. .68), com idade acima de 30 anos (P.R. .56), que residem há mais de 5 anos (P.R. .74), em contextos em que o /s/ é seguido de pausa (P.R. .68).

Barbosa (2022), com base na Amostra 2 do Projeto Acomodação (Oushiro, 2017), estratificada em Gênero (masculino; feminino), Idade de migração (até 19 anos; +20 anos) e Tempo de residência (até 9 anos; +10 anos), averiguou o emprego do /s/ em posição de coda a partir de quatro variantes: a Alveolar, a Palatalizada, a Aspirada e o Zero fonético. Assumindo que as formas Alveolar e Zero fonético são característicos do falar campineiro, Barbosa (2022) testou, com o auxílio da plataforma R, a correlação dessa variável com, além dos preditores sociais que estratificam a amostra, o Índice socioeconômico (variável contínua), o Grau de paulistidade (variável contínua), o Grau de nordestinidade (variável contínua), o Índice de rede do Nordeste (variável contínua), o Índice de hábitos (variável contínua). Barbosa (2022) também adicionou aos modelos estatísticos os preditores linguísticos Tonicidade (átona; tônica), Posição da sílaba (medial; final), Contexto fonológico precedente (aberta; meia aberta; meio fechada; fechada), Contexto fonológico seguinte (pausa; surdo; sonoro) e Classe morfológica (lexical; gramatical) e o preditor contextual Estilo (conversa; lista de palavras).

A distribuição das variantes na amostra revela que, de 6.512 observações, 76% foram da forma Alveolar, 11,6% da Aspirada, 10,1% de Zero fonético e 2,3% da variante Aspirada. Nas análises binárias, em que se considerou a variação entre a variante Palatal e a Alveolar, Barbosa (2022) notou que, de 2.263 ocorrências de /s/ em coda, 68,1% foram da forma Alveolar e 31,9% da Palatal. As análises multivariadas de regressão logística de efeitos mistos, com a inclusão do Item lexical e do Participante como efeito aleatório, evidenciaram que a Palatalização do /s/ em coda tende a ocorrer na fala dos migrantes alagoanos e paraibanos que chegaram

a Campinas (SP) com idade mais avançada (0,062 *logodds*,  $p= 0,05$ ), que residem há pouco tempo (-0,047 *logodds*,  $p= 0,05$ ), em contexto medial (2,240 *logodds*,  $p= 0,001$ ). Lendo esses resultados inversamente, verifica-se que os alagoanos e paraibanos que migraram mais cedo e que estão há mais tempo residindo em Campinas (SP) tendem a empregar a variante típica da comunidade anfitriã. Na comparação entre as variantes Zero fonético e Aspirada/Alveolar/Palatal, observou-se que, de 4.234 dados de /s/ em coda, 12,1% foram de Zero fonético e 87,9% das demais formas. Os resultados dos testes de regressão logística sugerem que o Zero fonético tende a ser desfavorecido pelas mulheres (-1,389 *logodds*,  $p= 0,001$ ) e pela posição medial (-2,799 *logodds*,  $p= 0,001$ ) e favorecido por quem migrou mais tarde (0,071 *logodds*,  $p= 0,01$ ) e pelo contexto fonológico seguinte cuja consoante é sonora (1,687 *logodds*,  $p= 0,001$ ). Ou seja, são os homens que migraram mais cedo que se adaptaram à variante da comunidade anfitriã. Na análise em que se ponderou a aplicação das formas Alveolar/Palatal e da Aspirada, a autora observou que, de 1.089 dados, 12,6% foram desta variante e 87,4% daquela. Os coeficientes obtidos dos testes de regressão logística indicaram não haver nenhuma correlação com os preditores analisados.

Hora e Wetzels (2010) observaram a fala de quatro migrantes paraibanos residentes em São Paulo (SP) para testar a hipótese de que, numa situação de fala mais monitorada, eles empregariam o Tepe, a variante de prestígio entre os paulistanos, no lugar da forma Aspirada, variante predominante na Paraíba. Além disso, os autores questionaram qual das propostas de estilo daria conta de explicar os seus resultados, se a de Labov (2006 [1966]), que relaciona o uso de variantes de prestígio ao contexto mais formal, se a de Bell (1984), que leva em consideração a *audience design*, ou se a de Eckert (2005), que atribui o uso de determinadas variantes à construção de uma *persona*. O resultado, contudo, foi contrário à hipótese aventada, isto é, mesmo em contextos de fala monitorada, os paraibanos não empregaram o Tepe. Hora e Wetzels (2010) concluem que a proposta de Labov (2006 [1966]) não daria conta desse resultado. Com base nas avaliações e atitudes dos paraibanos residentes em São Paulo (SP) sobre os sotaques paraibano e paulistano, os autores afirmam que apenas os falantes mais escolarizados têm consciência das distinções segmentais dos róticos entre a sua comunidade de origem e a comunidade anfitriã. Hora e Wetzels (2010) sugerem que as concepções de Estilo de Bell (1984) e de Eckert (2005) poderiam ser as mais indicadas para explicar esse resultado.

Bieler da Silva (2015) observou o uso do (-r) em posição de coda numa amostra de fala de 36 itanhanduenses migrantes e não migrantes, com base na hipótese de que o uso do Tepe e do Fricativo pelos itanhanduenses estaria relacionado a questões identitárias e ao contato dialetal. Para encontrar padrões do uso do (-r) em coda nessa amostra, Bieler da Silva (2015) testou, a partir de análises multivariadas efetuadas no Goldvarb X, a correlação dessa variável sociolinguística com as variáveis sociais Sexo/Gênero (masculino; feminino), Faixa etária (I, 18-30 anos; II, 31-45 anos; III, +46), Grau de identificação (- identificação; + iden-

tificação), Possuir parentes fora (possui; não possui), Tópico tratado (vida pessoal; opinião; cidade; produção/percepção linguística), Tempo fora de Itanhandu (nunca saiu; até 20% fora; 30% ou mais fora), com a variável contextual Estilo (leituras; entrevista) e com as variáveis linguísticas Classe morfológica da palavra (verbo; adjetivo/numeral; substantivo; conjunção/-preposição), Contexto fônico anterior (posterior; anterior; central), Contexto fônico seguinte (velar; bilabial/labiodental; alveolar/pós-alveolar; pausa), Tonicidade (tônica; átona), Posição (meio; fim), Frequência da palavra (até 39 vezes; de 40 a 180 vezes; 181 vezes ou mais). O resultado global dessa pesquisa mostrou que, do total de 10.916 observações, 90,8% foram da variante Retroflexa, 8,4% de Fricativo e 0,8% de Tepe. Os pesos relativos da rodada selecionada evidenciaram que o Tepe e o Fricativo tendem a ser mais usados por homens (P.R. .55), com idade igual ou superior a 46 anos (P.R. .76) e por jovens entre 18 e 30 anos (P.R. .57), que estão há 30% ou mais tempo fora de Itanhandu (MG) (P.R. .81) ou que estão há 29% do tempo fora (P.R. .69), em verbos (P.R. .57) ou adjetivo/numeral (P.R. .54), quando o /r/ é precedido por vogal posterior (P.R. .53) e tem como contexto seguinte consoante velar (P.R. .55) e bilabial/labiodental (P.R. .51), quando empregados nos tópicos produção/percepção linguística (P.R. .56) e opinião (P.R. .54).

Oliveira (2020) analisou o emprego do (-r) em posição de coda silábica na fala de 12 migrantes baianos residentes em Bauru (SP). A expectativa para esse estudo foi a de que tais migrantes, ao manterem contato com os bauruenses, usariam o /r/ Retroflexo, variante típica da comunidade anfitriã, no lugar da forma Aspirada, variante predominante em sua comunidade de origem. A amostra pesquisada está estratificada em Sexo/Gênero (homem; mulher) e Escolaridade (Ensino Médio; Ensino Superior). Além dessas variáveis sociais, Oliveira (2020) testou a correlação do (-r) em coda com a Idade do participante (faixa I, 22-29 anos; II, 38-44 anos; e III, 52-59 anos), o Tempo de residência (faixa I, até 10 anos; II, 11-20 anos; III, +20 anos), a Idade de chegada (faixa I, -13 anos; II, 14-26 anos; III, 27-47 anos; IV, +48 anos), o Contato com baianos (esporádico; frequente), o contexto de Estilo (-monitorado; +monitorado), as Atitudes em relação à comunidade anfitriã (atitude negativa; atitude positiva) e o Informante. Ele também analisou as variáveis linguísticas Contexto fônico precedente (vogal anterior; vogal central; vogal posterior), Contexto fônico seguinte (pausa; consoante oclusiva; consoante nasal; consoante africada; consoante lateral) e a Posição do /r/ na palavra (inicial; final).

O resultado geral dessa pesquisa apresenta que, do total de 503 ocorrências de (-r) em coda, 88% foram da forma Aspirada e 12% de Retroflexo. De acordo com os resultados das análises univariadas, feitas na plataforma R, o Retroflexo tendeu a ser mais usado por migrantes com idade entre 38 e 44 anos ( $\chi^2 = 126,77$  (2),  $p = 0,001$ ), que migraram com idade entre 14 e 26 anos ( $\chi^2 = 35,209$  (2),  $p = 0,0001$ ), que residem em Bauru (SP) há mais de 20 anos ( $\chi^2 = 48,887$  (2),  $p = 0,001$ ), que estudaram até o Ensino Médio ( $\chi^2 = 9,0701$  (1),  $p = 0,002$ ), que têm contato frequente com outros baianos ( $\chi^2 = 33,235$  (1),  $p = 0,001$ ), que têm atitudes positivas

para com a comunidade anfitriã ( $\chi^2 = 49,336$  (1),  $p = 0,001$ ), em estilo de conversa ( $\chi^2 = 35,209$  (2),  $p = 0,001$ ).

Guedes (2019 [2017])<sup>10</sup> pesquisou o uso do Artigo definido antes de possessivos, como em “A minha mãe deixa” e “Ø Minha mãe deixa”, a partir de uma amostra de fala de 32 participantes, sendo 12 paraibanos não migrantes, 8 paraibanos residentes em São Paulo (SP) da Amostra 1 do Projeto Acomodação (Oushiro, 2017) e 12 paulistanos. O propósito de Guedes (2019 [2017]) foi testar se os migrantes paraibanos residentes em São Paulo (SP) empregariam mais o determinante antes de pronomes possessivos, conforme o padrão da comunidade anfitriã. Para alcançar seus objetivos, Guedes (2019 [2017]) realizou, na plataforma R, modelos de regressão logística de efeitos mistos, com inclusão do Item lexical e do Participante como variáveis aleatórias, nos quais se testou a correlação da variável dependente com as variáveis linguísticas Pessoa do pronome possessivo (1<sup>a</sup>; 2<sup>a</sup>; 3<sup>a</sup> pessoa), Número do possessivo (singular; plural), Gênero do possessivo (masculino; feminino), Tipo de preposição (ausência de preposição; preposição que contrai; preposição que não contrai), Natureza semântica do núcleo do sintagma nominal (parte do corpo; parente; humano não parente; objeto próprio; objeto não próprio; abstração não única; abstração única) e Função sintática do SN (sujeito; tópico; adjunto adverbial; genitivo; objeto direto; objeto indireto; predicativo; especificidade do núcleo do SN), e com as variáveis sociais Faixa etária (I, 15-34 anos; II, 35-49 anos; III, 50+), Sexo/Gênero (masculino; feminino), Nível de escolaridade (Até o Ensino Médio; Ensino Superior), Idade de migração (até 19 anos; 20+ anos) e Tempo de permanência em SP (até 10 anos; entre 11-20 anos; 21+ anos). Os resultados gerais desse estudo revelaram que há diferença significativa entre as taxas de uso do Artigo definido diante de possessivo entre a amostra de fala dos migrantes paraibanos residentes em São Paulo (SP) (51%) e a amostra composta por paraibanos não migrantes (42%), mas não há diferença significativa entre a amostra dos migrantes e a dos paulistanos (54%). Os resultados das análises multivariadas de regressão logística sugeriram que o emprego do Artigo definido diante de possessivos na fala de migrantes paraibanos residentes em São Paulo (SP) é favorecido por preposição que não contrai (3,53 *logodds*,  $p = 0,001$ ) e desfavorecido por sintagmas que exercem a função de objeto direto (-2,52 *logodds*,  $p = 0,001$ ), predicativo (-2,28 *logodds*,  $p = 0,001$ ), sujeito (-1,43 *logodds*,  $p = 0,001$ ), tópico (-1,24 *logodds*,  $p = 0,05$ ) e por pronomes do gênero masculino (-0,77 *logodds*,  $p = 0,001$ ). Não houve correlação com nenhuma variável social, mas Guedes (2019 [2017]) notou que os migrantes paraibanos que têm paulistas em suas redes sociais tenderam a utilizar as mesmas regras abstratas da comunidade anfitriã em comparação com aqueles participantes que não têm paulistas em suas redes.

Guirelli (2018), guiando-se pelo trabalho de Guedes (2019 [2017]), analisou uma

<sup>10</sup> Este artigo resulta da pesquisa de Qualificação de Área apresentada ao IEL/UNICAMP em 2017, sob a orientação da Profa. Dra. Livia Oushiro.

amostra de fala composta por 40 participantes, sendo 16 migrantes paraibanos que integram a Amostra 2 do Projeto Acomodação (Oushiro, 2017), 12 paraibanos não migrantes e 12 paulistas, com o objetivo de averiguar se o uso do Artigo definido diante de possessivos, típico do falar paulista, seria utilizado pelos migrantes paraibanos ao manterem contato com a comunidade anfitriã. Para chegar aos seus resultados, Guirelli (2018) testou a correlação da variável dependente com as variáveis linguísticas Pessoa do pronome possessivo (1<sup>a</sup>; 2<sup>a</sup>; 3<sup>a</sup> pessoa), Número do pronome possessivo (plural; singular), Gênero do pronome possessivo (feminino; masculino), Tipo de preposição (ausência de preposição; preposição que contrai; preposição que não contrai), Natureza semântica do núcleo do sintagma nominal (parte do corpo; parente; humano não parente; objeto próprio; objeto não próprio; abstração não única; abstração única), Função sintática do SN (Sujeito; Tópico; Adjunto; Genitivo; Objeto direto; Objeto indireto; Predicativo), Especificidade do núcleo do SN (específico; não específico) e o Item lexical (variável aleatória). A autora também testou a correlação com as variáveis sociais Faixa etária (I, 15-34 anos; II, 35-50 anos), Idade de migração (até 19 anos; 20+ anos), Tempo de permanência em SP (menos de 9 anos; 10+ anos), e o Participante como variável aleatória. O resultado geral de Guirelli (2018) mostrou que não há diferença significativa entre as taxas de uso do Artigo antes de possessivos entre os migrantes paraibanos residentes em Campinas (SP) (44,84%) e os paraibanos não migrantes (43,97%), mas há diferença entre essas duas amostras e os dados de São Paulo (SP) (55,82%). As análises estatísticas de regressão logística mostraram que o Artigo definido antes de possessivos foi favorecido por sintagmas nominais com função de objeto indireto (0,74 *logodds*,  $p = 0,05$ ) e genitivo (3,86 *logodds*,  $p = 0,001$ ); por pronomes possessivos no singular (0,72 *logodds*,  $p = 0,01$ ); precedidos por preposição que se contrai (4,22 *logodds*,  $p = 0,001$ ) e por sintagma nominal não específico (0,51 *logodds*,  $p = 0,01$ ). Por outro lado, a variável dependente foi desfavorecida por sintagmas com a função de objeto direto (-2,14 *logodds*,  $p = 0,01$ ), predicativo (-1,95 *logodds*,  $p = 0,01$ ), sujeito (-1,32 *logodds*,  $p = 0,01$ ); por possessivos masculinos (-0,34 *logodds*,  $p = 0,01$ ); por sintagma nominal com natureza semântica de parente (-1,06 *logodds*,  $p = 0,05$ ) e de objeto próprio (-1,43 *logodds*,  $p = 0,01$ ). Não houve correlação da variável dependente com nenhuma variável social. Ao comparar esses resultados com os valores encontrados nas análises da amostra dos paraibanos não migrantes e com os valores vistos na amostra de paulistas, Guirelli (2018) notou que os migrantes paraibanos se adaptaram linguisticamente para as regras mais abstratas da variedade da comunidade anfitriã, tal qual o resultado de Guedes (2019 [2017]).

Souza (2019) estudou as formas imperativas, como em “Traga o livro” e “Traz o livro”, empregadas por migrantes baianos residentes em São Paulo no questionário *on-line* elaborado por Figueredo e Souza (2017) para coletar dados dessa variável sociolinguística. A expectativa para essa análise foi a de que os migrantes baianos residentes em São Paulo utilizariam o Imperativo com morfologia de indicativo, típico da comunidade anfitriã, em relação ao

Imperativo com morfologia de subjuntivo. O *corpus* explorado por Souza (2019) consta de respostas dadas por 73 baianos a 15 sentenças-alvo, sendo 41 não migrantes, usados apenas como grupo controle, e 32 residentes em São Paulo, perfazendo o total de 1.168 observações. Com base em análises de regressão logística de efeitos mistos, com a inclusão do Participante e da Sentença como variáveis aleatórias, feitas na plataforma R, Souza (2019) testou a correlação da variável dependente com as variáveis sociais Nível de escolaridade (Até o Ensino Médio; Nível Superior), Idade de migração (Faixa I, 13-21 anos; II, 22-24 anos; III, 25-35 anos) e Tempo em São Paulo (I, 0-7 anos; II, 8-15 anos; III, 16+), Tipo de relação (simétrica; assimétrica); e as variáveis linguísticas Intenção comunicativa (aconselhar; instruir; pedir; ordenar), Contexto temporal (imediatos; não-imediatos) e Polaridade (negativa; afirmativa). O resultado geral dessa análise mostrou que há diferença significativa do emprego do Imperativo com morfologia de indicativo entre os baianos residentes em São Paulo (34%) e os baianos não migrantes (23%), i.e., ao migrarem para São Paulo, os baianos tendem a usar mais a variante típica da comunidade anfitriã. Quanto aos preditores investigados, os dados depreendidos dos testes estatísticos revelaram que o Imperativo com morfologia de indicativo é desfavorecido por migrantes que estudaram até o Ensino Superior (-1,23 *logodds*,  $p = 0,0001$ ), por quem está entre 8-15 anos (-0,95 *logodds*,  $p = 0,0001$ ) e há mais de 16 anos (-1,14 *logodds*,  $p = 0,0001$ ) residindo em São Paulo, por relação simétrica (-0,45 *logodds*,  $p = 0,0001$ ), por contextos não imediatos (-0,31 *logodds*,  $p = 0,0001$ ) e por frase com polaridade negativa (-0,73 *logodds*,  $p = 0,0001$ ). Em compensação, essa variante é favorecida por baianos que migraram com idade entre 22-24 anos (1,12 *logodds*,  $p = 0,0001$ ) e entre 25-35 anos (0,19 *logodds*,  $p = 0,0001$ ) em relação aos que migraram com idade entre 13-21 anos; por contextos comunicativos de pedido (0,21 *logodds*,  $p = 0,0001$ ) e ordem (0,11 *logodds*,  $p = 0,0001$ ) em comparação com os de instrução (0,11 *logodds*,  $p = 0,0001$ ) e de aconselhamento. Em síntese, os migrantes baianos que possuem o Ensino Médio, que estão há menos tempo em São Paulo, que migraram mais tarde, em contextos de situações assimétricas, em contexto imediato, em sentenças de polaridade afirmativa, em contextos comunicativos de pedido e ordem tendem a usar mais a variante imperativa típica da variedade paulista. Destaca-se aqui o resultado para o Tempo de residência, porque a maioria dos trabalhos resenhados nesta Subseção apresentou a tendência contrária, i.e., quanto mais tempo o migrante está na comunidade anfitriã, maior é a frequência de uso das variantes típicas da comunidade anfitriã. De acordo com Souza (2019), a redução das taxas do Imperativo com morfologia de indicativo dos migrantes baianos residentes em São Paulo ao longo do tempo de estadia resulta de um relaxamento linguístico. Em outros termos, com o passar do tempo, os migrantes tendem a obliterar variantes para as quais se adaptaram no primeiro momento do contato, talvez por se sentirem engajados na comunidade anfitriã.

Numa análise de rede social, Bortoni-Ricardo (2011 [1985]) examinou o uso variável de traços da fala rural por parte de migrantes mineiros em Brasília (DF), com base na

hipótese de que os migrantes com maior integração urbana se adaptariam linguisticamente ao novo dialeto em comparação com os migrantes com baixa integração urbana. Para abordar os participantes de sua pesquisa, [Bortoni-Ricardo \(2011 \[1985\]\)](#), ao se inserir na comunidade averiguada, apresentava-se como amiga de X ou de Y, normalmente pessoas com certa distinção social naquele lugar. Depois de inserida, ela escolheu cinco âncoras para serem pontos de partida das redes e os entrevistou. Ao final de cada entrevista, ela perguntava “quais eram as três pessoas com as quais elas frequentemente conversavam fora as da casa delas”. Essa metodologia rendeu uma amostra de 118 falantes.

[Bortoni-Ricardo \(2011 \[1985\]\)](#) colocou a sua hipótese à prova, testando a correlação da Vocalização da lateral palatal /ʎ/, como em /mu'ʎer/ ~ /mu.j'jɛr/; a monotongação de ditongos, como em /'owru/ ~ /'oru/ e a marca Zero de concordância verbal (1ª (CV1) e 3ª pessoas do plural (CV3)), como em /fa'zemus/ ~ /fa'zemu/ com as variáveis sociais Sexo/Gênero (homem; mulher), Idade (jovem; adulto), Exposição à mídia (baixa; alta) e Categoria profissional (baixa, alta). Na análise da Vocalização da lateral palatal, na qual também se testou a sua correlação com a variável linguística Vogal anterior (V - qualquer vogal, exceto /i/; Y - vogal /i/), [Bortoni-Ricardo \(2011 \[1985\]\)](#) verificou que os migrantes homens, jovens, que têm alta exposição à mídia, que pertencem a categoria profissional alta, em contexto linguístico em que a lateral é precedida pela vogal Y tendem a usar a variante de prestígio da lateral. Na análise da monotongação de ditongos, na qual se analisou também a variável linguística Consoante precedente (fricativas/sonorantes; oclusivas; assimiladas em palatalização), [Bortoni-Ricardo \(2011 \[1985\]\)](#) notou que as mulheres, jovens, que pertencem a categoria profissional alta, que têm alta exposição à mídia, em contextos linguísticos em que o ditongo é precedido por consoantes assimiladas pela palatalização, tendem a realizar mais a forma padrão. Na análise de marca Zero da CV3, em cujo modelo se incluíram as variáveis linguísticas Classe morfológica (desinências não acentuadas; desinências acentuadas), [Bortoni-Ricardo \(2011 \[1985\]\)](#) viu que os migrantes homens, jovens, com alta exposição à mídia, em contexto linguístico em que as desinências verbais são acentuadas, em sentenças cujo sujeito está implícito e anteposto ao verbo tendem a realizar a forma padrão da CV3. Por fim, na análise da marca Zero da CV1, na qual se testou a sua correlação com as variáveis linguísticas Tonicidade (paroxítona, proparoxítona) e Posição do sujeito, [Bortoni-Ricardo \(2011 \[1985\]\)](#) observou que os homens, jovens, que têm alta exposição à mídia, em contexto linguístico em que as formas verbais são paroxítonas, em sentenças cujo está sujeito implícito tendem a realizar a variante padrão da CV1. Após as análises, [Bortoni-Ricardo \(2011 \[1985\]\)](#) concluiu que a mudança desses falantes da zona rural para a urbana fez com que eles se envolvessem numa rede “isolada”, constituída por parentes e vizinhos que pertenciam a uma rede urbana “integrada” e, a partir disso, estavam propensos a aderir ao padrão da comunidade anfitriã.

[Adant \(1989\)](#), a partir de uma amostra de fala de 40 alagoanos, 20 residentes em

Alagoas e 20 em Brasília (DF), contrastou a realização das Vogais médias pretônicas /e/ e /o/ e das Oclusivas alveolares /t, d/ diante de [i] para averiguar se os migrantes alagoanos, ao conviverem com os brasilienses, manteriam as características da sua fala nativa ou se adaptariam<sup>11</sup> aos aspectos da fala da comunidade anfitriã. Os resultados gerais dessa análise permitiram observar que os migrantes alagoanos apresentaram taxas de uso diferentes das variantes predominantes na fala dos alagoanos não migrantes. Os valores encontrados por Adant (1989) mostram que os migrantes alagoanos tenderam a usar mais a vogal média pretônica /e/ mais alta (63,39%) em comparação com os alagoanos não migrantes (42,86%). Por outro lado, os migrantes tenderam a empregar a Vogal média pretônica /o/ mais baixa (54,23%) em relação ao que não migraram (49,24%). Quanto ao uso de /t, d/ diante de [i], a autora notou que os migrantes alagoanos em Brasília (DF) tenderam a realizar menos as Oclusivas (72,58%) em comparação com os não migrantes (92,79%).

Santana (2023) avaliou o efeito de rede social na amostra de fala de 27 migrantes sergipanos residentes em São Paulo (SP) (Santana, 2018) quanto à realização de /t, d/ diante de [i] e da Negação sentencial, partindo da hipótese de que os migrantes sergipanos pertencentes a redes abertas estariam propensos a aderir ao padrão paulistano com relação à forma Palatalizada e à negação pré-verbal. Em seguida, Santana (2023) coadunou esses resultados com os que ela obteve na análise da Altura das vogais médias pretônicas (Santana, 2018) e verificou se as variantes típicas da variedade paulista covariam. Ou seja, se a variante Palatal, Neg1 e as pretônicas /e/ e /o/ mais altas se correlacionam. Para alcançar os objetivos de sua pesquisa, Santana (2023) testou, realizando análises estatísticas multivariadas de regressão logística de efeitos mistos, com inclusão do Participante e do Item lexical como variáveis aleatórias, a correlação de /t, d/ e de (Neg) com os preditores sociais Rede social (aberta; fechada), Idade de migração (variável contínua), Tempo de residência (variável contínua), Índice de integração à rede (variável contínua), Identificação do migrante em relação à comunidade anfitriã (variável contínua), Gênero (masculino; feminino) e Escolaridade (Fundamental I; Fundamental II; Ensino Médio). Depois de obter os seus resultados, Santana (2023) comparou seus dados com os valores alcançados das análises feitas a partir da amostra de fala de sergipanos não migrantes e com os dados de paulistanos não migrantes. Os resultados globais das análises de /t, d/ diante de [i] mostraram que, do total de 2.160 observações, 61,2% foram da variante Palatalizada e 38,8% da variante Oclusiva. Os resultados das análises multivariadas, nas quais se adicionaram aos modelos estatísticos os preditores linguísticos Sonoridade (sonora; surda), Tonicidade (tônica; postônica; pretônica inicial; pretônica não inicial), Contexto fonológico precedente (pausa; fricativa alveolar; nasal; rótico; semivogal; vogal anterior; vogal central; vogal posterior), Contexto fonológico seguinte (pausa; fricativa; líquida; nasal; oclusiva; vogal não posterior; vogal posterior) e Status da vogal

<sup>11</sup> O autor usa o termo “acomodar”.

(fonética; fonológica), [Santana \(2023\)](#) notou que as formas Palatalizadas tendem a acontecer mais na fala dos sergipanos residentes em São Paulo (SP) que migraram mais cedo e que estudaram até o Ensino Fundamental II ou até o Ensino Médio (0,299 *logodds*,  $p= 0,014$ ), em contexto linguístico em que a consoante é surda (1,212 *logodds*,  $p= 0,001$ ), em contextos em que /t, d/ são precedidos por consoante fricativa alveolar (1,244 *logodds*,  $p= 0,007$ ) e seguidos de vogal posterior (1,359 *logodds*,  $p= 0,027$ ).

Nas análises de (Neg), [Santana \(2023\)](#) verificou que, do total de 3.407 observações, 85,79% são de Neg1, 12,71% de Neg2 e 1,5% de Neg3. Ao contrapor esses valores com as taxas de (Neg) encontradas na amostra de fala de sergipanos não migrantes (79,3% de Neg1, 19,1% de Neg2 e 1,6% de Neg3) e com as proporções vistas por [Rocha \(2013\)](#) (ver a Seção 6) na fala de paulistanos (94% de Neg1, 5,8% de Neg2 e 0,2% de Neg3) reportados por [Santana \(2023\)](#), pode-se depreender que os migrantes sergipanos ao residirem em São Paulo (SP) tendem a reduzir, sobretudo, as taxas de Neg2 em favor do padrão paulistano em que prevalece o uso de Neg1. Baseando-se no trabalho de [Rocha \(2013\)](#), [Santana \(2023\)](#) amalgamou os dados de Neg2 e Neg3 e encontrou, depois da exclusão de dados duvidosos, 84,95% de Neg1 e 15,05% de Neg2/Neg3. Nas análises multivariadas de regressão logística de efeitos mistos, nas quais se incluíram os preditores linguísticos Ativação da proposição (direta; indireta), Marcador conversacional (ausente, presente) e Outro termo negativo (ausente; presente), [Santana \(2023\)](#) constatou que Neg2/Neg3 tendem a não ser usadas na fala dos migrantes sergipanos que integram redes mais abertas (-0,600 *logodds*,  $p= 0,024$ ), em contextos em que a ativação da proposição é ativada de maneira inferida (-0,649 *logodds*,  $p= 0,001$ ), em sentenças em que há marcador conversacional (-1,220 *logodds*,  $p= 0,001$ ) e há outro termo negativo (-0,785 *logodds*,  $p= 0,001$ ). Diferentemente desses resultados, na análise da amostra controle de sergipanos não migrantes, na qual se testaram os mesmos preditores linguísticos, [Santana \(2023\)](#) encontrou correlação de (Neg) apenas com a Ativação da proposição (-1,045 *logodds*,  $p= 0,001$ ), de maneira que Neg2/Neg3 tendem a não ser empregadas quando a proposição negada é ativada de forma inferida.

Os resultados da Ativação da proposição no uso de Neg2/Neg3 nas duas amostras chamam bastante atenção, porque há diferenças nas taxas de aplicação entre elas. Na amostra controle, [Santana \(2023\)](#) verificou 30% de Neg2/Neg3 no contexto em que a ativação da proposição é ativada de forma direta e 12% no contexto em que ela foi inferida, uma diferença de 18 pontos percentuais. Em contrapartida, na amostra da fala dos migrantes sergipanos, [Santana \(2023\)](#) constatou 21% de aplicação de Neg2/Neg3 no contexto de ativação direta e 10% no contexto de ativação inferida, diferença de 11 pontos percentuais. Em ambas as amostras, portanto, Neg2/Neg3 tendem a ser desfavorecidas quando a ativação é inferida. Contudo, na fala dos migrantes sergipanos, vê-se que Neg2/Neg3 são mais impactadas quando a proposição é ativada de maneira direta, uma vez que há uma redução de 30% da amostra controle para 21%

em comparação com os dados da ativação inferida, que saiu de 12% da amostra controle para 10%.

[Figueiredo \(2014\)](#), à luz da Dialetologia Pluridimensional, investigou os efeitos do contato intervareietal na fala de gaúchos residentes em Mato Grosso, analisando uma amostra composta por 16 participantes, sendo 8 topodinâmicos e 8 topostáticos, estratificados em Sexo (homem; mulher), Grupo etário (GI, 18-36 anos; GII, +50 anos), Procedência ético-linguística (português; alemão; italiano) e Escolaridade (até o Ensino Médio e Ensino superior). [Figueiredo \(2014\)](#) analisou a relação das dimensões Diageracional, Diassexual, Diastrática, Diafásica, Diatópica, Diareferencial e Contatual com um conjunto de variáveis de diferentes naturezas, a saber, a Palatalização das dentais /t, d/ diante de [i, e], como em *mentira*, *noite*, *dia*, *tarde*; a realização da vibrante: em posição intervocálica, como em *arroz*; em posição inicial, como em *revólver*; em posição final, como em *varrer*; e antes de consoante sonora, como em *corda* e *fervido*; a realização da lateral: seguindo consoante na mesma sílaba, como em *bicicleta*; no final da sílaba, como em *calção*; como palatal, como em *colher*; em posição final de sílaba tônica, como em *revólver*; na posição de sílaba e palavra, como em *sal*; em posição final de sílaba prôtônica, como em *calção*. [Figueiredo \(2014\)](#) observou também a ditongação diante de sibilante em monossílabos com vogal [a, e, u], como em *paz*, *dez* e *cruz*; a realização do ditongo decrescente nasal *-ão* tônico final, como em *procissão* e 20 variáveis lexicais. Especificamente a respeito das análises descritivas do uso das variantes menos gaúchas [-RS] e mais gaúchas [+RS] em relação às dimensões Diageracional, Diassexual, Diastrática, [Figueiredo \(2014\)](#) verificou que acerca da dimensão Diageracional, o GII, o grupo dos mais velhos, tende a manter os traços da comunidade matriz [+RS] em relação ao GI, o grupo dos jovens, que incorporam os aspectos linguísticos do novo meio [-RS]. Quanto à dimensão Diassexual, [Figueiredo \(2014\)](#) notou que as mulheres tendem a utilizar as variantes [-RS] em comparação com os homens. Em relação à dimensão Diastrática, os dados de [Figueiredo \(2014\)](#) apresentaram não haver diferença entre os participantes com menos escolaridade e mais escolaridade.

[Meurer \(2022\)](#), também com base na metodologia da Dialetologia Pluridimensional, contrastou o comportamento linguístico de dois grupos de falantes: o topodinâmico, constituído por cinco migrantes gaúchos (três homens e duas mulheres), com o topostático, formado por cinco participantes nortistas (dois homens e três mulheres). A intenção desse estudo foi saber se os participantes topodinâmicos substituiriam as variantes da comunidade matriz pelas variantes do novo meio. Para tanto, [Meurer \(2022, p. 189\)](#) observou os efeitos do contato intervareietal na altura das Vogais médias pretônicas /e/, como em *ferver*, e /o/, como em *procissão*; a fricatização da vibrante em sílaba inicial, como em *revólver*, intervocálica, como em *arroz*, e em coda interna, como em *tarde*; a Palatalização de /s/ em coda precedendo dentais /t/ e /d/, como em *nordestino* e *estilingue*; a Vocalização da lateral em coda, como *agricultor*; a sonorização das oclusivas /b/, /d/ e /g/, como em *urubu*, *bodoque*, *bergamota* e o vocabulário

diferenciado. Além disso, a partir de análises descritivas, a autora correlacionou as variáveis linguísticas às dimensões Diatópica (Sul; Norte), Diatópico-cinético (topodinâmicos/migrantes sulistas; topostáticos/sulistas da matriz de origem), Dialingual (topodinâmico/migrantes sulistas; topostáticos/nordestinos), Diageracional (GI - 48 a 58 anos; GII - +60), Diastrática (Ca - classe alta + ocupação; Cb - Classe baixa + ocupação), Diagenérica (homem; mulher) e Diarreferencial (respostas objetivas; comentários metalinguísticos). Dos resultados encontrados por [Meurer \(2022\)](#), evidenciam-se aqui apenas os obtidos nas análises da dimensão Diageracional, Diagenérica e Diastrática. Sobre a primeira dimensão, os dados sugerem que o GII é grupo conservador das variantes da comunidade matriz, enquanto o GI é que mais apresentou mudanças em favor da variedade do novo meio, à exceção da realização das Vogais médias pretônicas [e] e [o] e do /S/ em coda, realizado como [s] sibilante alveolar. Para as 40 variáveis lexicais estudadas, a autora reporta que o GI pouco utiliza as variantes sulistas, tampouco utiliza as variantes nortistas. Segundo ela, o GI tem utilizado variantes lexicais menos marcadas regionalmente, com uma leve tendência para o léxico do novo meio. A análise qualitativa de [Meurer \(2022, p. 284–285\)](#) sobre as dimensões Diagenérica e Diastrática apresenta que as mulheres tendem a ser mais conservadoras que os homens. Em outras palavras, foram os homens que aderiram aos traços do novo meio. Ao cruzar esse dado com as características sociais, a autora notou que as mulheres são donas de casa e, portanto, têm atividades mais restritas que os homens e, talvez por isso, tenham mantido o uso dos itens lexicais da comunidade matriz.

[Oliveira \(2023, inédito\)](#) analisou /t, d/ diante de [i], o /s/ em posição de coda e o /r/ em posição de coda. A amostra utilizada por [Oliveira \(2023, inédito\)](#) é composta por 32 alagoanos que residiram em São Paulo e voltaram ao seu lugar de origem. Aliás, esse parece ser um trabalho pioneiro, tendo em vista que, pelo menos até o momento de produção desta Tese, não havia notícias sobre pesquisas com esse escopo. A amostra estudada por [Oliveira \(2023, inédito\)](#) foi estratificada em Sexo, Tempo de residência em São Paulo (menos de 5 anos; mais de 5 anos), Tempo de retorno (menos de 5 anos; mais de 5 anos). Na análise do /s/ diante de [t, d], contrastou-se o uso da Fricativa palatal (típica de Alagoas) com o uso da Fricativa alveolar (típica de São Paulo). Além das variáveis estratificadoras da amostra, [Oliveira \(2023, inédito\)](#) testou a correlação dessa variável sociolinguística com Idade (variável contínua), Anos de escolaridade (variável contínua), Local de residência (rural; urbano), Tipo de imóvel (alugado; financiado; próprio), Renda familiar (Até 1 salário mínimo (SM); 2SM–3SM; acima de 3SM), Estilo (entrevista; leitura), Contexto fonético precedente (as vogais [i, e, ε, a, ɔ, o, u]), Contexto fonético seguinte (t; d), Classe morfológica (adjetivo; advérbio; pronome; substantivo; verbo) e Tonicidade (átona; tônica). Nas análises de regressão multinível, viu-se que a manutenção da variante paulista tende a acontecer mais na fala de regressos mais velhos e que voltaram há mais tempo, residentes na zona rural (P.R. .76), que possuem renda familiar entre 2SM–3SM (P.R. .86), em contexto de leitura (P.R. .83).

Para o estudo de /t, d/ diante de [i], em cuja análise comparou-se o emprego da forma Oclusiva (típica de Alagoas) com a forma Palatalizada (típica de São Paulo), Oliveira (2023, inédito) testou a correlação dessa variável sociolinguística com as mesmas variáveis sociais e contextual da análise do /s/ em coda e com as variáveis linguísticas Contexto fonético precedente (vogal alta; vogal média; vogal baixa; fricativa; fronteira inicial de palavra), Consoante alvo (t; d), Classe morfológica (adjetivo; advérbio; interjeição; numeral; preposição; pronome; substantivo; verbo) e Tonicidade (monossílabo átono; pretônica; tônica; postônica). Os resultados estatísticos dos modelos de regressão multinível mostram que os regressos de São Paulo com mais anos de escolaridade e em contexto de leitura tendem a empregar a forma Palatalizada, na consoante alvo /t/ ( P.R. .62), quando precedida por fricativa (P.R. .83) e em sílabas pretônicas.

Na análise do (-r) em coda, em qual se confrontou a variante Glotal (típica de Alagoas) com as variantes vibrante simples e retroflexa, Oliveira (2023, inédito) testou a correlação de (-r) com as mesmas variáveis sociais e contextual investigadas no estudo do /s/ diante de /t, d/ e de /t, d/ diante de [i], além das variáveis linguísticas Contexto fônico precedente (as vogais [i, e, ε, a, ɔ, o, u]), Contexto fônico seguinte (coronal; dorsal; labial), Classe morfológica (adjetivo; advérbio; conjunção; numeral; substantivo; verbo) e Tonicidade (átono; tônica). Os resultados estatísticos indicam que os regressos de São Paulo para Alagoas com menos anos de escolaridade (P.R. -0,0365), que residiram há mais tempo em São Paulo (P.R. 0,0211), em contexto de entrevista (P.R. .65), em contextos em que o /r/ é seguido de consoante coronal (P.R. .64) tendem a manter a variante da comunidades de onde retornaram.

Os resultados dessas pesquisas contribuíram para a tomada de decisões relacionadas à escolha: das variáveis sociolinguísticas a serem observadas (ver a Introdução), da metodologia utilizada para abordar os participantes, dos preditores sociais e suas hipóteses e dos testes estatísticos utilizados para testá-las (ver a Seção 3). Além disso, a análise desses dados permitiu a formulação de generalizações adicionais sobre a fala dos migrantes em situação de contato dialetal, como evidenciado na Seção 7.

### 2.3 Síntese

Nesta Seção, fez-se uma breve revisão da literatura sobre o contato entre dialetos mutuamente inteligíveis. Ao retomar algumas pesquisas, foi possível observar que: (i) Trudgill (1986) explica os efeitos do contato dialetal a partir da Teoria da Acomodação (Giles; Taylor; Bourhis, 1973); Chambers (1992) e Siegel (2010) discordam de Trudgill (1986) quanto ao uso dessa teoria nesse tipo de pesquisa e propõem que as adaptações linguísticas na fala de migrantes em situação de contato dialetal resultam da aquisição de segundo dialeto; (ii) o Projeto Acomodação (Oushiro, 2017) tem analisado sistematicamente a fala de migrantes alagoanos e paraibanos residentes em São Paulo (SP) e na Região Metropolitana de Campinas (SP), ampliando as discussões acerca

---

desse tópico; (iii) a grande maioria dos trabalhos realizados sobre a fala de migrantes em situação de contato dialetal tem se concentrado no estudo de apenas uma variável sociolinguística, comumente de natureza fonético-fonológica ou morfossintática; (iv) raras são as análises sobre itens lexicais ou a análise de múltiplas variáveis.

## 3

# Materiais e métodos

Na presente Seção, descrevem-se os procedimentos metodológicos utilizados na pesquisa sobre a fala de migrantes baianos residentes na RMSP. A sua estrutura está assim organizada: na Subseção 3.1, mostra-se como, aplicando o método de rede social (Bortoni-Ricardo, 2011 [1985]; Santana, 2018), recrutaram-se os participantes da amostra. Na Subseção 3.2, detalham-se as perguntas do roteiro usado nas entrevistas sociolinguísticas realizadas com eles a fim de coletar os dados linguísticos necessários para testar as hipóteses aventadas para este estudo. Na Subseção 3.3, apresentam-se os participantes da amostra. Na Seção 3.4, destacam-se os aspectos sociais de maior relevância dos participantes para esta pesquisa. Na Subseção 3.5, mencionam-se as ferramentas e os métodos adotados para manipular e analisar estatisticamente os dados coletados da amostra de fala dos migrantes baianos residentes na RMSP. Na Subseção 3.6, listam-se as hipóteses aventadas para os preditores sociais. Por fim, na Subseção 3.7, apresenta-se o resumo desta Seção.

### 3.1 Constituição da amostra

Pode-se constituir uma amostra de fala de maneira longitudinal ou transversal (Woods; Fletcher; Hughes, 1986). O primeiro modo é realizado a partir da observação dos mesmos indivíduos em diferentes períodos, como no estudo de Chambers (1992) sobre seis jovens imigrantes de duas famílias canadenses que se mudaram para o sul da Inglaterra e cuja fala foi gravada em dois momentos entre 1983 e 1984 (ver a Seção 2). O segundo é feito com diferentes pessoas que representam os perfis de interesse do pesquisador. Um exemplo desse método é o que foi empregado na pesquisa de Oushiro (2020b), que observou uma amostra estratificada em Sexo/Gênero, Idade de migração e Tempo de residência de 40 migrantes alagoanos e paraibanos residentes na Região Metropolitana de Campinas (SP) (ver a Seção 2). Considerando essas duas possibilidades metodológicas, constituiu-se, transversalmente, uma amostra de fala com 50 migrantes baianos residentes na Região Metropolitana de São Paulo para alcançar os objetivos pretendidos nesta pesquisa (ver a Seção de Introdução). A coleta foi feita por este pesquisador,

também baiano, no período de abril a agosto de 2018, após a aprovação do projeto pelo *Comitê de Ética em Pesquisa - IEL/UNICAMP*<sup>12</sup>. O fato de o pesquisador e os participantes serem da mesma proveniência os coloca numa relação simétrica quanto ao aspecto regionalista.

Diferentemente da abordagem utilizada por Chacon (2012), Oliveira (2020), Oushiro (2020b) e Possatti (2020) (ver a Seção 2), que compuseram amostras de fala de migrantes com representantes das variáveis sociais de seu interesse, os baianos residentes na RMSP foram recrutados a partir do método de *redes sociais* (Bortoni-Ricardo, 2011 [1985]; Santana, 2018). Essa técnica, inicialmente utilizada em pesquisas da Antropologia, da Sociologia e da Psicologia Social, é, hodiernamente, aplicada em outras áreas, como na Ciência da Informação e na Sociolinguística.

Em estudos sociolinguísticos, conforme se viu na Seção 2, essa metodologia consiste em mapear as pessoas com as quais os participantes interagem. Para Bortoni-Ricardo (2011 [1985]), a análise de redes sociais pode fornecer às pesquisas sociolinguísticas meios de conhecer as inter-relações pessoais numa comunidade e faz com que se percebam quais elementos linguísticos são compartilhados entre os indivíduos, tendo em vista que as regras de usos são constructos sociais e, por isso, as diferenças entre as redes podem explicar as diferentes normas aplicadas em cada uma delas.

De modo diferente aos estudos de Bortoni-Ricardo (2011 [1985]) e Santana (2018), que escolheram os âncoras de suas pesquisas, os âncoras da amostra coletada para o presente estudo foram indicados por baianos não migrantes, denominados aqui como *intermediários*<sup>13</sup>. Da Bahia, cada intermediário (três) entrou em contato com um baiano residente na Região Metropolitana de São Paulo (normalmente parente ou amigo) e apresentou o pesquisador como alguém próximo que precisava entrevistá-los para um trabalho acadêmico sobre como era a vida dos baianos fora do seu lugar de origem. O único critério estabelecido foi o de que, como na pesquisa de Santana (2018), os âncoras não poderiam se conhecer.

Assim, dos indicados, três pessoas concordaram em participar do estudo: (1) AlanaS<sup>14</sup>, de 37 anos, esteticista, moradora do bairro da Sapopemba, e está em São Paulo há 11 anos; (2) MarieneS, de 47 anos, assessora da empresa da família, residente no Jaraguá e está em São Paulo há 29 anos; e (3) JoãoS, 42 anos, monitor escolar numa instituição privada, residente no bairro da Bela Vista e vive em São Paulo há 10 anos. Para observar o modo de vida, as características da fala e as redes sociais desses migrantes, foi realizada uma entrevista

<sup>12</sup> Aprovado em 22 de fevereiro de 2018, sob o parecer substanciado de número 2.484.322 e o CAAE 80645217.4.0000.8142.

<sup>13</sup> Os intermediários não podem ser considerados âncoras, porque não fazem parte da comunidade migrante. Logo não têm o perfil de interesse desta pesquisa.

<sup>14</sup> Os nomes dos participantes foram substituídos por nomes fictícios a fim de assegurar o seu anonimato, conforme orientam as normas de transcrição do Projeto SP2010 (Mendes; Oushiro, 2013) e do Projeto Acomodação (Oushiro, 2017).

sociolinguística (Labov, 2006 [1966]; Tagliamonte, 2011; Becker, 2013) com cada um deles, seguindo o roteiro descrito na Subseção 3.2. Com esse tipo de coleta, pôde-se observar a fala semiespontânea e monitorada dos participantes e capturar fenômenos linguísticos variáveis.

### 3.2 Roteiro utilizado nas entrevistas sociolinguísticas

Para coletar os dados linguísticos dos migrantes baianos residentes na RMSP, adaptou-se o roteiro utilizado nas entrevistas sociolinguísticas realizadas no Projeto Acomodação (Oushiro, 2017), cuja versão consta de três partes (ver o Apêndice A). Na primeira delas, as questões conduziram à busca de informações pessoais dos migrantes, como as características do bairro onde moram; os relatos da infância; as pessoas que constituem a sua família; o(s) motivo(s) da migração; o trabalho/ocupação e o lazer. Perguntas como essas, segundo Labov (1972), contribuem para que o falante monitore menos a fala e algo próximo do vernáculo possa ser coletado.

Na segunda parte, expõem-se os estímulos sociolinguísticos usados para observar algumas variáveis dependentes, especialmente o (-r) em posição de coda em situação de fala monitorada, como na leitura de um texto e de uma lista de palavras<sup>15</sup> (ver o Apêndice A). A hipótese que fundamenta esse método é a de que há uma gradação quanto ao monitoramento da fala em diferentes contextos de estilo: conversação → leitura de texto → leitura de palavras (Labov, 2001). Nessa parte do roteiro, estão enumeradas as Questões Semântico-Lexicais (QSL), produzidas com base no *Atlas Linguístico do Brasil – ALiB* (Cardoso et al., 2014), com vistas a observar possíveis adaptações na fala dos migrantes baianos quanto ao uso de lexias tipicamente paulistas, como “Mexerica” em vez de “Tangerina”; “Mandioca” no lugar de “Aipim”; “Marmita” em substituição a “Quentinha” e “Serviço” ao invés de “Trabalho” (ver o Apêndice A). Elencam-se também, nessa seção, questionamentos sobre a percepção e a avaliação dos falares baiano e paulista. Na última parte, listam-se as perguntas, desenvolvidas a partir de Oushiro (2017) e Santana (2018), sobre o *modus vivendi* (ver o Apêndice B) dos migrantes. São questões que permitem aferir o grau de autoidentificação declarada deles, especificamente sobre como eles se denominam quanto à estadualidade<sup>16</sup>, se como baianos ou já como paulistas; sobre suas redes sociais e a respeito de seus hábitos alimentares e culturais; e sobre os vínculos com os familiares e amigos que deixaram na Bahia.

Pouco antes de encerrar a entrevista com cada migrante baiano, de modo semelhante a Bortoni-Ricardo (2011 [1985]) e Santana (2018), questionou-se “quais eram as dez pessoas com as quais eles mais conversavam/conviviam, incluindo as da casa”, com o intuito de obter

<sup>15</sup> Replicou-se a lista de palavras utilizada por Oushiro (2017), que controlou os contextos fônicos antecedente e seguinte do (-r) em coda, das vogais médias pretônicas /e/ e /o/ e do /t, d/ diante de [i].

<sup>16</sup> Para se referir à autoidentificação dos baianos, será usado o termo “estadualidade” no lugar de “baianidade”, já que este é comumente usado para fazer referência à identidade soteropolitana.

uma amostra de sua rede social e também de angariar outros participantes. Das pessoas citadas, convidaram-se para participar da pesquisa apenas aquelas que eram baianas; as que eram de outras proveniências foram computadas para controle a fim de estimar se os migrantes baianos residentes na RMSP convivem mais com outros baianos, com paulistas ou com pessoas de outros lugares (ver a Subseção 3.4).

As gravações foram feitas nos locais sugeridos pelos participantes (na própria residência, no trabalho ou em espaço público) para deixá-los mais à vontade durante a conversa. Devido a isso, os áudios, algumas poucas vezes, não ficaram livres de ruídos, sons de automóveis, sobreposição de vozes etc. Esses pequenos problemas foram ponderados durante o estudo, de maneira que os dados de algumas gravações<sup>17</sup> não entraram nas análises estatísticas.

### 3.3 Os participantes da amostra

O principal objetivo de usar a abordagem por rede social em pesquisas sociolinguísticas consiste em verificar “quem conversa com quem” (Bortoni-Ricardo, 2011 [1985]) e, com esse mapeamento, identificar quais são os fenômenos linguísticos compartilhados. Isso justifica o emprego da pergunta sociométrica<sup>18</sup> “quais são as dez pessoas com as quais você mais conversa/convive no dia a dia?”. Para além disso, é de interesse, numa análise de rede social, observar a densidade das relações das pessoas que integram as redes; se ela é mais ou menos densa. Parafraseando Mitchell (1969) e Milroy (1987 [1980]), Bortoni-Ricardo (2011 [1985], p. 91) descreve que a densidade da rede é equivalente ao “número de vínculos que existem de fato como uma proporção do número máximo de vínculos que poderiam existir”; e mostra a fórmula, vista no exemplo (1), usada por Barnes (1969 *apud* Bortoni-Ricardo, 2011 [1985]).

(1)

$$D = \frac{200a}{n(n-1)}$$

Genericamente, essa fórmula é uma simples divisão, de modo que 200 é um número estabelecido,  $a$  é a quantidade real de ligações e  $n$  o número total de integrantes da rede (Bortoni-Ricardo, 2011 [1985], p. 91). De modo igual, Santana (2018) criou um *índice de integração* à rede a fim de correlacioná-lo ao fenômeno estudado (ver a Seção 2), conforme a fórmula em (2).

<sup>17</sup> Nas Seções correspondentes aos resultados das variáveis sociolinguísticas reporta-se a quantidade de informantes observados em cada uma das análises.

<sup>18</sup> A sociometria é um conjunto de procedimentos quantitativos usados para mensurar os vínculos dos integrantes da rede. Ela engloba desde técnicas de coleta até a visualização de dados (Bortoni-Ricardo, 2011 [1985], p. 86–87).

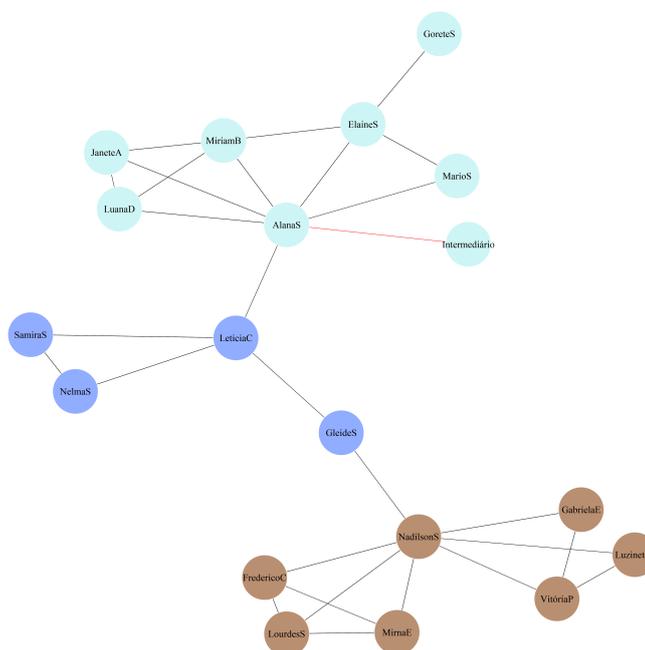
(2)

$$I = \frac{\sum c}{n}$$

Para ter o índice de integração da rede (I), ela somou o número de sergipanos citados na entrevista (g) com a quantidade de vezes que cada um desses migrantes foi citado (c) e dividiu pelo total de participantes na rede (n) (Santana, 2018, p. 65). Diferentemente dessas duas pesquisas, contabilizou-se a densidade de contato dos migrantes baianos na RMSP com base no total (0 a 10) das pessoas/proveniências com as quais cada participante disse conviver, independentemente de os indicados terem sido entrevistados ou não.

Para visualizar as amostras de rede social obtidas por meio de perguntas sociométricas, Bortoni-Ricardo (2011 [1985]) sugere que elas sejam observadas com matrizes e/ou sociogramas, como os *grafos*, por exemplo. Este tipo de gráfico é formado por nós, que representam os atores da comunicação; e por arestas, que refletem a densidade e/ou os tipos de vínculos que eles têm. Por serem mais intuitivos que as matrizes, as redes sociais dos migrantes baianos residentes na RMSP estão assim apresentadas. Para produzi-los, usaram-se os pacotes *Network* e *Igraph* da plataforma R (R Core, 2023). Para Sadler (2017), o R é o programa mais robusto para a análise de redes em relação ao *Gephi* e ao *Cytoscape*, por exemplo, devido à grande variedade de pacotes que dão suporte às nuances de cada pesquisa.

Na Figura 2, visualiza-se a rede social de AlanaS, na qual constam 18 participantes, sendo 15 mulheres e 3 homens, cujo recrutamento aconteceu logo depois que um dos intermediários sugeriu a participação dela nesta pesquisa.

**Figura 2** – Rede 1 – AlanaS

Fonte: Elaboração própria.

Os integrantes dessa rede estão representados por nós etiquetados com os pseudônimos e estão interligados por arestas, as quais retratam os vínculos que eles têm. Cabe ressaltar que a aresta que liga o intermediário a AlanaS está em vermelho para separar o elo que eles têm em relação ao dela com as outras pessoas do seu convívio. Os agrupamentos do grafo da Figura 2 estão assim apresentados: o primeiro, de nós azul-claros, é constituído de AlanaS (37 anos); LuanaD (19), sua filha; MarioS (71) e ElaineS (62), seus tios; JaneteA (27) e MiriamB (30), suas primas; e GoreteS (70), amiga íntima de ElaineS. Indicada pela âncora, sua conhecida LeticiaC (31 anos) inicia o segundo agrupamento, caracterizado pelos nós azul-escuros, que é composto de LeticiaC; SamiraS (54), sua mãe; NelmaS (43), sua tia; e GleideS (33), sua amiga. O terceiro agrupamento, representado pelos nós marrons, é encabeçado por NadilsonS (43), prestador de serviço de GleideS, em que constam MirnaE (44), sua esposa; LourdesS (44), sua irmã; FredericoC (50), seu cunhado e esposo de LourdesS; GabrielaE (47), sua cunhada e irmã de MirnaE; VitoriaP (52) e LuzineteE (35), suas amigas e conterrâneas.

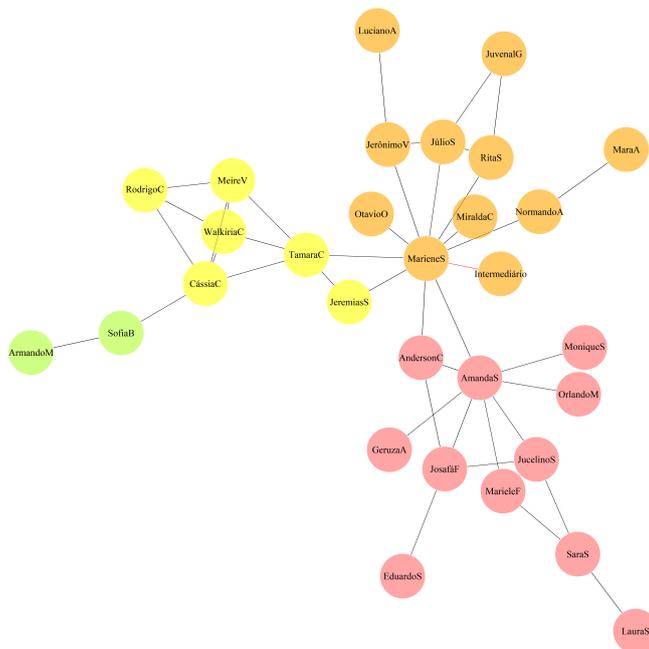
É importante ratificar que o fato de GoreteS e GleideS não terem, nesse gráfico, outros desdobramentos de contato com baianos não significa que elas convivam mais com paulistas e/ou pessoas de outras proveniências. Mas, por questões metodológicas, seja devido à logística da pesquisa de campo, seja por causa da indisponibilidade dos indicados, não foi pos-

sível entrevistá-las. De todo modo, a configuração que se apresenta nessa e nas outras redes é uma amostra da maneira como os migrantes baianos se relacionam na RMSP.

O grafo da Figura 2 é constituído de pequenos grupos de migrantes baianos (nós azul-claros, nós azul-escuros e nós marrons), os quais, em sua maioria, restringem-se a núcleos familiares e, portanto, são consideradas redes de troca (Milroy, 1987 [1980]). Graficamente, esses agrupamentos não foram estabelecidos aleatoriamente, mas a partir de comandos dos pacotes usados para posicionar os nós (Sadler, 2017). Basicamente, o que o pacote do R (R Core, 2023) faz é ponderar o número de vezes que a pessoa foi indicada e a quantidade de retribuições durante a indicação. Em outras palavras, especificamente na Rede 1, depois de AlanaS ter sido indicada pelo intermediário, constata-se que a maioria dos baianos sugeridos por ela se aponta como sendo as pessoas com as quais elas mais mantêm contato, portanto elas integram o mesmo núcleo. Assim também, embora LeticiaC tenha sido indicada por AlanaS, ela não retribuiu a indicação, logo o programa calculou que o vínculo entre elas é fraco e, por isso, pertencem a núcleos diferentes.

A Figura 3 apresenta os vínculos dos migrantes baianos da Rede 2. Este grupo é constituído de 29 participantes (15 mulheres e 14 homens) e teve como ponto de partida MarieneS. De modo igual à Rede 1, ela é formada por um conjunto de redes de troca, representadas pelos nós laranjas, amarelos, verde-claros e vermelho pastel.

**Figura 3 – Rede 2 – MarieneS**



Fonte: Elaboração própria.

Os nós laranjas retratam as pessoas que têm maior vínculo com MarieneS. Neste caso, MiraldaC (47), sua melhor amiga; OtavioO (63), seu vizinho; NormandoA (55), seu amigo e conterrâneo; MaraA (52), também sua conterrânea e esposa de NormandoA; JeronimoV (41), seu primo; LucianoA (65), seu conhecido e vizinho de JeronimoV; RitaS (77), sua tia; JuvenalG (45), seu primo e sobrinho de RitaS; JulioS (49), seu conterrâneo e amigo de JuvenalG e de RitaS. Apesar de ter sido indicado por sua tia MarieneS, JeremiasS (21) apresenta maior vínculo com TamaraC (33), sua esposa. Essa informante é a referência do grupo representado pelos nós amarelos que, além desse casal, é constituído por CassiaC (31) e WalkiriaC (23), suas irmãs; MeireV (28), sua prima, e RodrigoC (27), seu amigo e esposo de MeireV. CassiaC indicou SofiaB (44), sua conhecida e conterrânea, porém ela não retribuiu a indicação. Dos indicados por esta informante, apenas ArmandoM (49), seu esposo, foi entrevistado, o qual a apontou como pessoa com quem mais tem contato. O fato de esse casal não ter indicado outros participantes da rede dos nós amarelos sugere que eles constituem outro núcleo, por isso os nós estão em verde-claro.

Os nós em tom vermelho pastel apresentam os vínculos de AmandaS (45), irmã de MarieneS. Embora sejam irmãs, vê-se que elas pertencem a núcleos diferentes. Talvez a ausência de mutualidade durante a indicação seja porque elas moram um pouco distante uma da outra, não se vendo pessoalmente com muita frequência. Enquanto MarieneS mora no bairro do Jaraguá em São Paulo (SP), AmandaS e todos os demais que constam no seu núcleo moram no município de Francisco Morato (SP), na Região Metropolitana de São Paulo. Além de AmandaS, essa rede é formada por AndersonC (44 anos), seu primo; JosafaF (48), seu esposo; GeruzaA (61), MoniqueA (60) e OrlandoM (74), seus vizinhos; JucelinoS (54), seu pastor; MarieleF (52) e SaraS (46), suas colegas de igreja e irmãs de JucelinoS; LauraS (27), prestadora de serviço de SaraS e EduardoS (26), conhecido de JosafaS.

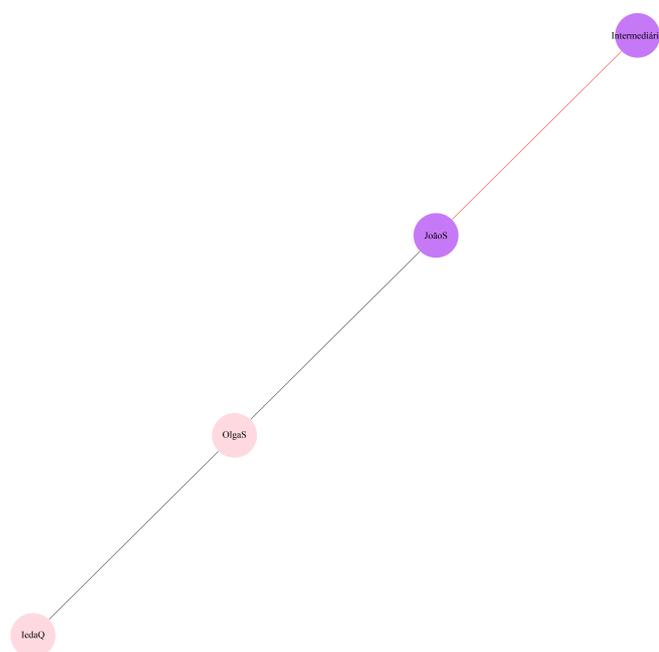
O design dessas redes chama bastante atenção, porque é o reflexo da maneira como acontece a migração de nordestinos para São Paulo. Segundo [Fontes \(2008\)](#), a grande maioria dos nordestinos que migraram para São Paulo (SP) em meados do século XX eram solteiros ou recém-casados que, depois de instalados e com certa “estabilidade”, serviam de apoio para outros familiares que também buscavam melhores condições de vida. Acerca da organização desse processo migratório, veja-se o relato de RitaS, de 77 anos, que mora em São Paulo (SP) há 53, no exemplo (3).

- (3) S1: *eu vim morar aqui porque a minha irmã já morava aqui né / e eu tinha vontade de vim pra São Paulo né / e minhas duas irmã a T. e D. já tinham vindo aqui né / e elas gostaram e eu tinha vontade de vim* (RitaS)

O depoimento de RitaS, em (3), demonstra que a migração para São Paulo (SP) não era/é desordenada, mas “meticulosamente pensada e preparada da melhor forma possível

tanto no âmbito familiar como no da comunidade” (Fontes, 2008, p. 55). Apesar de poucas, há também situações em que o migrante encontra o suporte para a mudança com nativos da comunidade anfitriã, ou seja, não necessariamente ele precisa de uma pessoa da sua família que também resida no lugar de destino. Este é o caso de JoãoS, âncora da Rede 3, que se mudou para São Paulo (SP) com o objetivo de morar com o namorado. O grafo da Figura 4 retrata as relações dessa rede.

**Figura 4** – Rede 3 – JoãoS



Fonte: Elaboração própria.

Assim que foi indicado por um intermediário, JoãoS foi contatado. Como resposta à pergunta sociométrica, ele relatou ter contato apenas com paulistas. Entretanto, como queria muito ajudar na pesquisa, sugeriu duas pessoas conhecidas do seu esposo, OlgaS (27) e IedaQ (30), as quais também foram entrevistadas. Dessa maneira, a Rede 3 é composta do âncora; de OlgaS e de IedaQ, que são colegas de sala num curso técnico.

### 3.4 Distribuição dos participantes nas categorias sociais investigadas

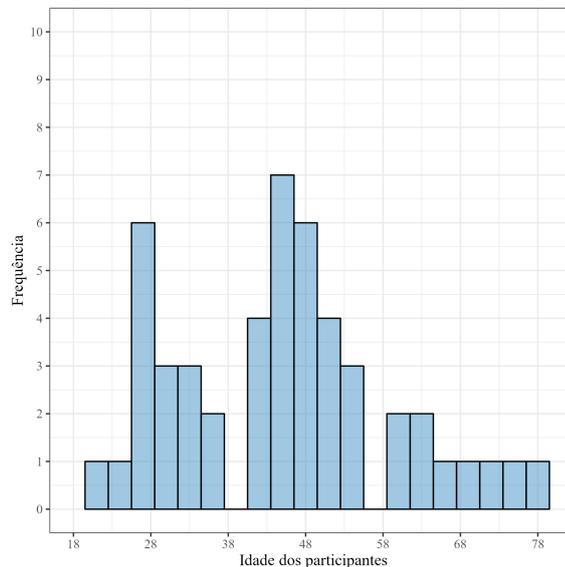
Tencionou-se, com base nos trabalhos resenhados na Seção 2, especialmente nos que estão vinculados ao Projeto Acomodação (Oushiro, 2017), testar a correlação dos dados linguísticos dos migrantes baianos na RMSP com os principais preditores por eles analisados, a saber, Sexo/Gênero, Idade, Nível de escolaridade, Idade de migração, Grau de autoidentificação declarada,

Tempo de residência e Rede social (ver a Seção 2). Contudo, como se pôde notar, a amostra de fala examinada na presente pesquisa não é estratificada aos moldes da Sociolinguística tradicional<sup>19</sup>, isto é, não há total controle da distribuição dos participantes pelos preditores sociais do interesse do pesquisador. Dessa maneira, antes de aventar as hipóteses para cada preditor social e testá-las estatisticamente, precisou-se avaliar o equilíbrio do número de participantes nas categorias sociais investigadas.

Do total de 50 participantes desta pesquisa, 32 (64%) são mulheres e 18 (36%) são homens, por conseguinte o preditor Sexo/Gênero está razoavelmente balanceado quanto à distribuição dos participantes entre seus níveis. Segundo Chaves (2012, p. 149), as mulheres têm migrado mais depois que houve uma “queda nos níveis de fecundidade, a intensa urbanização, o aumento dos níveis de escolaridade e a incorporação de novos papéis sociais” por parte delas; além de serem atraídas pelo serviço doméstico e/ou por terem acompanhado sua família, seja como filha, seja como esposa.

A amostra de fala analisada neste estudo se restringe à população adulta de migrantes baianos na RMSP, uma vez que se estabeleceu 18 anos como idade mínima para participar dela. Na Figura 5, na qual se apresenta o histograma da idade que os migrantes tinham quando foram entrevistados, observa-se não só que há representantes de várias idades mas também que o preditor Idade, nos testes estatísticos, pode ser analisado como uma variável contínua, i.e., sem a necessidade de agrupar as idades em faixas.

**Figura 5** – Frequência da idade dos migrantes baianos residentes na RMSP participantes da amostra



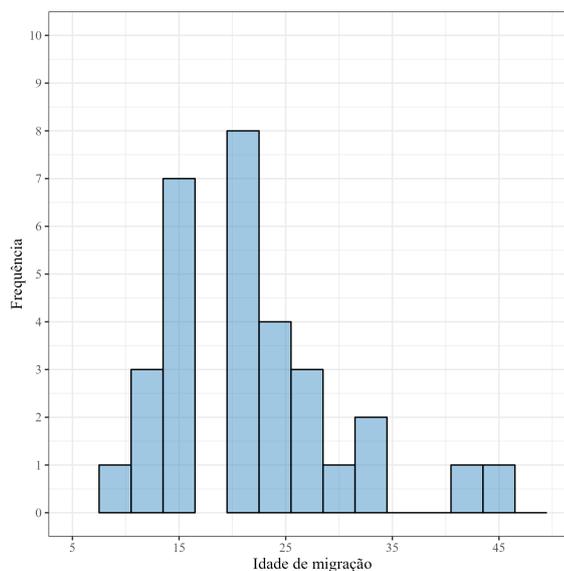
Fonte: Elaboração própria.

<sup>19</sup> Sobre a constituição de amostras na Sociolinguística tradicional do Brasil, recomenda-se a leitura dos capítulos 8 e 9 de Guy e Zilles (2007).

A respeito do Nível de escolaridade, vê-se que, dos 50 migrantes que compõem a amostra deste trabalho, 30 deles estudaram até o Ensino Fundamental, 17 estudaram até o Ensino Médio e apenas 3 eram universitárias: MiriamB cursava o 3º semestre de Arquitetura e MeireV e LuadaD tinham começado o Ensino Superior, mas tiveram que trancar os cursos. Considerando que essas três participantes estavam em processo de formação universitária, para as análises estatísticas, elas foram incorporadas ao montante dos participantes que estudaram até o Ensino Médio. Desse modo então, nota-se que a distribuição dos participantes entre os níveis desse preditor também está satisfatoriamente equilibrada, sendo 30 migrantes baianos que possuem o Ensino Fundamental e 20 que possuem o Ensino Médio.

Apesar de poucos migrantes baianos residentes na RMSP terem dado continuidade aos estudos em cursos do Ensino Superior, há aqueles que fizeram cursos profissionalizantes de curta duração, como é o caso de AlanaS, JulioS, LeticiaC, OlgaS, IedaQ, CassiaC. Outros que nem cursaram o Ensino Médio participam de outras práticas de letramento. AmandaS, por exemplo, é engajada na política de Francisco Morato (SP); MiraldaC, MarieneS, NadilsonS, MerieleF, JucelinoS são evangélicos praticantes; LucianoA lê bastante e se denomina “amante da língua portuguesa” e GoreteS frequenta periodicamente museus e espaços culturais que oferecem oficinas de arte.

Quanto à Idade de migração, conforme se vê no histograma da Figura 6, a grande maioria dos migrantes baianos residentes na RMSP chegou à comunidade anfitriã com idade entre 15 e 25 anos, período em que, comumente, os jovens estão interessados em ingressar no mercado de trabalho, especialmente com o objetivo de conquistar a sua independência financeira. Ademais, é possível visualizar, na Figura 6, que há baianos que migraram para a RMSP quando eram crianças e tantos outros que migraram tardiamente. Assim, com base nessas informações, verifica-se que não só há certo balanceamento da frequência absoluta dos participantes da presente pesquisa segundo a Idade de migração, como também se avalia que esse preditor pode ser analisado estatisticamente como uma variável contínua.

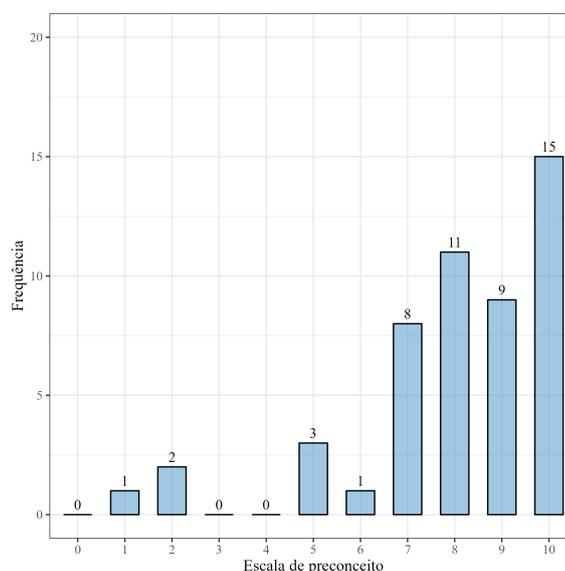
**Figura 6** – Frequência da idade de migração dos baianos residentes na RMSP participantes da amostra

Fonte: Elaboração própria.

A maior parte desses migrantes baianos saiu da zona rural, alguns migraram do nordeste da Bahia, muitas vezes, devido às consequências da seca, e outros do centro-sul, em busca, sobretudo, de melhores oportunidades de trabalho. Exceto quando migraram ainda crianças por causa da mudança da família, como LuanaD, que migrou aos seis anos, muitos deles só aguardaram atingir a maioridade para ir para São Paulo. Em algumas situações, basicamente com mulheres, a ida era antes mesmo de completar os 18 anos. GoreteS, por exemplo, aos 15 anos foi para São Paulo (SP) trabalhar como doméstica.

Assim que chegaram a São Paulo, os migrantes baianos da amostra aqui pesquisada passaram por um difícil processo de adaptação social. Além da saudade dos familiares e amigos que ficaram na terra natal, o fato de terem saído do Nordeste, uma região que carrega o estigma de ser “pobre” e “atrasada”, impactou a forma como foram recebidos pelos nativos da comunidade anfitriã. A xenofobia é uma marca que muitos desses baianos carregam na memória. Os dados da Figura 7 resultam das notas atribuídas à pergunta do questionário de identidade “de 0 a 10, quanto você acha que baiano sofre preconceito em São Paulo?”. Com base na frequência dos participantes que deram notas acima de 7, percebe-se que os migrantes baianos sofrem bastante preconceito na RMSP.

**Figura 7** – Frequência da escala de preconceito contra os migrantes baianos residentes na RMSP participantes da amostra



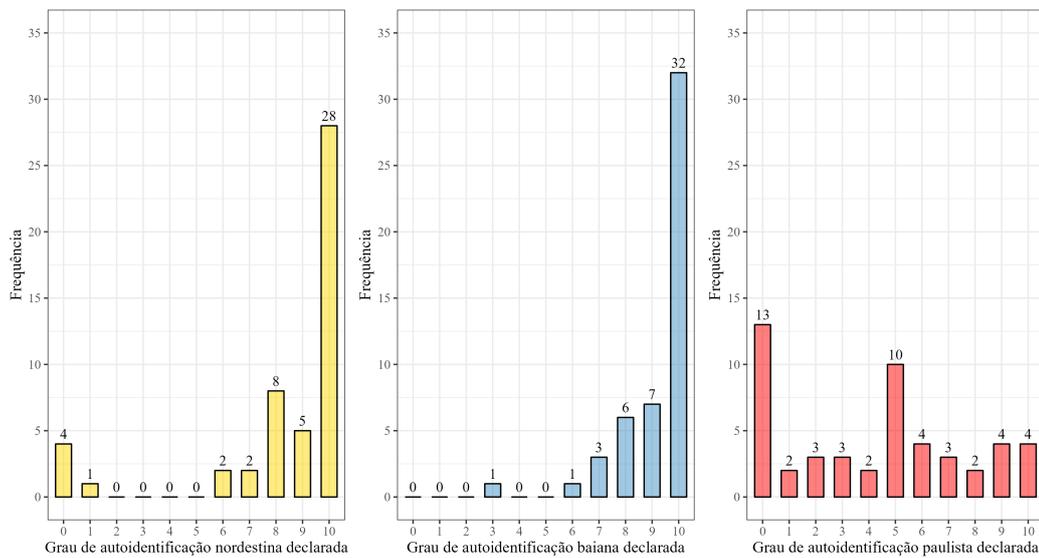
Fonte: Elaboração própria.

A autoavaliação feita acerca da xenofobia sofrida por eles foi, muitas vezes, acompanhada de relatos pessoais. As respostas dadas à pergunta “você já sofreu algum preconceito por ser baiano?” mostram que muitos deles foram vítimas de discriminação, principalmente devido à forma de falar e de se vestir, conforme se observa no depoimento de AlanaS no exemplo (4).

- (4) D1: *você já sofreu algum tipo de preconceito por ser baiana?*  
 S1: *já / muitas (AlanaS)*  
 D1: *pode dar pelo menos uma situação aqui?*  
 S1: *uma das **mais fortes** foi no **meu primeiro trabalho** né? não sei se era pelo **meu jeito de falar** ou **meu jeito de se vestir** / uma das gerentes falou que eu não servia pra estar naquele cargo e sim para lavar banheiro.*

Ressalta-se fortemente que os casos de violência mencionados pelos migrantes baianos ocorreram sobretudo no período inicial de sua chegada. No depoimento de AlanaS, em (4), por exemplo, ela enfatiza que o caso de preconceito “mais forte” foi em seu “primeiro trabalho” em São Paulo (SP). Talvez esse seja um dos muitos motivos que façam com que alguns migrantes baianos deixem de se autoidentificar como nordestinos e/ou como baianos e passem a se autoidentificar como paulistas, conforme sugerem dados explicitados nos gráficos da Figura 8.

**Figura 8** – Frequência do grau de autoidentificação nordestina, baiana e paulista declarada dos migrantes baianos residentes na RMSP participantes da amostra

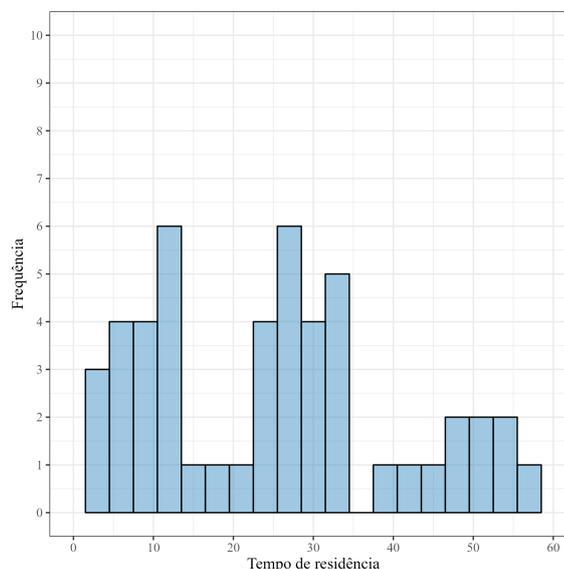


Fonte: Elaboração própria.

Nos gráficos da Figura 8, apresentam-se, respectivamente da esquerda para a direita, as frequências das respostas dadas às perguntas do questionário de identidade “de 0 a 10, quanto você se considera nordestino?”; “de 0 a 10, quanto você se considera baiano?” e “de 0 a 10, quanto você já se considera paulista?”. Vê-se que os dois primeiros gráficos têm um design semelhante, de maneira que a grande maioria dos migrantes baianos na RMSP se autodeclara bastante nordestina e baiana. No último gráfico, por sua vez, a frequência se distribui entre todas as notas, demonstrando muita variação nas respostas. Levando em conta que essa distribuição é a mais balanceada em comparação com as duas primeiras, decidiu-se que o preditor Grau de autoidentificação paulista declarada seria testado nos modelos estatísticos.

Acerca do tempo de residência dos migrantes baianos, partindo dos dados expostos na Figura 9, detecta-se que a amostra está relativamente equilibrada em relação ao número de participantes por cada ano de estadia na comunidade anfitriã. Uma vez que são raras as lacunas vistas nesse cruzamento de informações, considerou-se, nos testes estatísticos, o preditor Tempo de residência como variável contínua.

**Figura 9** – Frequência do tempo de residência dos migrantes baianos residentes na RMSP participantes da amostra



Fonte: Elaboração própria.

Grande parte dos baianos da amostra aqui analisada migrou para a RMSP na expectativa não só de trabalhar mas também de estudar. Contudo esse desejo poucas vezes pôde ser realizado devido à escassez de tempo e ao cansaço gerado pelo trabalho. Durante o tempo de estadia em São Paulo, alguns migrantes conseguiram ascender socialmente em relação ao status econômico de quando chegaram, especialmente quanto à estabilidade financeira e à aquisição da casa própria. GleideS e FredericoC, por exemplo, são funcionários públicos; CassiaC, AlanaS, LeticiaC e JosafaF são pequenos empresários engajados no bairro onde moram; os demais ou estão ativos no mercado de trabalho (formal ou informal) ou são aposentados.

Antes apresentar a rede social de cada participante, é pertinente salientar que o termo “rede social”, neste trabalho, está sendo empregado para situar o método de abordagem dos participantes, conforme visto na Subseção 3.1, mas também para reportar a todas as pessoas do convívio dos migrantes baianos. No que diz respeito à rede social de cada participante da amostra, como foi explicado na Subseção 3.4, considerou-se a proveniência das pessoas com as quais os migrantes baianos residentes na RMSP mantêm contato. Na Tabela 1, reporta-se o número de participantes de acordo com a proveniência (bairanas, paulistas ou de outros lugares) das pessoas citadas na resposta à pergunta sociométrica “quais são as dez pessoas com as quais você mais convive?”. A título de exemplo, o valor 3, na terceira linha da segunda coluna, indica que, dos 50 migrantes baianos, apenas três disseram conviver com nenhum baiano na RMSP. Por outro lado, o valor visto na segunda linha da sexta coluna sugere que 12 deles disseram conviver na RMSP com quatro baianos. Na 13ª coluna, mostra-se o total de participantes sem resposta (S.R.) e, portanto, que não foram incluídos nas análises em que se testaram as corre-

lações com este preditor. Ao contrastar a distribuição dos participantes segundo a origem de seus contactantes, verificou-se que a quantidade de contactantes baianos é mais distribuída em comparação com os paulistas e com as pessoas de outros lugares. Dessa maneira, para os testes estatísticos, analisou-se o preditor Rede de contato com baianos, que também foi considerado como uma variável contínua. É oportuno salientar que boa parte dos paulistas mencionados nas respostas dadas à pergunta sociométrica corresponde aos filhos dos próprios migrantes baianos.

**Tabela 1** – Distribuição dos migrantes baianos residentes na RMSP participantes da amostra segundo a proveniência dos integrantes de suas redes sociais

Proveniência dos contactantes	Número de contactantes											S.R.
	0	1	2	3	4	5	6	7	8	9	10	
Baianos	3	1	7	6	12	1	5		2	3	6	3
Paulistas	19	2	10	2	5	2	3	2		1		3
Outros nordestinos	30	10	1	3	2							3
Outros sudestinos	44	1	1									3
Outros lugares	44	1			1							3

Fonte: Elaboração própria.

Os dados encontrados sobre a proveniência das pessoas com as quais os migrantes têm contato na RMSP sustentam a ideia de que o processo de migração tem o auxílio da família ainda no lugar de origem e a ajuda de alguém (parentes e/ou amigos) que já mora na comunidade anfitriã, fornecendo o apoio necessário para a instalação deles (Fontes, 2008). Os grafos das redes 01 (ver a Figura 2) e 02 (ver a Figura 3) são formados por nichos familiares e/ou de conterrâneos e sugerem essas *redes de apoio*. Elas, num contexto de migração, têm papéis fundamentais no processo de instalação na comunidade anfitriã, porque são capazes de acolher novos migrantes e ampará-los em qualquer necessidade durante a estadia (Fontes, 2008). A vivência de AlanaS, em (5), é um exemplo. É relevante comentar que, com a possibilidade de fazer videochamadas nos *smartphones*, os migrantes baianos residentes na RMSP têm mantido contato diariamente com os seus parentes e amigos que permaneceram na Bahia.

- (5) D1: *hoje você mora somente com L., né? mas você falou que morou já com as suas primas*  
 S1: *isso (AlanaS)*  
 D1: *foi você que veio primeiro depois elas?*  
 S1: *eu vim primeiro e fui trazendo a leva toda*  
 D1: *entendi / quem veio primeiro?*  
 S1: *eu*

D1: *não / depois que você veio / você trouxe /*

S1: *ai depois veio a M. / depois veio a E. / minha irmã / depois veio a C. que é irmã da M. e depois veio a J.*

D1: *entendi / a sua adaptação aqui foi muito difícil foi fácil? você teve um acolhimento de outras pessoas?*

S1: *ah! não foi fácil / mas assim tive ajuda de bastante pessoas / tipo minha tia né?... primeiro lugar foi minha tia / ela que abriu as portas assim pra tudo / morei com ela um mês / depois acabou num dando muito certo / eu fui e aluguei um quarto com as menina*

### 3.5 Manipulação e análise de dados

Os áudios da amostra de fala dos migrantes baianos residentes na RMSP têm entre 00h30min e 01h40min. As gravações acima de 50' são entrevistas feitas com mais de um participante. Depois de coletadas, as amostras de fala foram transcritas e codificadas quanto às variáveis sociolinguísticas no Elan (Hellwig Birgit; Geerts, 2020), um *software* desenvolvido pelo *Instituto Max Planck de Psicolinguística* que possibilita, além de outras vantagens, o alinhamento do arquivo de mídia com a transcrição/anotação (Oushiro, 2014). Essa ferramenta auxilia o pesquisador na busca automática de dados, uma vez que as transcrições podem ser feitas em diferentes trilhas e, assim, separar as falas do documentador e do informante. Além disso, esse programa pode gerar arquivos de dados com extensões para outros *softwares* utilizados em análises linguísticas, como o Praat e o R (Oushiro, 2014).

Com os dados tabulados, foi possível visualizá-los, organizá-los, codificá-los quanto aos preditores linguísticos, sociais e analisá-los estatisticamente no R (R Core, 2023). Segundo Levshina (2015), os testes estatísticos aplicados a pesquisas linguísticas de cunho quantitativo se fazem necessários a fim de livrar o pesquisador de generalizações ruidosas. Dessa maneira, as variáveis sociolinguísticas Itens lexicais, (-r) em posição de coda e Negação sentencial, por serem de natureza categórica, foram submetidas a testes de regressão logística univariada (Itens lexicais) e multivariada de efeitos mistos, com inclusão do Participante e do Item lexical como variável aleatória ((-r) em coda silábica, Negação sentencial).

Depois que se alcançaram os resultados das adaptações linguísticas dos migrantes baianos na RMSP, recorreu-se a trabalhos que analisaram essas mesmas variáveis sociolinguísticas em amostras de fala de baianos não migrantes e na de paulistas não migrantes com o objetivo de contrastar seus padrões sociais e linguísticos. Valeu-se ainda, quando encontradas, de pesquisas que estudaram essas mesmas variáveis dependentes em amostras de fala de outras populações migrantes.

### 3.6 Preditores sociais e hipóteses

Conforme dito na Subseção 3.4, almejou-se, neste estudo, obter resultados comparáveis com trabalhos que utilizaram amostras de fala de migrantes, a fim de ampliar as discussões sobre o contato dialetal. Em vista disso, testou-se a correlação das variáveis sociolinguísticas Itens lexicais, (-r) em coda e Negação sentencial, tendo como valor de referência as variantes tipicamente paulistas em relação às variantes tipicamente baianas (ver a Introdução), com os preditores sociais, testados na maioria das pesquisas mencionadas na Seção 2, fundamentando-se nas seguintes hipóteses:

(i) A respeito de Sexo/Gênero, no estudo de Oushiro (2020a), em que se analisaram múltiplas variáveis numa amostra de fala de migrantes alagoanos e paraibanos residentes na Região Metropolitana de Campinas (SP), verificou-se que esse preditor se correlacionou com as variáveis /t, d/ diante de [i] e a Concordância nominal, de modo que as mulheres tenderam a usar mais as variantes palatal e concordância padrão em relação aos homens. Porém ele não se correlacionou com a Altura das vogais médias pretônicas /e/ e /o/, nem com o (-r) em posição de coda nem com a Negação sentencial. No estudo de Possatti (2020), em que se investigou a não palatalização do /s/ em coda de migrantes cariocas residentes em João Pessoa (PB), viu-se que as mulheres tenderam a se adaptar ao padrão da comunidade anfitriã, isto é, à forma não palatal. Os resultados de Figueiredo (2014) sobre a fala de migrantes gaúchos em Mato Grosso, em cujo estudo também se analisou um conjunto de lexias, mostraram que as mulheres tenderam a empregar mais as variantes menos gaúchas [-RS] do que os homens. Já na pesquisa de Meurer (2022), contudo, as mulheres tenderam a utilizar mais as variantes mais gaúchas [+RS]. Por outro lado, no trabalho de Bieler da Silva (2015), no qual se analisou a realização do (-r) em coda da fala de Itanhanduenses, notou-se que os homens que estão há mais tempo fora de Itanhandu (MG) tendem a empregar mais o Tepe em relação às mulheres. Assim, com respaldo nesses dados, aventa-se a hipótese de que as mulheres se adaptam mais às variantes típicas de São Paulo (SP) em comparação com os homens, malgrado os diferentes níveis linguísticos das variáveis sociolinguísticas.

(ii) Em relação à Idade, na pesquisa de Martins (2008) sobre a palatalização das oclusivas dentais /t, d/ diante de [i] na fala de paraibanos residentes no Rio de Janeiro (RJ), notou-se a correlação dessa variável dependente com a Idade, de modo que os migrantes paraibanos com até 25 anos tendem a empregar a forma palatal em relação dos migrantes com idade acima de 26 anos. Na investigação de Oliveira (2020) acerca do uso do Retroflexo na fala de migrantes baianos em Bauru (SP), observou-se que os migrantes baianos com idade entre 38 e 44 anos tendem a empregar mais o Retroflexo do que os migrantes com idade entre 22 e 29 anos e entre 52 e 59 anos. No estudo de Chacon (2012) quanto à palatalização do /s/ diante de /t/ e /d/ na fala de migrantes paulistas radicados em João Pessoa (PB), constatou-se que os migrantes paulistas com idade entre 19 e 25 anos tendem a usar mais a variante palatal em relação aos

migrantes que têm mais de 30 anos de idade. Nos resultados de [Possatti \(2020\)](#), observou-se que a não palatalização do /s/ tende a acontecer mais na fala dos migrantes cariocas com idade acima de 30 anos em comparação com os migrantes com idade entre 18 e 29 anos. Na pesquisa de [Figueiredo \(2014\)](#), o GII (grupo dos mais velhos) tendeu a conservar mais o uso das lexias da comunidade matriz [+RS] em relação ao GI (grupo dos jovens). Por outro lado, os resultados de [Meurer \(2022\)](#) evidenciaram que o GII tendeu a conservar mais as variantes [+RS], mas o GI adotou um léxico comum às duas comunidades (matriz e novo meio), com uma sutil tendência para o uso dos itens lexicais típicos do novo meio. Ao que parece, com base nos resultados das correlações de variáveis dependentes com a idade do migrante, os mais jovens tendem a se adaptar mais às variantes típicas da comunidade anfitriã em comparação com os migrantes mais velhos. Dessa maneira, conjectura-se que os migrantes baianos mais jovens residentes na RMSP usam mais as variantes tipicamente paulistas do que os migrantes baianos mais velhos.

(iii) Acerca do Nível de escolaridade, [Alves \(1979\)](#), em seu estudo sobre as atitudes linguísticas de pernambucanos e baianos residentes em São Paulo, aventou a hipótese de que os migrantes de baixa escolarização se adaptam mais aos aspectos linguísticos da comunidade anfitriã do que os mais escolarizados. No estudo de [Santana \(2023\)](#), em que se analisaram a palatalização de /t, d/ diante de [i] e a Negação sentencial na fala de migrantes sergipanos residentes em São Paulo (SP), observou-se a correlação de /t, d/ com o Nível de escolaridade em interação com a Idade de migração. O resultado de tal interação sugere não haver diferença entre os migrantes sergipanos que possuem o Ensino Fundamental I, independentemente se migraram mais cedo ou mais tarde. Em contrapartida, os migrantes sergipanos residentes em São Paulo (SP) que possuem o Ensino Fundamental II ou o Ensino Médio tendem a empregar mais a variante palatal se eles tiverem migrado mais cedo. Na pesquisa de [Oliveira \(2020\)](#), verificou-se que o Retroflexo, na fala dos migrantes baianos residentes em Bauru (SP), tende a ser mais empregado pelos migrantes que possuem o Ensino Médio do que os têm Ensino Superior. No trabalho de [Oushiro \(2020c\)](#) sobre a realização de múltiplas variáveis na fala de migrantes alagoanos e paraibanos radicados em Campinas (SP), notou-se que o Nível de escolaridade não se correlacionou com a Altura das vogais /e/ e /o/ nem com /t, d/ diante de [i]. Por outro lado, esse preditor social se correlacionou com: o (-r) em posição de coda, de maneira que o Tepe e o Retroflexo tenderam a ser mais empregados pelos migrantes de Ensino Médio em comparação com os de Nível Superior; a Negação sentencial, de modo que o uso de Neg2/Neg3 foi favorecido por migrantes com o nível Fundamental; a Concordância nominal, sendo que os migrantes que possuem o Ensino Superior tendem a realizar mais a variante padrão em relação aos migrantes com o nível Médio. Nos dados de [Figueiredo \(2014\)](#) quanto à dimensão Diastrática, a autora observou que não havia diferença no uso das variantes [-RS] entre os mais e os menos escolarizados. Por outro lado, nos resultados de [Meurer \(2022\)](#), verificou-se que as mulheres menos escolarizadas tenderam a usar mais as variantes [-RS] em comparação com os que estu-

daram até o Ensino Médio. Souza (2019), na análise sobre o uso das formas do imperativo de migrantes baianos residentes em São Paulo, viu que os migrantes com menor nível de escolaridade tendem a empregar mais a forma imperativa com morfologia de indicativo, conforme o padrão da comunidade anfitriã. Ao contrastar os resultados obtidos das correlações de dados linguísticos, oriundos de amostras de fala de migrantes, com o Nível de escolaridade, nota-se que os migrantes com menor nível de escolaridade tendem a se adaptar às variantes típicas da comunidade anfitriã. Sendo assim, levanta-se a hipótese de que os migrantes baianos residentes na RMSP que possuem o nível de escolaridade mais baixo se adaptam mais às variantes típicas de São Paulo (SP), sem fazer distinção da natureza delas, em relação aos migrantes que têm o nível de escolaridade mais alto.

(iv) Quanto à Idade de migração, no trabalho de Oushiro (2020c), esse preditor não se correlacionou com as variáveis morfossintáticas (Negação sentencial e Concordância Nominal), mas se correlacionou com as variáveis fonético-fonológicas (Altura das vogais médias pretônicas, (-r) em coda e /t, d/ seguido de [i]), de forma que todas elas apresentaram o mesmo padrão social: quanto mais cedo os migrantes alagoanos e paraibanos chegaram a Campinas (SP), maior foi o uso das variantes típicas da comunidade anfitriã. Na pesquisa sobre a Altura das vogais médias pretônicas na fala de migrantes sergipanos residentes em São Paulo (SP), Santana (2018) obteve esse mesmo padrão, os migrantes sergipanos que chegaram mais cedo à comunidade anfitriã tenderam a empregar mais as Vogais médias pretônicas mais altas, seguindo o padrão paulistano. Em seu trabalho sobre a realização de /t, d/ seguido de [i] e da Negação sentencial nessa mesma amostra de fala, Santana (2023) notou que a Idade de migração se correlacionou apenas com a variável fonético-fonológica, cujo resultado evidenciou que os sergipanos que migraram mais cedo e que possuem o Ensino Fundamental II e Ensino Médio tenderam a usar mais a variante palatal. Na investigação de Barbosa (2022) sobre a realização do /s/ em coda na fala de alagoanos e paraibanos residentes em Campinas (SP), verificou-se que as variáveis Palatalização e Apagamento se correlacionam com a Idade de migração, cujo resultado mostrou que as variantes da comunidade anfitriã (formas alveolar e zero fonético) tendem a ser mais usadas por aqueles que migraram mais cedo em comparação aos que migraram mais tarde. Os resultados de Figueiredo (2014) e Meurer (2022), para os quais não se aplicaram testes estatísticos, revelam que os sulistas que migraram mais cedo se adaptaram às variantes do novo meio em comparação com os que migraram mais tarde. Assim, com base nesses resultados, aventa-se a hipótese de que os migrantes baianos residentes na RMSP que migraram mais cedo se adaptam mais às variantes (seja qual natureza for) da comunidade anfitriã em relação aos que migraram mais tarde.

(v) No que diz respeito ao Grau de autoidentificação (paulista declarada), conforme se viu na Seção 2, a intenção de testar esse preditor em análises linguísticas que envolvem migrantes não é aferir a identidade dos falantes, mas tentar quantificar as suas atitudes a respeito

da comunidade de origem e da comunidade anfitriã (Oushiro, 2024, no prelo[b]). Em Oushiro (2020b), observou-se a correlação do Grau de estadualidade com o (-r) em posição de coda e com (Neg), de maneira que os migrantes alagoanos e paraibanos residentes em Campinas (SP) que atribuíram notas mais elevadas para a identificação com o seu lugar de origem tenderam a não usar as variantes típicas da comunidade anfitriã em relação aos que se deram notas mais baixas. Em compensação, o Grau de paulistidade se correlacionou com /t, d/ diante de [i] e com a Concordância nominal, cujos resultados mostraram que os migrantes que se deram notas mais elevadas para a identificação com São Paulo tenderam a empregar mais as variantes típicas da comunidade anfitriã em comparação com os que apresentaram notas mais baixas. No estudo de Oliveira (2020), viu-se que o Retroflexo tendeu a ser mais empregado na fala dos migrantes baianos residentes em Bauru (SP) que demonstraram ter atitudes positivas para com a comunidade anfitriã em relação aos que manifestaram atitudes negativas. Partindo desses dados, depreende-se que os migrantes que manifestam atitudes positivas para com a comunidade anfitriã tendem a se adaptar linguisticamente a ela. Desse modo, levanta-se a hipótese de que os migrantes baianos residentes na RMSP que se autoidentificam declaradamente mais paulista usam mais as variantes (ainda que sejam variáveis de diferentes natureza) típicas da comunidade anfitriã em relação aos migrantes que se autoidentificam como menos paulista.

(vi) Sobre o Tempo de residência, nos resultados de Oushiro (2020c), notou-se que tal preditor se correlacionou com o (-r) em coda, de maneira que o Tepe e o Retroflexo tendem a ser mais usados pelos migrantes alagoanos e paraibanos que residem em Campinas (SP) há mais de 30 anos, em comparação aos que residem entre 11 e 29 anos. No estudo de Barbosa (2022) sobre o /s/ em coda, reparou-se que apenas a Palatalização se correlacionou com o Tempo de residência, cujo resultado sugere que os migrantes alagoanos e paraibanos que estão há mais tempo em Campinas (SP) tendem a empregar mais a variante típica da comunidade anfitriã. Nas pesquisas de Chacon (2012), Oliveira (2020) e Possatti (2020), observou-se esse mesmo padrão sociolinguístico. Por outro lado, nas análises de Souza (2019), viu-se que os migrantes baianos que residem entre 0-7 anos favorecem o uso do imperativo com morfologia de indicativo em relação aos que residem entre 8-15 anos ou há mais de 16 anos. Desse modo, considerando que, na maioria desses resultados, o maior tempo de estadia favorece a adaptação linguística dos migrantes quanto às variantes típicas da comunidade anfitriã, conjectura-se que os migrantes baianos que residem há mais na RMSP se adaptam mais às variantes típicas do falar paulista em relação aos que residem há menos tempo, sem fazer distinção do nível linguístico delas.

(vii) No tocante à Rede de contato (com baianos), verificou-se, no estudo de Bortoni-Ricardo (2011 [1985]) sobre a fala de mineiros rurais residentes em Brasília (DF), que os migrantes pertencentes a redes mais abertas tendem a utilizar mais as variantes da comunidade anfitriã em relação aos que pertencem a redes mais fechadas. Na pesquisa de Guedes (2019 [2017]) acerca do uso do artigo definido antes de possessivos na fala de alagoanos e paraiba-

nos residentes em Campinas (SP), verificou-se que os migrantes que convivem com paulistas tendem a empregar mais o artigo definido precedendo pronomes possessivos do que os que convivem com pessoas de mesma procedência. Nos dados de Santana (2023), esse preditor se correlacionou apenas com o uso de (Neg), de modo que Neg2/Neg3 tendem a ser menos frequentes na fala de migrantes sergipanos residentes em São Paulo (SP) que integram redes abertas em comparação com os que pertencem a redes mais fechadas. Com base nessas informações, aventa-se a hipótese de que os migrantes baianos que integram redes abertas (com poucos baianos) usam mais as variantes típicas da comunidade anfitriã em relação aos que pertencem a redes fechadas (com mais baianos).

### 3.7 Síntese

Descreveram-se, nesta Seção, os procedimentos metodológicos adotados para as análises sobre a fala de migrantes baianos residentes na RMSP. Com base em Bortoni-Ricardo (2011 [1985]) e Santana (2018), utilizou-se, como método de recrutamento dos participantes para a composição da amostra a ser investigada, o de rede social. Esse método consistiu em, a partir de perguntas sociométricas, feitas inicialmente aos âncoras e depois com cada participante da pesquisa, mapear a proveniência das pessoas com as quais os migrantes baianos convivem.

Considerando que a amostra não foi estratificada aos moldes da Sociolinguística Tradicional, mostraram-se as tomadas de decisões quanto à maneira como os dados foram manipulados e analisados estatisticamente. Depois de fazer a inspeção do balanceamento dos participantes nas categorias sociais de interesse desta pesquisa, verificou-se que os preditores Sexo/Gênero, Idade, Nível de escolaridade, Idade de migração, Grau de autoidentificação paulista declarada, Tempo de residência e Rede de contato com baianos poderiam ser analisados estatisticamente sem correr o risco de um viés. Outrossim, viu-se que tais preditores, exceto o Sexo/Gênero e o Nível de escolaridade, poderiam ser analisados nos modelos estatísticos como variáveis contínuas.

Tendo em vista a natureza categórica das variáveis respostas, realizaram-se análises univariadas e multivariadas de regressão logística de efeitos mistos com inclusão do Participantes e dos Itens lexicais como variáveis aleatórias. Para a realização dos testes estatísticos, ponderaram-se as hipóteses de que, independentemente do nível linguístico das variáveis respostas, (i) as mulheres se adaptam mais às variantes tipicamente paulistas do que os homens; (ii) os migrantes baianos residentes na RMSP mais jovens usam mais as variantes tipicamente paulistas em relação aos migrantes baianos mais velhos; (iii) os migrantes baianos residentes na RMSP que possuem o nível de escolaridade mais baixo se adaptam mais às variantes tipicamente paulistas em relação aos migrantes que têm o nível de escolaridade mais alto; (iv) os migrantes baianos residentes na RMSP que migraram mais cedo se adaptam mais às variantes da comunidade anfitriã em relação aos que migraram mais tarde; (v) os migrantes baianos resi-

---

dentem na RMSP que se autoidentificam declaradamente mais paulista usam as variantes típicas da comunidade anfitriã em relação aos migrantes que se autoidentificam como menos paulista; (vi) os migrantes baianos que residem há mais tempo em São Paulo se adaptam às variantes típicas do falar paulista em comparação com quem reside há menos tempo; (vii) os migrantes baianos que integram redes abertas (com poucos baianos) usam mais as variantes típicas da comunidade anfitriã em relação aos que pertencem a redes fechadas (com muitos baianos).

Os resultados dessas análises estão apresentados nas Seções 4, 5, 6 e 7.

## 4

# **Adaptação em nível lexical: Tangerina/Mexerica, Aipim/Mandioca, Quentinha/Marmita e Trabalho/Serviço**

Poucas ainda são as pesquisas sociolinguísticas sobre a variação lexical nas variedades do português brasileiro. A grande maioria dos estudos dessa área tem se debruçado especialmente sobre a variação de fenômenos do nível fonético-fonológico e do nível morfossintático. Ao que parece os trabalhos que analisam as adaptações lexicais causadas pelo contato entre dialetos mutuamente inteligíveis não existem, conforme visto na Seção 2.

A fim de preencher essa lacuna, o estudo apresentado nesta Seção teve como objetivo investigar a adaptação lexical dos migrantes baianos residentes na RMSP a partir do uso das variáveis sociolinguísticas Tangerina/Mexerica, Aipim/Mandioca, Quentinha/Marmita e Trabalho/Serviço. A expectativa para essa investigação foi a de que essa população migrante utilizaria mais os itens lexicais tipicamente paulistas “Mexerica”, “Mandioca”, “Marmita” e “Serviço” em resposta às questões semântico-lexicais (QSL) utilizadas nas entrevistas sociolinguísticas (ver o Apêndice A) no lugar das lexias classificadas como tipicamente baianas “Tangerina”, “Aipim”, “Quentinha” e “Trabalho”, respectivamente. Além desse objetivo, almejou-se também descobrir como se dá o encaixamento social dos Itens lexicais paulistas na fala dessa população migrante. Para tanto, utilizaram-se modelos de regressão logística com a finalidade de testar a correlação dos Itens lexicais com os preditores sociais reportados na Seção 3.

Esta Seção assim se subdivide: descreve-se, na Subseção 4.1, a variação lexical em alguns dialetos do português brasileiro, focalizando o uso das variáveis Tangerina/Mexerica, Aipim/Mandioca, Quentinha/Marmita e Trabalho/Serviço nos dialetos baiano e paulista; na Subseção 4.2, detalham-se os procedimentos metodológicos usados para coletar, manipular e analisar estatisticamente os dados obtidos com as entrevistas sociolinguísticas; na Subseção 4.3, apresentam-se e discutem-se os resultados encontrados a partir das análises descritivas de cada

item lexical bem como dos modelos de regressão logística; por fim, na Subseção 4.4, resumem-se os achados desse estudo.

#### **4.1 O uso variável de Tangerina/Mexerica, Aipim/Mandioca, Quentinha/Marmita e Trabalho/Serviço no português brasileiro: Bahia e São Paulo em destaque**

O *Atlas Linguístico do Brasil – ALiB* (Cardoso et al., 2014) tem sido o principal responsável pelo mapeamento dos usos lexicais no território brasileiro. Para a constituição de sua amostra, o projeto, com base na Geolinguística Pluridimensional, utilizou como critério para a seleção dos participantes o de que eles deveriam ser falantes prototípicos da localidade analisada, isto é, eles e seus pais deveriam ter nascido naquela comunidade, onde também deveriam ter passado pelo menos dois terços de suas vidas (Cardoso et al., 2014). A abordagem do Projeto ALiB rendeu uma amostra com 1.100 participantes de 250 pontos do território nacional (sendo quatro de cidades do interior e oito das capitais de 25 estados) estratificados quanto ao gênero (550 homens e 550 mulheres), à faixa etária (faixa I, 18-30 anos e faixa II, 50-65 anos), e, nas capitais, à escolaridade (Fundamental e Universitária). Estritamente para a coleta dos dados lexicais, durante as entrevistas de caráter sociolinguístico (Labov, 2008 [1972]), os pesquisadores utilizaram perguntas do questionário semântico-lexical, cujo instrumento consta de 202 questões distribuídas em 14 campos semânticos (Cardoso et al., 2014, p. 126).

As recolhas do Projeto ALiB têm permitido comparar o uso dos itens lexicais no Brasil. Além dos trabalhos monográficos produzidos a partir dessa base de dados (ver o site <https://alib.ufba.br/>), o *Atlas Linguístico do Brasil* (Cardoso et al., 2014), publicado em 2014, apresenta cartografias regionais a respeito dessa distribuição. Para efeito de análise, retomam-se aqui dois casos de variação lexical diatópica pesquisados pelo ALiB, a saber, o uso variável de Tangerina/Mexerica e o de Aipim/Mandioca entre os estados da Bahia e de São Paulo. Ambas as variáveis pertencem ao campo semântico de Atividades agropastoris, cujas questões podem ser localizadas, respectivamente, no QSL 39 e 50.

A depender da área dialetal, as variáveis lexicais podem apresentar variantes formadas por significantes diferentes. As Cartas L05b e L05c do *Atlas Linguístico do Brasil* (Cardoso et al., 2014, p. 425–427) listam as denominações para a “fruta que se parece com uma laranja, que se descasca com as mãos e deixa um cheiro forte nelas” registradas na Região Nordeste e Sudeste, nesta mesma ordem. Os dados encontrados da variável Tangerina/Mexerica evidenciam que, em Salvador (BA), há registros de “Tangerina” e “Mexerica”, enquanto em São Paulo (SP) constam as variantes “Mexerica”, “Tangerina”, “Poncã”, “Carioquinha” e “Maricote”. Aguilera (2009), com base numa amostra de respostas de 200 participantes que integram o Projeto ALiB, mapeou a diatopia dos usos de Córrego/Riacho, Tangerina/Mexerica, Galinha d’Angola/Guiné/Cocar, Meleca/Tatu, Toco de cigarro e Bala. Especificamente quanto ao uso de Tangerina/Mexerica, a autora mostra que “Tangerina” é predominante na Bahia e que “Mexerica” é

mais recorrente em São Paulo. Dito de outra maneira, Bahia e São Paulo compartilham do uso das lexias “Tangerina” e “Mexerica” para se referirem à mesma coisa, porém, neste estado, predomina o uso de “Mexerica”, e naquele o emprego de “Tangerina”.

Esses estados, contudo, não compartilham das mesmas variantes quando utilizam a variável Aipim/Mandioca. A Carta L08 (Cardoso et al., 2014, p. 439) enumera as variantes utilizadas para nomear a “raiz branca por dentro coberta com uma casca marrom que se cozinha para comer” nas capitais brasileiras. De acordo com os dados cartografados, em Salvador (BA) emprega-se “Aipim”, já em São Paulo (SP) utiliza-se “Mandioca”. Romano e Cá (2020), a partir de uma amostra de respostas de 316 participantes do Projeto ALiB, residentes de 79 municípios da Região Sudeste, analisou o emprego das lexias “Mandioca”, “Aipim” e “Macaxeira” a partir das respostas dadas ao QSL 50. Os resultados totais desse estudo revelaram que, em São Paulo, há 93,75% de uso do item “Mandioca”, 2,50% de “Aipim”, 3,75% de outras variantes (“Mandioca vassourinha”, “Mandioca amarela”, “Mandioca pão”, “Mandioca cacau” e “Macaxeira”) e 0% de não respostas. Em Minas Gerais, há 93,62% de “Mandioca”, 1,06% de “Aipim”, 1,06% de outras variantes e 4,26% de não respostas. No Rio de Janeiro, os dados mostraram que há 32,39% de “Mandioca”, 64,79% de “Aipim”, 0% de outras variantes e 2,82% de não respostas. No Espírito Santo, o uso de “Mandioca” foi de 39,13%, de “Aipim” foi de 56,52%, 0% de outras variantes e 4,35% de não respostas.

É pertinente ressaltar que o item “Mandioca” apresenta variação de significado em diferentes áreas dialetais. De acordo com os dados apresentados na Carta L10 (Cardoso et al., 2014, p. 441), alcançados com o QSL 51, as variantes para designar “a raiz que não serve para comer e se rala para fazer farinha (polvilho, goma)” são “Mandioca”, “Mandioca brava” e “Macaxeira brava”, de maneira que, em Salvador (BA), usa-se “Mandioca” e, em São Paulo (SP), empregam-se “Mandioca” ou “Mandioca brava”. Sendo assim, conclui-se que, na capital baiana, emprega-se “Aipim” para a raiz comestível e “Mandioca” para o tubérculo utilizado para fazer farinha. Por outro lado, na capital paulista, usa-se “Mandioca” para ambos os tipos de raiz.

Muitos dados lexicais coletados pelo ALiB ainda carecem ser descritos e outras lexias que têm distinção diatópica precisam ser levantadas. Durante a sua estadia em São Paulo, este pesquisador notou que as variáveis Quentinha/Marmita e Trabalho/Serviço parecem apresentar diferenças dialetais. Por causa da ausência de trabalhos sobre essas variáveis, coletaram-se dados da Bahia e de São Paulo a fim de comparar com os resultados dos migrantes baianos. Com a impossibilidade de fazer a recolha presencialmente, devido às determinações sanitárias durante a pandemia da Covid-19, replicou-se o questionário eletrônico utilizado por [Figuereido e Souza \(2017\)](#) para angariar dados sobre a variação do modo imperativo de baianos migrantes em São Paulo e de baianos não migrantes. O questionário foi reelaborado no *Google Forms* e consta de 23 perguntas, sendo 4 sentenças-alvo e 19 distratoras (ver o Apêndice C). Cada

sentença-alvo teve por objetivo permitir que o participante escolhesse uma das variantes das variáveis Tangerina/Mexerica, Aipim/Mandioca, Quentinha/Marmita e Trabalho/Serviço.

No período entre maio e junho de 2020, publicou-se o formulário nas mídias sociais (*Facebook, Instagram e WhatsApp*) a fim de recrutar participantes. Com as postagens, conseguiu-se a participação de 150 respondentes, sendo 90 de cidades da Bahia e 60 de cidades de São Paulo, entre homens (total de 51) e mulheres (99) com diferentes níveis de escolaridade, Ensino Fundamental (1), Ensino Médio (51) e Ensino Universitário (98). Cabe destacar que, ao contrário do critério utilizado pelo ALiB, o questionário não se restringiu a falantes prototípicos de cada localidade, embora se tenha controle das informações de onde os participantes e seus pais nasceram e onde moram. Os resultados globais da análise dessa amostra evidenciam dissimilitudes dialetais entre a Bahia e São Paulo quanto à escolha de cada item lexical, conforme se verifica na Tabela 2.

**Tabela 2** – Frequência e proporção do emprego das variáveis lexicais Tangerina/Mexerica, Aipim/Mandioca, Quentinha/Marmita, Trabalho/Serviço nas variedades baiana e paulista

Variáveis	Variantes	BA		SP	
		N (90)	%	N (60)	%
Tangerina/Mexerica	Tangerina	87	97%	5	8%
	Mexerica	3	3%	55	92%
Aipim/Mandioca	Aipim	80	89%	0	0%
	Mandioca	10	11%	60	100%
Quentinha/Marmita	Quentinha	77	85%	1	2%
	Marmita	13	15%	59	98%
Trabalho/Serviço	Trabalho	88	98%	50	83%
	Serviço	2	2%	10	17%

Fonte: Elaboração própria.

De acordo com a Tabela 2, para o emprego de Tangerina/Mexerica na Bahia, houve 97% (87/90) de respostas para “Tangerina” e 3% (3/90) para “Mexerica”. Por seu turno, em São Paulo, houve 8% (5/60) de “Tangerina” e 92% (55/60) de “Mexerica”. Quanto à variável Aipim/Mandioca, na Bahia, o item “Aipim” teve 89% (80/90) e “Mandioca” 11% (10/90)<sup>20</sup>; já

<sup>20</sup> Das dez respostas com essa variante, quatro foram de pessoas que nasceram em São Paulo e as outras seis foram de pessoas que convivem com elas. Esse dado sugere que possivelmente os migrantes paulistas residentes na Bahia utilizam não só as formas da comunidade de origem como também estão exercendo o papel de propagadores da lexia “Mandioca” com o valor semântico de “Aipim”.

em São Paulo, a utilização de “Mandioca” foi categórica (100%). Acerca da variável Quentinha/Marmita, na Bahia, o item “Quentinha” apresentou 85% (77/90) de aplicação e “Marmita” 15% (13/90). Em São Paulo, “Quentinha” teve 2% (1/60) de uso e “Marmita” 98% (59/60). Quanto à variável Trabalho/Serviço, na Bahia, o item “Trabalho” apresentou 98% (88/90) de realização e o item “Serviço” teve 2% (2/90), enquanto em São Paulo houve 83% (50/60) da utilização de “Trabalho” e 17% (10/60) de uso do item “Serviço”.

## 4.2 Manipulação dos dados

Devido à dificuldade de coletar lexias intercambiáveis numa amostra de fala semiespontânea, apanharam-se os dados dos itens lexicais dos migrantes baianos residentes na RMSP a partir do questionário semântico-lexical aplicado a eles nas entrevistas sociolinguísticas (ver o Apêndice A). A expectativa para essa análise foi a de que essa população migrante utilizaria como primeira resposta às questões semântico-lexicais os itens “Mexerica”, “Mandioca”, “Marmita” e “Serviço”, variantes mais comuns em São Paulo, em vez de “Tangerina”, “Aipim”, “Quentinha” e “Trabalho”, respectivamente, variantes mais frequentes na Bahia.

Para as análises estatísticas, levou-se em conta apenas a primeira resposta de cada pergunta, porque quando o participante falava mais de uma lexia para a mesma pergunta, comumente, era numa análise metalinguística. Ou seja, os migrantes respondiam com o item lexical  $x$  e relatavam a existência de  $y$  e  $z$ . Isso foi recorrente no QSL dos itens Tangerina/Mexerica e Aipim/Mandioca, sugerindo que essas variáveis estão acima do nível de consciência dos falantes. As ausências de respostas foram computadas, mas não entraram na análise. Assim, dos 50 participantes, um não soube responder à pergunta sobre Aipim/Mandioca; duas pessoas não responderam acerca de Quentinha/Marmita e três sobre Trabalho/Serviço; onze migrantes responderam outros itens para esta variável, que também não foram contabilizados nas análises estatísticas. Desse total, oito participantes falaram o verbo “trabalhar”, um disse “ir para a obra”, um falou que “ia para a igreja” (o que é pastor) e outro que “ia para o trampo”. Logo se avalia que essa questão apresenta ruído e, para estudos futuros, precisará ser reelaborada. Depois dessa filtragem, restaram 183 dados de Itens lexicais (baianos e paulistas).

Após analisar cada item de maneira quantitativa, juntaram-se os itens lexicais baianos e os itens lexicais paulistas e testou-se, a partir de modelos de regressão logística de efeitos fixos, a correlação dos Itens lexicais com os preditores sociais, segundo as hipóteses arroladas na Seção 3. Escolheu-se esse tipo de tratamento estatístico devido à impossibilidade de testar os preditores contínuos com o teste de *qui-quadrado*. Aplicaram-se testes univariados, porque a quantidade de dados não suportaria um modelo mais complexo, como a análise multivariada, por exemplo.

### 4.3 Resultados

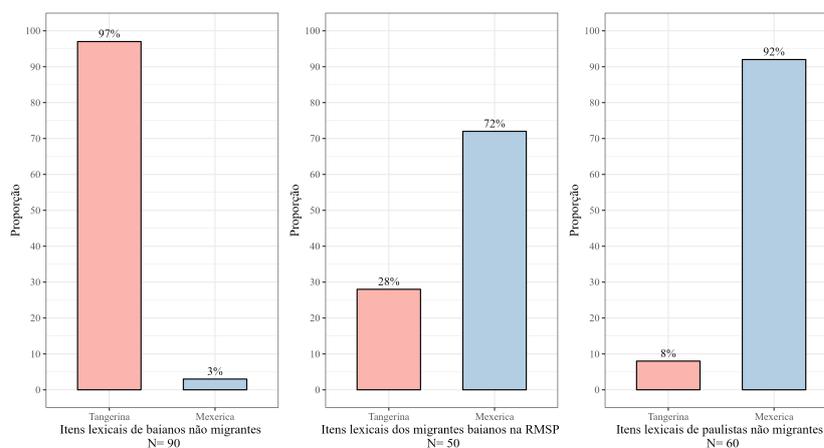
Esta Subseção está dividida em duas partes. Na primeira, descrevem-se as análises quantitativas dos Itens lexicais e, na segunda, apresenta-se a interpretação de seus resultados.

#### 4.3.1 Descrição dos achados

Primeiramente, com o auxílio de gráficos, comparam-se as taxas de uso das variantes das variáveis Tangerina/Mexerica, Aipim/Mandioca, Quentinha/Marmita e Trabalho/Serviço dos migrantes baianos residentes na RMSP com os dados dos baianos não migrantes e com os dados dos paulistas não migrantes, ambos coletados com o questionário *online* (ver Subseção 4.1). Na sequência, descrevem-se os coeficientes dos modelos de regressão logística de efeitos fixos nos quais se testou a correlação dos Itens lexicais com os preditores sociais. Por fim, faz-se a análise dos achados.

Em relação ao uso da variável Tangerina/Mexerica, os valores da Figura 10 mostram que, na Bahia (gráfico à esquerda), o uso de Tangerina para designar a “fruta menor que a laranja, que se descasca com a mão e, normalmente, deixa um cheiro na mão” é semicategórico, com 97% de aplicação em relação ao item Mexerica, com 3%. Inversamente, na variedade paulista, o emprego de Mexerica é de 92% e o de Tangerina é de 8% (gráfico à direita). Na fala dos migrantes baianos na RMSP, observa-se que a realização de Mexerica é maior, 72% (36 das 50 ocorrências) em comparação com Tangerina, 28% (14/50) (gráfico do meio). Desse modo, nota-se que os migrantes baianos na RMSP tendem a utilizar mais a lexia Mexerica no lugar de Tangerina, como se previu.

**Figura 10** – Comparação das proporções dos itens lexicais Tangerina e Mexerica de baianos não migrantes, de migrantes baianos na RMSP e de paulistas não migrantes



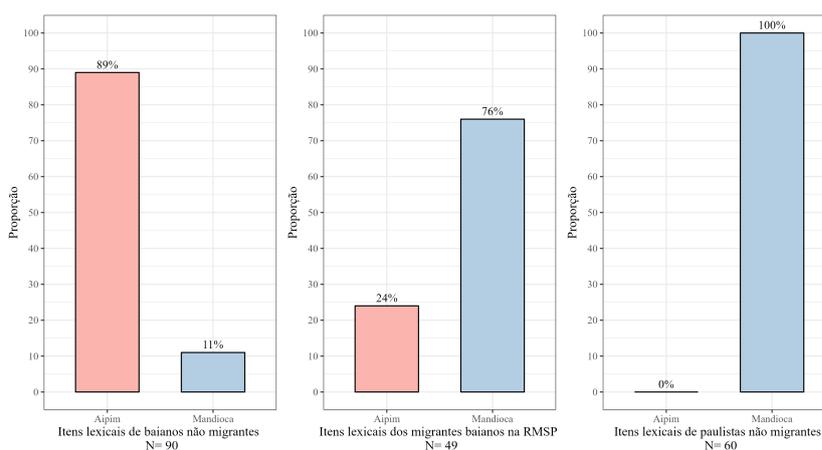
Fonte: Elaboração própria.

Ao que tudo indica, a variação diatópica do uso de Tangerina e Mexerica é algo saliente para o migrante baiano, porque nove dos 50 participantes fizeram comentários metalinguísticos a respeito disso. Tais comentários, como em (6), mencionam que a lexia “Tangerina” é mais comum na Bahia e “Mexerica” é mais recorrente em São Paulo.

- (6) D1: *M. como é que chama aquelas frutas que são menores que uma laranja que a gente descasca com as mãos e que normalmente deixa um cheiro nelas?*  
 S1: *onde eu moro? (MeireV)*  
 D1: *não aqui em São com/ como você usa aqui em São Paulo?*  
 S1: *eh tangerina mexerica*  
 D1: *aqui você usa qual?*  
 S1: *mais mexerica*  
 D1: *e lá?*  
 S1: *tangerina*

Sobre os resultados da variável Aipim/Mandioca, as taxas do gráfico à esquerda da Figura 11 indicam que, na Bahia, é mais comum empregar a lexia Aipim (89%) para se referir “à raiz comestível branca coberta por uma casca marrom”. Já em São Paulo, segundo o gráfico à direita, Mandioca é categoricamente (100%) utilizada para designar essa mesma raiz. Os migrantes baianos residentes na RMSP, de acordo com o gráfico do meio, por sua vez, empregaram 76% (37/49) da lexia Mandioca e 24% (12/39) da variante Aipim para nomear esse tipo de tubérculo, conforme já era previsto.

**Figura 11** – Comparação das proporções dos itens lexicais Aipim e Mandioca de baianos não migrantes, de migrantes baianos na RMSP e de paulistas não migrantes



Fonte: Elaboração própria.

Esse caso de adaptação merece destaque por duas razões. A primeira é que, assim como a variável Tangerina/Mexerica, a variação diatópica entre Aipim e Mandioca parece ser

também saliente, já que oito participantes fizeram comentários metalinguísticos, semelhantes ao proferido por JuvenalG no exemplo (7).

- (7) D1: *e como é que chama aquela raiz que é branca por dentro coberta por uma casca marrom que a gente cozinha pra comer?*  
 S1: *aqui é mandioca lá no nordeste é aipim (JuvenalG)*  
 D1: *mas aqui você usa mais o quê?*  
 S1: *já aqui tem que chamar de mandioca né porque senão*  
 D1: *mandioca né*  
 S1: *se eu chegar na fila eu zoo com os caras que eu conheço me dá um aipim aí/ que eu já conheço os cara da feira ali/ a feira que é de (xxx) né foi dia de sábado*  
 D1: *mas se não conhece é mandioca?*  
 S1: *é/ vai ser mandioca mesmo*

A segunda razão para o destaque é porque há diferença semântica do item Mandioca entre Bahia e São Paulo, cuja distinção também está na consciência do migrante baiano, como se observa no exemplo (8). Como se viu na Subseção 4.1, de acordo com a *Carta L10* do ALiB (Cardoso et al., 2014, p. 441), em Salvador (BA), “Mandioca” é o termo utilizado para designar “a raiz venenosa usada apenas na produção de farinha e ração animal” e, em São Paulo (SP), é empregado tanto para a raiz venenosa quanto para a “raiz não venenosa consumida frita, assada ou cozida e de que também se fazem doces e bolos”.

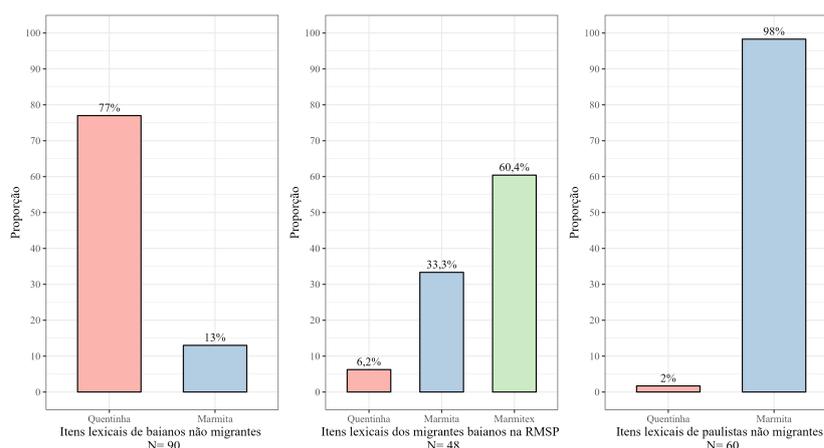
- (8) D1: *como se chama aquela raiz branca coberta por uma casca marrom que se cozinha pra comer?*  
 S1: *mandioca (LucineideE)*  
 S3: *aipim/ não mandioca (VitóriaP)*  
 S2: *mandioca (GabrielaE)*  
 S3: *é mandioca/ mas é aipim né?*  
 S2: *mandioca*  
 S1: *nós usa mandioca / que nós já conhece de lá/ então mandioca é também né?*  
 S2: *é mandioca*  
 S3: *é mandioca mas lá tem aipim também não tem?*  
 S4: *tem aipim / é porque já lá... aqui é mandioca lá é aipim (NadilsonS)*  
 S2: *lá é outra né?*  
 S1: *é aipim?*  
 S3: *é*  
 S4: *falar mandioca lá pensa que é a mandioca de fazer farinha*  
 S1: *aquela lá*  
 S2: *é isso mesmo/ que se você chegar lá dizendo quero mandioca vai achar que é a mandioca de fazer farinha*

S1: *pra mim tudo é mandioca*

No diálogo de GabrielaE, VitóriaP, LucineideE e NadilsonS, em (8), os participantes reproduzem a diferença semântica dos itens Aipim/Mandioca nas variedades baiana e paulista, conforme fora apresentado na *Carta L10 ALiB* (Cardoso et al., 2014). A resposta de LucineideE e de GabrielaE à pergunta dialetológica é “Mandioca”, já a de VitóriaP é “Aipim”. NadilsonS faz uma comparação dialetal diferenciando semanticamente o uso do item “Mandioca” em São Paulo e na Bahia. Segundo ele, “*aqui*” (São Paulo) é “Mandioca” e “*lá*” (Bahia) é “Aipim”. Ele reitera explicando que “*falar “Mandioca” lá pensa que é a “Mandioca” de fazer farinha*”. GabrielaE ratifica a análise dele, mas, para LucineideE, “Mandioca” é uma lexia genérica que serve para nomear tanto a raiz tóxica quanto a raiz comestível, conforme o significado desse item em São Paulo.

Quanto à variável Quentinha/Marmita, confrontando o gráfico à esquerda da Figura 12 com o da direita, observa-se que, na Bahia, o termo mais usado para nomear a “porção de comida que compramos, normalmente num recipiente laminado ou plástico, num restaurante ou em lugares afins quando não temos tempo de cozinhar” é “Quentinha”, com 77% de aplicação, e, em São Paulo, é “Marmita”, com 98% de uso. Nos dados dos migrantes baianos (gráfico do meio), porém, nota-se o uso de três variantes: Quentinha, com 6,2% (3/48) de uso; Marmita, com 33,3% (16/48); e Marmitex, com 60,4% (29/48) de aplicação. Desse modo, verifica-se que, no lugar de Quentinha, os migrantes baianos na RMSP tendem a empregar mais as lexias Marmita e/ou Marmitex.

**Figura 12** – Comparação das proporções da realização dos itens lexicais Quentinha e Marmita de baianos na Bahia, de migrantes baianos na RMSP e de paulistas não migrantes



Fonte: Elaboração própria.

Assim como o item Mandioca, a variação de Quentinha e Marmita chama atenção, porque, na Bahia, a lexia Marmita tem outros significados. É muito comum nesse estado em-

pregar Marmita para nomear a vasilha com compartimentos utilizada para transportar alimentos ou também a comida que é levada de casa para outros lugares, como para o trabalho, por exemplo<sup>21</sup>. É oportuno comentar ainda que o item Marmitex é bastante comum na variedade paulista equiparado ao dialeto baiano. Nessa variedade, até 2018, essa lexia era raramente ou quase nunca empregada. No entanto, recentemente, verifica-se que ela tem se propagado, sobretudo depois da pandemia da Covid-19, por meio dos aplicativos usados para pedir comida, como o *Ifood*, *Quero Delivery* e outros. Sendo assim, para trabalhos futuros, sugere-se que esse item seja examinado com mais detalhes na Bahia.

Outro dado que precisa ser destacado é que a variação diatópica de Quentinha/Marmita/Marmitex é um caso de variação que parece estar acima do nível da consciência. Durante as entrevistas, apenas um participante comentou a diferença dialetal dessas variantes entre o Nordeste e São Paulo. MarioS, em (9), diz que, “no nordeste”, as pessoas usam “Quentinha”, mas em São Paulo (SP), os paulistanos chamam “Marmitex”.

(9) D1: *como é que se chama aquela porção de comida que a gente compra eh normalmente que vem no recipiente de alumínio ou de plástico e que a gente compra normalmente quando não tem tempo pra cozinhar?*

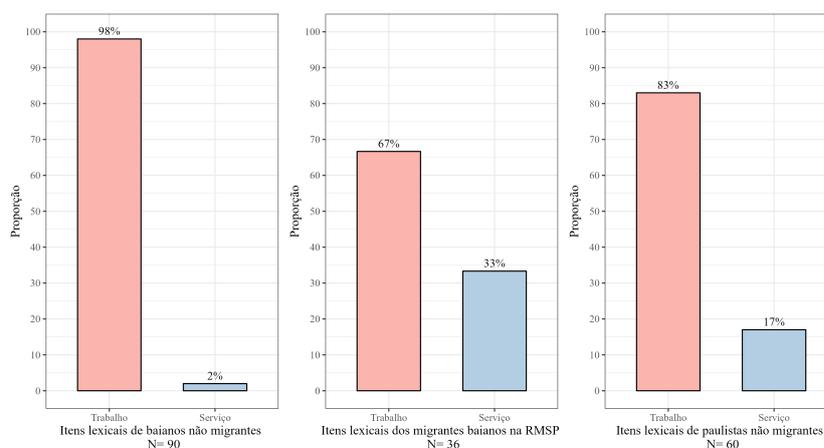
S1: *lá / lá no nordeste a turma chama “Quentinha” né? mas aqui nós chama marmitex*  
(MarioS)

D1: *marmitex né?*

No que tange à variável Trabalho/Serviço, em conformidade com os dados da Figura 13, na Bahia (à esq.), o emprego do item Trabalho é semicategórico (98%) em relação ao item Serviço (2%). Em São Paulo, de acordo com o gráfico da direita, as porcentagens dessas variantes são de 83% e 17%, respectivamente. Contrastando essas duas variedades do português, constata-se que a lexia Serviço, na variedade paulista, é mais produtiva em comparação com o dialeto baiano. Os dados dos migrantes baianos na RMSP (gráfico do meio), por seu turno, revelam não só que há uma adaptação no emprego dessa variável seguindo na direção do padrão da comunidade anfitriã quanto ao emprego de Serviço mas também que a taxa de 33% (12/36) de aplicação dele ultrapassa o percentual dos paulistas (17%).

<sup>21</sup> Devido à ausência de trabalhos que analisam esse caso de variação, esta descrição está sendo feita com base nas experiências deste pesquisador.

**Figura 13** – Comparação das proporções de uso dos itens lexicais Trabalho e Serviço de baianos não migrantes, de migrantes baianos na RMSP e de paulistas não migrantes

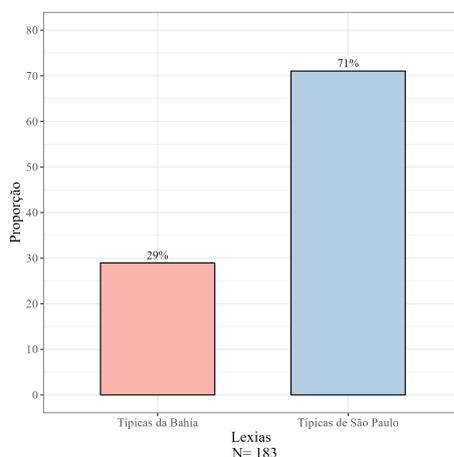


Fonte: Elaboração própria.

Cabe lembrar que a pergunta do QSL para capturar esse dado apresentou ruído (ver a Subseção 4.2) e, por isso, o número de observações foi menor (36) em relação aos outros itens lexicais. Alguns tópicos do roteiro utilizado nas entrevistas permitiram coletar dados intercambiáveis de Trabalho e Serviço na fala semiespontânea dos participantes. Fora os dados das respostas dialetológicas e de “Trabalho” (verbo), contabilizaram-se 23 observações dessas variantes, sendo 26% (6) do item Trabalho e 74% (17) de Serviço. Esses dados se coadunam com os resultados alcançados com o QSL sobre essa variável e permitem estimar que, no dia a dia, os migrantes baianos na RMSP tendem a usar mais o item Serviço em vez de Trabalho. De maneira oposta às outras variáveis, a variação diatópica de Trabalho/Serviço não foi mencionada em comentários metalinguísticos. Logo se presume que esse caso de variação está abaixo do nível da consciência do migrante.

Após analisar as adaptações de cada variável lexical dos migrantes baianos na RMSP, agruparam-se as lexias Aipim, Tangerina, Trabalho e Quentinha, recodificadas como Itens lexicais baianos, e Mandioca, Mexerica, Serviço, Marmita e Marmitex como Itens lexicais paulistas. Em seguida, testou-se a correlação dos Itens lexicais (baianos e paulistas) com os preditores sociais Sexo/Gênero, Idade, Nível de escolaridade, Idade de migração, Tempo de residência, Rede de contato com baianos e Grau de autoidentificação paulista declarada. Depois da recodificação, consoante o gráfico da Figura 14, verificou-se que, do total de 183 observações, 29% (53) são de Itens lexicais baianos e 71% (130) de Itens lexicais paulistas. Portanto, ao migrarem para São Paulo, os baianos tendem a empregar mais as lexias típicas do dialeto paulista em relação às lexias do dialeto baiano.

**Figura 14** – Proporção dos Itens lexicais baianos e paulistas utilizados pelos migrantes baianos residentes na RMSP



Fonte: Elaboração própria.

Os resultados dos testes de regressão logística estão apresentados nas Tabelas 3 a 9. Antes de ler e interpretar esses dados, porém, é conveniente entender como os valores estão distribuídos nessas tabelas. A título de exemplo, na primeira coluna da Tabela 3, há o preditor (em **negrito**) e seus níveis testados no modelo de regressão logística. Na segunda coluna, reporta-se o *input* dos dados, isto é, a frequência absoluta (Apl./N) de cada nível do preditor, seguido da sua frequência relativa (%) na terceira coluna. A partir da quarta coluna, consta o *output* do modelo estatístico. Nessa coluna está a estimativa (Est.) calculada em *logodds*, medida estatística que tem o 0 (zero) como referência neutra, os valores positivos indicando o favorecimento da variante de aplicação pelo nível do preditor independente e os valores negativos indicando o desfavorecimento em relação ao *intercept* (Oushiro, 2022). Na quinta coluna, está o erro padrão (EP), medida que indica a precisão das estimativas. Na sexta coluna, exibe-se o valor-z, obtido a partir da razão da estimativa pelo erro padrão. Na sétima coluna está o valor-p (*p*), que mostra se a diferença das estimativas é significativa (< 0,05) ou não (> 0,05) em relação ao *intercept*, cuja estimativa está localizada na linha superior da tabela. E a fórmula aplicada ao conjunto de dados está no rodapé dela. Algumas tabelas apresentam estruturas diferentes, como a Tabela 4 que não possui a coluna da frequência absoluta nem da frequência relativa, porque esse preditor foi analisado como uma variável contínua.

**Tabela 3** – Resultados do modelo de regressão logística de efeitos fixos quanto à aplicação dos Itens lexicais paulistas dos migrantes baianos na RMSP segundo o Sexo/Gênero

*Intercept = 0,810.*

Variável	Apl./N	%	Est.	EP	Valor-z	p
<b>Sexo/Gênero</b>						
mulheres ( <i>ref.</i> )	81/117	69%				
homens	49/66	74%	0,247	0,345	0,717	0,47

Modelo: glm (ITENS.LEXICAIS ~ SEXO, data = dados)

Fonte: Elaboração própria.

**Tabela 4** – Resultados do modelo de regressão logística de efeitos fixos para a realização dos Itens lexicais paulistas dos migrantes baianos na RMSP segundo a Idade

*Intercept= -0,150.*

Variável	Est.	EP	Valor-z	p
<b>Idade</b> ( <i>contínua</i> )	0,023	0,012	1,902	0,057

Modelo: glm (ITENS.LEXICAIS ~ IDADE, data = dados)

Fonte: Elaboração própria.

**Tabela 5** – Resultados do modelo de regressão logística de efeitos fixos para a realização dos Itens lexicais paulistas dos migrantes baianos na RMSP segundo a Rede de contato com baianos

*Intercept= 0,785.*

Variável	Est.	EP	Valor-z	p
<b>Rede de contato</b> ( <i>contínua</i> )	0,015	0,056	0,272	0,785

Modelo: glm (ITENS.LEXICAIS ~ REDE, data = dados)

Fonte: Elaboração própria.

**Tabela 6** – Resultados do modelo de regressão logística de efeitos fixos para a realização dos Itens lexicais paulistas dos migrantes baianos na RMSP segundo o Grau de autoidentificação paulista declarada

*Intercept= 0,551.*

Variável	Est.	EP	Valor-z	<i>p</i>
<b>Autoidentificação</b> ( <i>contínua</i> )	0,087	0,049	1,763	0,077

Modelo: glm (ITENS.LEXICAIS ~ IDENTIDADE, data = dados)

Fonte: Elaboração própria.

Os resultados das Tabelas 3, 4, 5 e 6 indicam que não há correlação dos Itens lexicais (baianos e paulistas) com os preditores Sexo/Gênero (0,247 *logodds*,  $p = 0,47$ ); Idade (0,023 *logodds*,  $p = 0,057$ ); Rede de contato com baianos (0,015 *logodds*,  $p = 0,785$ ) e Grau de autoidentificação paulista declarada (0,087 *logodds*,  $p = 0,077$ ). Por outro lado, as Tabelas 7, 8 e 9 mostram que houve correlação dos Itens lexicais com o Nível de escolaridade (-1,013 *logodds*,  $p = 0,009$ ); com a Idade de migração (-0,058 *logodds*,  $p = 0,01$ ) e com o Tempo de residência (0,038 *logodds*,  $p = 0,01$ ). Os coeficientes da regressão logística que mostram a relação do Nível de escolaridade com o uso dos Itens lexicais estão expostos na Tabela 7.

**Tabela 7** – Resultados do modelo de regressão logística de efeitos fixos quanto à aplicação dos Itens lexicais paulistas dos migrantes baianos na RMSP segundo o Nível de escolaridade

*Intercept = 1,363.*

Variável	Apl./N	%	Est.	EP	Valor-z	<i>p</i>
<b>Nível de escolaridade</b>						
Fundamental ( <i>ref.</i> )	86/108	80%				
Médio	44/75	59%	-1,013	0,334	-3,026	0,002 **

Modelo: glm (ITENS.LEXICAIS ~ NIVEL.ESCOLARIDADE, data = dados)

Fonte: Elaboração própria.

Com base nos coeficientes da Tabela 7, o valor negativo da estimativa do fator Ensino Médio (-1,013 *logodds*,  $p = 0,002$ ) indica que os migrantes desse nível de escolaridade tendem a não empregar os itens lexicais paulistas em relação aos de Ensino Fundamental. Com esse resultado, acata-se a hipótese de que os migrantes com menos escolaridade tendem a empregar mais os itens lexicais paulistas em relação aos migrantes baianos com mais escolaridade.

Os coeficientes da correlação com a Idade de migração estão na Tabela 8. Com base no valor negativo da estimativa (-0,058 *logodds*,  $p = 0,009$ ), observa-se que quem migrou

mais tarde tende a não utilizar os Itens lexicais paulistas em comparação com quem chegou mais cedo à comunidade anfitriã. Interpretando de forma contrária, os migrantes baianos que migraram mais cedo tendem a utilizar mais as lexias típicas de São Paulo em relação a quem migrou mais tarde, conforme a hipótese aventada.

**Tabela 8** – Resultados do modelo de regressão logística de efeitos fixos para a realização dos Itens lexicais paulistas dos migrantes baianos na RMSP segundo a Idade de migração

*Intercept= 2,100.*

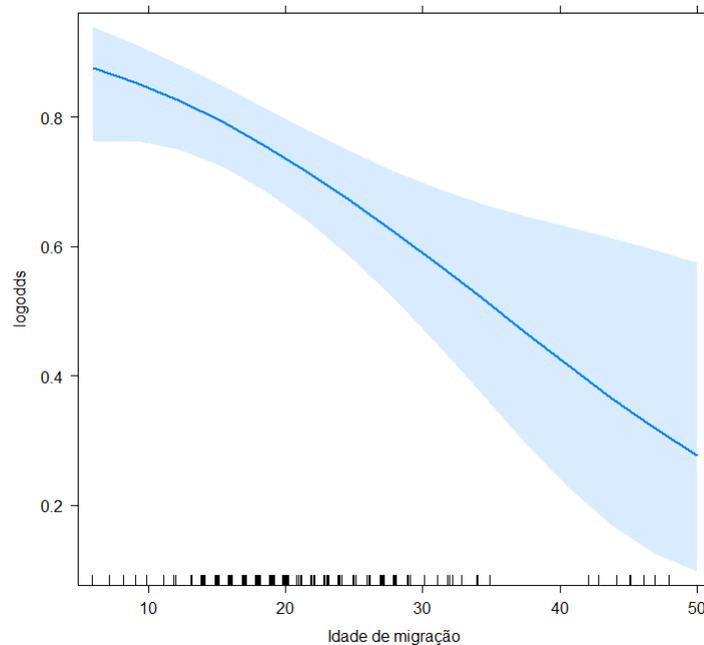
Variável	Est.	EP	Valor-z	p
<b>Idade de migração</b> ( <i>contínua</i> )	-0,058	0,022	-2,587	0,009 **

Modelo: glm (ITENS.LEXICAIS ~ IDADE.MIGRACAO, data = dados)

Fonte: Elaboração própria.

Os valores da regressão logística plotados na Figura 15 configuram uma linha descendente e atestam que os migrantes que chegaram mais cedo a RMSP tendem a empregar as lexias paulistas relativamente mais do que os que chegaram mais tarde.

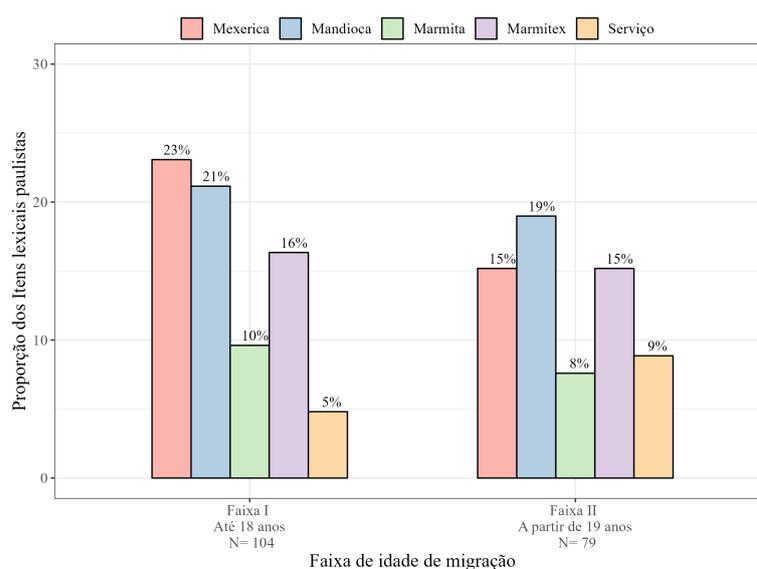
**Figura 15** – Correlação dos Itens lexicais paulistas e a Idade de migração dos baianos residentes na RMSP



Fonte: Elaboração própria.

Para uma descrição mais detalhada das adaptações lexicais dos migrantes baianos na RMSP, é cabível questionar qual foi a taxa de uso de cada Item lexical paulista computada no repertório linguístico dos baianos que migraram com menos idade e com mais idade. Para responder a esse questionamento, agruparam-se os participantes em duas faixas de idade de migração, uma com representantes dos que chegaram à RMSP com até 18 anos de idade (N= 29) e outra com os que chegaram a partir de 19 anos (N= 21). Na tentativa de equilibrar a quantidade de participantes nas duas faixas, usou-se como critério a maioridade, uma vez que, relativamente, até 18 anos, os jovens supostamente estão sob a tutela de um adulto, enquanto os que têm a partir de 19 anos provavelmente já não necessitariam dessa supervisão. A expectativa para essa análise foi a de que todos os itens lexicais apresentariam as proporções maiores na Faixa I em relação à Faixa II, contudo o resultado dessa comparação mostra uma relativa diferença no uso do item Serviço, conforme se observa na Figura 16.

**Figura 16** – Proporção dos Itens lexicais paulistas de acordo com a Faixa de idade de migração dos baianos residentes na RMSP



Fonte: Elaboração própria.

Com as informações da Figura 16, constata-se que o item Mexerica apresentou 23% (24 ocorrências) na Faixa I e 15% (12) na Faixa II; a lexia Mandioca teve 21% (22) na Faixa I e 19% (15) na Faixa II; a variante Marmita teve 10% (10) na Faixa I e 8% (seis) na Faixa II; o item Marmitex evidenciou 16% (17) na Faixa I e 15% (12) na Faixa II; por fim Serviço apresentou 5% (cinco) na Faixa I e 9% (sete) na Faixa II. Contrastando os valores de ambas as faixas de idade de migração, verifica-se que, à exceção do item Serviço, que apresentou menor proporção na Faixa I, os demais itens lexicais tendem a ser mais empregados na Faixa I em comparação com a Faixa II com diferenças bastante relativas.

A respeito da correlação com o Tempo de residência, os coeficientes da regressão logística (0,038 *logodds*,  $p = 0,001$ ) explicitados na Tabela 9 e o gráfico da Figura 17 sugerem acatar a hipótese de que quem está há mais tempo residindo na RMSP tende a utilizar os Itens lexicais paulistas em relação a quem está há menos tempo.

**Tabela 9** – Resultados do modelo de regressão logística de efeitos fixos da realização dos Itens lexicais paulistas segundo o Tempo de residência dos migrantes baianos na RMSP

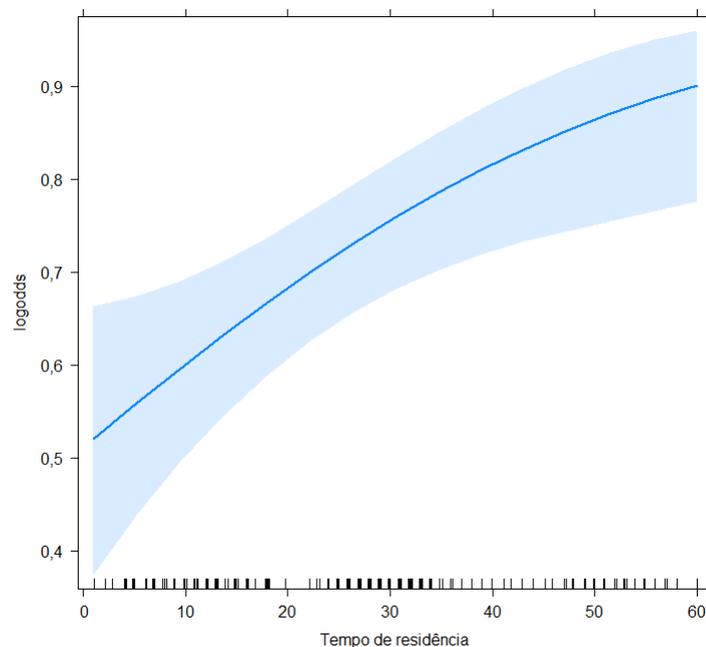
*Intercept= 0,032.*

Variável	Est.	EP	Valor-z	<i>p</i>
<b>Tempo de residência</b> ( <i>contínua</i> )	0,038	0,012	3,163	0,001 **

Modelo: glm (ITENS.LEXICAIS ~ TEMPO.RESIDENCIA, data = dados)

Fonte: Elaboração própria.

**Figura 17** – Gráfico da correlação dos Itens lexicais paulistas e o Tempo de residência dos migrantes baianos na RMSP



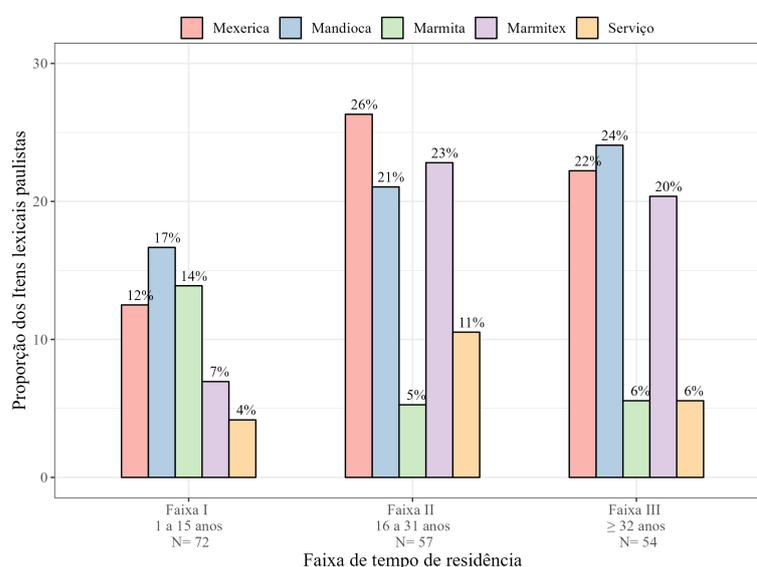
Fonte: Elaboração própria.

Na tentativa de encontrar algum padrão da adaptação de cada Item lexical paulista quanto ao tempo de estadia do migrante baiano na comunidade anfitriã, agruparam-se os participantes em Faixas de tempo de residência. Na Faixa I estão os participantes que residem na

RMSP no período entre 1 e 15 anos (N= 19), na Faixa II estão aqueles que residem entre 16 e 31 anos (N= 16) e na Faixa III aqueles que estão com o tempo de residência  $\geq 32$  anos (N= 15). O período estabelecido para cada faixa foi definido visando apenas a um equilíbrio da quantidade de participantes em cada uma delas.

A expectativa para essa análise foi a de que as proporções de cada item lexical se mostrariam ascendentes, isto é, que os valores de todos os Itens lexicais paulistas seriam menores na Faixa I em relação às Faixas II e III e que os valores da Faixa II seriam menores que os da Faixa III. Os resultados dessa análise, porém, indicam diferentes gradações para cada item lexical, conforme se observa nos dados da Figura 18.

**Figura 18** – Proporção dos Itens lexicais paulistas de acordo com a Faixa de tempo de residência dos baianos na RMSP



Fonte: Elaboração própria.

As proporções do item Mexerica demonstram um aumento no uso dessa variante da Faixa I (12%, N= 9) para a Faixa II (26%, N= 15) e um declínio desta para a Faixa III (22%, N= 12). Os dados do item Mandioca mostraram um aumento na realização dessa variante da Faixa I (17%, N= 12) para a Faixa II (21%, N= 12) e desta para a Faixa III (24%, N= 13). Os valores do item Marmita explicitaram um declínio no emprego dessa variante da Faixa I (14%, N= 10) para a Faixa II 5% (três) e um leve aumento desta para a Faixa III (6%, N= 3). Os resultados do item Marmitex mostraram um aumento no uso dessa variante da Faixa I (7%, N= 5) para a Faixa II (23%, N= 13) e uma leve queda dessa para a Faixa III (20%, N= 11). As taxas do item Serviço evidenciaram um aumento dessa variante da Faixa I (4%, N= 3) para a Faixa II (11%, N= 12) e uma redução desta para a Faixa III (6%, N= 3).

### 4.3.2 Discussão dos achados

O estudo das adaptações lexicais dos migrantes baianos residentes na RMSP amplia as discussões sobre os efeitos do contato entre dialetos mutuamente inteligíveis, uma vez que, com os resultados, depreenderam-se duas principais informações. A primeira é a de que a saliência e a não saliência social das variantes implica diferentes taxas de adaptação dos Itens lexicais paulistas. A segunda informação é a de que cada variável sociolinguística lexical perpassa diferentes processos de adaptação durante o contato dialetal.

A respeito da saliência das variantes, vê-se que as proporções mais elevadas de adaptação em favor do padrão paulista foram de Mandioca, com 76%, de Mexerica, com 72%, e de Marmitex, com 60,4%, e também foram as que se destacaram por apresentar comentários metalinguísticos acerca de diferenças diatópicas. O item Serviço, por outro lado, teve 33% de aplicação e nenhum comentário desse tipo. Para Chambers (1992) e Trudgill (1986), as alterações linguísticas resultantes do contato entre dialetos mutuamente inteligíveis acontecem inicialmente nos itens lexicais porque eles são mais salientes. Contudo, se se considerar que os baianos, ao migrarem para São Paulo, também aumentaram a frequência de uso do item Serviço, que parece não estar acima do nível da consciência, verifica-se que os itens lexicais menos salientes socialmente também são impactados com o contato dialetal, mas, ao que tudo indica, de maneira menos frequente em relação aos itens mais salientes.

Constatou-se que há distintos processos de adaptação lexical quando a comunidade anfitriã e a comunidade de origem compartilham das mesmas variantes de uma variável e quando não compartilham delas. O item Mexerica está presente no repertório linguístico dos baianos não migrantes (3%) (ver a Tabela 2). Contudo, quando migram para São Paulo, eles tendem a *potencializar* esse uso (72%), acompanhando as taxas encontradas na fala dos paulistas não migrantes (92%). O item Mandioca teve 11% de aplicação nos dados dos participantes residentes na Bahia<sup>22</sup>. Nos dados dos migrantes baianos em São Paulo, há 76% do uso dessa variante, seguindo o padrão dos paulistas não migrantes (100%).

Essa adaptação, contudo, não é um caso de potencialização como aconteceu com o item Mexerica. Como se pôde ver na Subseção 4.1, a lexia Mandioca consta no dialeto baiano, mas com o significado mais restrito em relação aos sentidos atribuídos a esse mesmo significante em São Paulo. Na Bahia, utiliza-se o termo Mandioca para designar a raiz não comestível, enquanto em São Paulo essa lexia serve tanto para nomear a raiz comestível quanto a raiz tóxica. O que se verifica com os resultados dos migrantes baianos em relação ao uso da variável Aipim/Mandioca, portanto, é que eles ampliam os sentidos de uma lexia já existente em seu repertório linguístico. Essa mesma *ampliação semântica* ocorreu com o item Marmita,

<sup>22</sup> Novamente, esse total advém dos respondentes que nasceram em São Paulo e migraram para a Bahia e de baianos que pertencem à rede desses migrantes.

tendo em vista que, na Bahia, essa lexia é usada para nomear o recipiente com compartimentos utilizado para transportar alimentos ou então a porção de comida levada de casa para outros lugares. Esse processo, porém, não se aplica ao item Marmitex. Presumivelmente, esse item entra no repertório linguístico do migrante pelo processo de substituição (*replacement*), uma vez que, segundo Chambers (1992), a *substituição* acontece quando duas palavras distintas são usadas para se referir a uma mesma coisa nas duas variedades. Entretanto, para chegar a essa conclusão, seria necessária a ampliação de dados de baianos e paulistas não migrantes.

Em relação ao encaixamento social das adaptações lexicais, os modelos de regressão logística mostraram que os migrantes baianos que estudaram até o Ensino Fundamental, que migraram com menos idade e que estão há mais tempo em São Paulo tendem a utilizar mais os Itens lexicais paulistas. É possível que esse resultado tenha alguma relação com as mudanças na educação brasileira quanto ao ensino de Língua Portuguesa nestes últimos 25 anos. Com a publicação dos Parâmetros Curriculares Nacionais – PCNs em 1997 (Brasil, 1997b), o ensino de língua materna focalizou, além de outros tópicos, o respeito à diversidade linguística e o combate ao preconceito linguístico. Dessa maneira, provavelmente o trabalho em sala de aula sobre as diferenças diatópicas pode ter feito com que os participantes que estudaram até o Ensino Médio, que migraram com mais idade e que estão há menos tempo em São Paulo não tendessem a usar os itens lexicais paulistas.

Na adaptação de cada item lexical segundo a Faixa de idade de migração, conforme se viu na Figura 16, à exceção do item Serviço, as demais lexias apresentaram taxas de uso maiores do que os que migraram até 18 anos e os que migraram a partir de 19 anos de idade. A diferença no emprego do item Serviço entre essas faixas de idade de migração talvez tenha a ver com o fato de os adultos já estarem inseridos no mercado de trabalho. Ou seja, presumivelmente é um item muito frequente em suas interlocuções.

Na disposição de cada item lexical de acordo com a Faixa de tempo de residência, notou-se, com base na Figura 18, que as taxas de uso de Mexerica, Mandioca, Marmitex e Serviço tiveram aumento da Faixa I para a II. Em compensação, a taxa de Marmita reduziu da Faixa I para a II. Da Faixa II para a III, à exceção do item Mandioca e do item Marmita, que apresentaram um aumento relativo, os demais itens tiveram redução em suas taxas.

A coexistência dos itens Marmita e Marmitex chama bastante atenção, porque na Faixa I há 14% de Marmita e 4% de Marmitex, já na Faixa II, essas variantes apresentaram 5% e 23%, respectivamente, e, na Faixa III, Marmita teve um leve aumento (6%) e Marmitex atingiu 20%. A partir desses dados, pode-se inferir que, no primeiro momento do contato, a lexia Quentinha deixa de ser usada e Marmita assume o seu lugar, quiçá como um item neutro. Mais tarde, Marmitex tende a substituir o item Marmita.

Talvez a obliteração do item Quentinha tenha a ver com a inteligibilidade da lexia. Assim que este pesquisador se instalou em Campinas (SP) para cursar o Doutorado, foi comprar

o almoço num mercado e, ao pedir uma “Quentinha”, a atendente não sabia o que era. Depois que ele apontou para o que queria, ela disse que era uma “Marmita”. Com base nesse fato, possivelmente a ampliação semântica de Marmita aconteça na primeira fase do contato para garantir a comunicação com a população anfitriã, uma vez que a lexia Quentinha parece ser pouco ou nada conhecida em São Paulo. É oportuno ressaltar que os três únicos dados de Quentinha foram empregados por LeticiaC, que está em São Paulo (SP) há dois anos; por NelmaS que mora nessa cidade há apenas um ano; e por SamiraS que migrou há nove anos. Elas pertencem à mesma rede social e foram entrevistadas juntas. Considerando a ordem das respostas dadas durante a entrevista (LeticiaC – NelmaS – SamiraS), pode-se conjecturar que a resposta das duas últimas tenha sido um paralelismo, isto é, elas podem ter replicado a resposta da primeira. Isso sendo verdade e considerando que LeticiaC reside em São Paulo (SP) há apenas dois anos, sustenta-se a hipótese de que o item Quentinha é rapidamente obliterado na primeira fase do contato.

Os valores dos itens Mexerica, Marmitex e Serviço na Faixa III chamam atenção, porque a expectativa era a de que, assim como no item Mandioca, eles aumentassem em relação à faixa anterior. Esses dados se assemelham aos resultados vistos por Souza (2019) na análise do uso do imperativo com morfologia de indicativo de migrantes baianos em São Paulo. Na correlação da variável resposta com o preditor Tempo de residência em São Paulo, Souza (2019) verificou que os migrantes das Faixas II (8-15 anos) e III (+16 anos) tendem a não empregar a variante típica da comunidade anfitriã em comparação com a Faixa I (0-7 anos). Fundamentando-se nesses dados, é plausível inferir que certas variantes típicas da comunidade anfitriã deixam de ser utilizadas pelos migrantes com o passar do tempo de sua estadia. Se se considerar que, na amostra examinada neste estudo, os migrantes que estão há mais de 32 anos residindo na RMSP são os que têm idade mais avançada e aposentados, não estaria errado assumir que o mercado de trabalho tem forte influência na adaptação linguística dos migrantes na comunidade anfitriã.

#### 4.4 Síntese

Os resultados das análises das adaptações lexicais dos migrantes baianos residentes na RMSP quanto ao uso das variáveis sociolinguísticas Tangerina/Mexerica, Aipim/Mandioca, Quentinha/Marmita e Trabalho/Serviço revelam que (i) as lexias salientes do ponto de vista diatópico, isto é, aquelas das quais os migrantes têm consciência de que pertencem ao léxico paulista, tendem a explicitar as taxas mais elevadas de uso em comparação às lexias que, aparentemente, não são salientes; (ii) há distintos processos de adaptação dialetal no léxico do migrante, a saber, potencialização, ampliação semântica e substituição; (iii) segundo os resultados das análises de regressão logística, os migrantes baianos que estudaram até o Ensino Fundamental, que migraram com menos idade e que estão há mais tempo residindo na RMSP tendem a utilizar mais

os Itens lexicais paulistas; (iv) cada Item lexical paulista, ao que tudo indica, é implementado de modo diferente em relação à Idade de migração e ao Tempo de residência. (v) a adaptação lexical parece ter uma relação estreita com certos contextos sociais nos quais os migrantes estão inseridos, como no mercado de trabalho, por exemplo.

## 5

# Adaptação em nível fonético-fonológico: o (-r) em coda silábica

O (-r) em posição de coda silábica, como em “porta” e “deixar”, é uma variável proeminente, cujas variantes diferenciam dialetos do português brasileiro (Callou; Moraes; Leite, 2002 [1996]; Oliveira, 1999) e, por serem salientes, permitem identificar a origem de seus falantes (Oushiro, 2015). Sendo São Paulo (SP) um nicho de migrantes, essa cidade seria um lugar onde provavelmente essas variantes estariam em contato (Oushiro, 2015). Considerando que a migração de baianos para essa capital foi e ainda é bastante expressiva em relação a migrantes de outras proveniências (Fontes, 2008; IBGE, 2010; IPEA, 2011), questiona-se se eles, ao manterem contato com os paulistas, utilizariam o Tepe e o Retroflexo, róticos típicos da comunidade anfitriã (Oushiro, 2015), no lugar das Fricativas velar e glotal, típicas de sua região de origem (Callou; Moraes; Leite, 2002 [1996]; Oliveira, 1999; Guedes, 2014), nos contextos de coda silábica. Se a resposta for sim, é possível ainda indagar quais fatores sociais e linguísticos estão correlacionados com esse uso.

Nesta Seção, apresentam-se os resultados das análises multivariadas de regressão logística de efeitos mistos, com a inclusão do Participante e do Item lexical como variáveis aleatórias, nas quais se testou a correlação da variável (-r) em coda como Retroflexo/Tepe, em relação à forma Velar/Glotal, com os preditores sociais já apresentados na Seção 3; com o preditor contextual Estilo (conversação e leitura); e com os preditores linguísticos Contexto fônico precedente (anterior, central e posterior), Contexto fônico seguinte (coronal, dorsal, labial e pausa), Tonicidade da sílaba (tônica e átona), Posição da sílaba (medial e final) e Classe morfológica ([+gramatical] e [-gramatical]).

Esta Seção está assim dividida: na Subseção 5.1, contrastam-se os padrões de uso do (-r) em coda encontrados na Bahia a partir dos estudos de Callou, Moraes e Leite (2002 [1996]), Oliveira (1999) e Guedes (2014) com os de São Paulo registrados nas pesquisas de Callou, Moraes e Leite (2002 [1996]), Leite (2010), Oushiro e Mendes (2014), Oushiro (2015)

e Carreão (2019); na Subseção 5.2, descreve-se a maneira como foram manipulados os dados; na Subseção 5.3, leem-se e interpretam-se os coeficientes das análises de regressão logística; e, na Subseção 5.4, resumem-se os achados deste estudo.

### 5.1 A variação do (-r) em posição de coda silábica em variedades do português brasileiro: Bahia e São Paulo em evidência

O (-r) em posição de coda silábica é uma variável sociolinguística cujas variantes distinguem dialetos e, conseqüentemente, podem revelar o lugar de origem de seus falantes. Com o objetivo de delimitar áreas dialetais do português do Brasil a partir da realização dessa variável sociolinguística, Callou, Moraes e Leite (2002 [1996], p. 464) analisaram um *corpus* formado por 30 inquéritos do tipo DID (diálogo entre informante e documentador) e D2 (diálogo entre dois informantes) de amostras estratificadas em Faixa etária (I, 25-35 anos; II, 36-55 anos; e III,  $\geq$  56 anos) e Gênero (homem; mulher) de cinco capitais brasileiras que constituem o Projeto Norma Urbana Culta (NURC) (Castilho, 1990), a saber: Recife (PE), Salvador (BA), Rio de Janeiro (RJ), São Paulo (SP) e Porto Alegre (RS). Os resultados gerais alcançados por Callou, Moraes e Leite (2002 [1996]) indicam que, do total de 4.334 ocorrências de /r/ em coda, a variante predominante nessa amostra é a Vibrante simples (32%), seguida do Apagamento (26%), da Fricativa velar (21%) e da Aspirada (18%). A Vibrante uvular, a Vibrante múltipla e a Aproximante retroflexa apresentaram-se marginais, com 1% cada. Quanto às delimitações dialetais, Callou, Moraes e Leite (2002 [1996], p. 468) verificaram que: (i) o Apagamento, especialmente quando o /r/ está na coda da sílaba final da palavra, tende a ocorrer mais em Salvador (BA) (62%) em comparação com Recife (PE) (50%), São Paulo (SP) (49%), Rio de Janeiro (RJ) (47%) e Porto Alegre (RS) (37%); (ii) que a Vibrante simples tende a ser mais usada em São Paulo (SP) e Porto Alegre (RS); (iii) que as Fricativas velar e aspirada tendem a ser mais empregadas em Salvador (BA), Recife (PE) e Rio de Janeiro (RJ); e (iv) que o uso da Vibrante múltipla e da Aproximante retroflexa se restringem a São Paulo (SP) e a Porto Alegre (RS). Esses autores testaram, utilizando o programa estatístico Varbrul, a correlação do emprego do /r/ na posição medial e final separadamente com as variáveis linguísticas Tonicidade da sílaba, Tonicidade do vocábulo na cadeia fônica, Dimensão do vocábulo, Vogal antecedente, Segmento seguinte (ponto + modo de articulação do contexto fônico seguinte), Classe gramatical da palavra e com as variáveis sociais Faixa etária e Gênero. Como valor de aplicação, eles consideraram a variante com maior proporção em cada posição (medial e final) nas ocorrências de cada cidade. Desse modo, focalizando apenas os dados de São Paulo (SP) e de Salvador (BA), consideraram-se nas análises dos dados da capital paulista, por exemplo, a realização da Vibrante simples em posição medial e do Apagamento em posição final; enquanto nos dados da capital baiana, analisaram-se a Fricativa velar na posição interna e o Apagamento na posição final. Na análise dos dados de Salvador (BA), viu-se que os homens pertencentes às Faixas etárias III e II tendem a utilizar mais a Fri-

cativa velar no contexto medial quando o /r/ é antecedido por vogal posterior (P.R. .60) e vogal central (P.R. .53) e quando é seguido de oclusiva, fricativa, nasal, velar, alveolar e labial. No que concerne aos dados de Apagamento na posição final, observou-se que, segundo os autores, ocorrem mais entre os jovens e em verbos (P.R. .72).

Conforme os resultados de Callou, Moraes e Leite (2002 [1996]), Salvador (BA) apresenta o maior percentual de Apagamento do /r/ (62%) da amostra analisada. Contudo em Oliveira (1999), verifica-se que essa supressão, na fala culta da capital baiana, tem distintas proporções em diferentes contextos de Estilo. Especificamente sobre o Apagamento do rótico em relação ao emprego do /r/ implosivo na fala culta de Salvador (BA), Oliveira (1999) examinou uma amostra de fala do Projeto NURC formada por 32 informantes, sendo 16 inquiridos do tipo EF (elocução formal) e 16 do tipo DID. A partir de análises multivariadas realizadas no Varbrul, Oliveira (1999) testou a correlação do uso de (-r) em coda silábica de cada contexto de Estilo (EF e DID) separadamente com as variáveis sociais Sexo (homem; mulher) e Faixa etária (Faixa I, 25-34 anos; II, 35-44 anos; III, 45-54 anos; IV,  $\geq$  55 anos) e com as variáveis linguísticas Posição da sílaba, Vogal precedente, Contexto subsequente, Tonicidade da sílaba, Classe de palavra e Carga semântica do segmento. Os resultados globais de Oliveira (1999, p. 42) mostram que, do total de 7.607 ocorrências de /r/ em posição de coda nos dados de EFs, 32% foram de Apagamento e 68% da forma Implosiva. Por outro lado, de 5.204 observações dessa variável nos dados de DIDs, 46% foram de Supressão e 54% da Aspirada. Sendo assim, observa-se que, na variedade culta de Salvador (BA), os percentuais de realização do /r/ Implosivo são maiores que os de Apagamento; e que os participantes tendem a suprimir mais a pronúncia do /r/ em coda silábica em situações informais, como os DIDs, em comparação com contextos mais formais, como as EFs.

Os resultados das análises multivariadas dos EFs evidenciam que o Apagamento do /r/ tende a ocorrer em numeral (P.R. .98), em conjunção (P.R. .84), em verbo no infinitivo (P.R. .58) e em preposição (P.R. .54), em posição final (P.R. .94), em sílaba pretônica (P.R. .62), quando o /r/ é precedido pela vogal [u] (P.R. .86) ou pela vogal [ɛ] (P.R. .55) e seguido de consoante alveolar (P.R. .61) ou de consoante posterior (palatalizadas, palatais e velares). Os resultados das análises multivariadas realizadas com os dados de DIDs assemelham-se aos valores vistos nas EFs. A supressão do /r/ tende a acontecer em conjunção (P.R. .91), em preposição (P.R. .54) e em verbo (P.R. .66), na coda da sílaba final do vocábulo (P.R. .93), em sílaba pretônica (P.R. .64), quando o /r/ é precedido pelas vogais [u] (P.R. .69), [ɔ] (P.R. .63) e [ɛ] (P.R. .56) e seguido de consoante alveolar (P.R. .63). Sobre as variáveis sociais, com os dados de EFs e DIDs amalgamados, os resultados apontam para a mudança em progresso em favor da supressão do /r/ em coda, uma vez que esse fenômeno tende a acontecer mais entre os participantes da faixa I (P.R. .59) em comparação com a faixa II (P.R. .52), com a faixa III (P.R. .47) e com a faixa IV (P.R. .34). Outro dado relevante dessa análise é o de que, no cruzamento

das variáveis faixa etária e sexo, os resultados salientam que os homens da faixa I encabeçam essa mudança (P.R. .66).

No interior da Bahia, o processo de apagamento do rótico parece estar mais acelerado em relação à capital baiana. Levando em conta uma amostra de fala de 12 participantes do projeto A Língua Portuguesa do Semiárido Baiano (Almeida; Carneiro, 2008), estratificada em Sexo/Gênero (masculino; feminino), Faixa etária (I, 25-35 anos; II, 45-55 anos; III,  $\geq$  65 anos) e Nível de escolaridade (Fundamental I; Nível universitário), Guedes (2014) observou o apagamento desse segmento fônico em relação às Fricativas velar e glotal, focalizando cada uma das posições da coda silábica no vocábulo, medial e final. Para tanto, ela testou, a partir do Goldvarb X e em cada um desses contextos, a correlação do (-r) com as variáveis sociais Sexo/Gênero, Faixa etária e Nível de escolaridade e com as variáveis linguísticas Vogal precedente, Zona de articulação da consoante seguinte, Modo de articulação da consoante seguinte, Classe morfológica, Sonoridade e Tonicidade da sílaba. O resultado geral dessa análise revelou que o Apagamento é semicategórico quando o /r/ está na posição final do vocábulo, perfazendo o total de 98% do universo de 2.960 observações dessa variável. Na posição medial, nota-se que, do total de 2.130 ocorrências de /r/, o emprego do Zero fonético foi de 20% em comparação com as Fricativas velar/glotal, com 80%. Os valores da análise multivariada evidenciam que o Apagamento do /r/ em coda tende a ocorrer em não verbos (P.R. .52), quando esse fonema é precedido pela vogal [i] (P.R. .67) e seguido de consoante surda (P.R. .56), produzida na zona palatal (P.R. .91), alveolar (P.R. .87) e labiodental (P.R. .77) e articulada no modo fricativo (P.R. .90). Na análise feita com os dados de /r/ na posição final, verificou-se que o Apagamento tende a acontecer em não verbos (P.R. .64) e quando o rótico é precedido por vogal [a] (P.R. .72) e [o] (P.R. .59).

No Estado de São Paulo, por outro lado, o padrão de uso do (-r) em coda tende a ser diferente do encontrado na Bahia. Conforme já mencionado, segundo os dados de Callou, Moraes e Leite (2002 [1996]), as variantes predominantes na cidade de São Paulo (SP) são a Vibrante simples em posição medial (87%) e o Zero fonético em posição final (49%). As variantes Retroflexa e Aspirada, nos resultados dessa pesquisa, apresentaram-se marginais. Acerca do uso da Vibrante simples no contexto interno, verificou-se que essa variante tende a ser mais realizada por homens (P.R. .64), em vocábulos polissílabos (P.R. .67) e trissílabos (P.R. .62) e quando o /r/ é seguido de Fricativa alveolar (P.R. .74). Os valores da Faixa etária, segundo os autores, indicam um caso de variação estável. Em relação ao Apagamento na posição final, os resultados apresentam que esse fenômeno tende a se efetuar na fala das mulheres (P.R. .69), em verbos (P.R. .64), em contextos em que o /r/ é precedido por vogal central (P.R. .79) e por anterior (P.R. .78) e seguido de consoantes (P.R. .56).

Oushiro e Mendes (2014) investigaram a variação do Apagamento vs. Realização do (-r) em posição de coda silábica na fala de paulistanos, cuja análise levou em consideração

seis subconjuntos de dados (verbos com morfema de infinitivo; conjunção/preposição; verbos finitos/raiz; advérbios; adjetivos; substantivos) em diferentes contextos de Estilo (conversaço; depoimento; leitura de jornal; leitura de palavras). Para chegar aos resultados, esses autores examinaram uma amostra de fala de 118 paulistanos, estratificada em Sexo/Gênero (masculino; feminino), Faixa etária (20-34 anos; 35-59 anos;  $\geq$  60 anos), Nível de escolaridade (médio; universitário) e Região de residência (bairros mais centrais; mais periféricos). O resultado geral desse estudo evidencia que, do total de 68.764 dados de (-r) em coda, a variante predominante é o Apagamento (56,3%), depois o Tepe/Vibrante (29,2%), a Aproximante retroflexa/alveolar (14,1%) e a Fricativa velar/glotal (0,4%).

Além dessas informações, Oushiro e Mendes (2014) notaram a existência de uma gradação das taxas de Apagamento do /r/ em relação aos contextos de Estilo, de maneira que quanto menor foi o monitoramento da fala, maiores foram as taxas de Apagamento, sendo 97% em conversaço, 33% em leitura de depoimento, 10% em leitura de notícia de jornal, 5% em lista de palavras. Especificamente nos dados da conversaço, Oushiro e Mendes (2014, p. 255) descobriram uma gradação também quanto ao estatuto morfológico do vocábulo. De acordo com as taxas encontradas por esses autores, o Apagamento nesse contexto de Estilo tende a se efetivar em verbos com morfema de infinitivo (97%), depois em conjunções/preposições (52%), seguido de verbos finitos ou da presença do /r/ na raiz do verbo (33%), em advérbios (12%), em adjetivos (10%) e, por fim, em substantivos (3%).

Oushiro (2015), fazendo uso desse mesmo *corpus*, observou, entre outros traços linguísticos que identificam a fala paulistana, a realização variável do (-r) em coda silábica como Tepe ou Retroflexo. Embora o mapeamento das formas de realização de /r/ em São Paulo (SP) corrobore o levantamento feito por Callou, Moraes e Leite (2002 [1996]). Oushiro (2015) destaca o aumento da Aproximante retroflexa na capital paulista, saindo de 2% (Callou; Moraes; Leite, 2002 [1996]) para 14,1%. Utilizando o programa computacional Rbrul, Oushiro (2015) testou a correlação do uso do Tepe e do Retroflexo em coda silábica com as variáveis linguísticas Contexto fônico precedente, Contexto fônico seguinte, Tonicidade da sílaba, Posição da sílaba e Classe morfológica; com a variável contextual Estilo; e com as variáveis sociais Sexo/Gênero, Faixa etária/idade, Nível de escolaridade, Região de residência, Classe social/Índice socioeconômico, Origem dos pais e Mobilidade. Os resultados das análises multivariadas, nas quais a autora incluiu o Participante e o Item lexical como variáveis aleatórias, indicaram que a Aproximante retroflexa, em relação ao Tepe, tende a ser mais utilizada por homens (P.R. .59), paulistanos filhos de migrantes nortistas, nordestinos (P.R. .82) e filhos do interior de São Paulo/Minas Gerais/Paraná (P.R. .60), por residentes em região mais periférica (P.R. .67), que estudaram até o Ensino Médio (P.R. .61), pertencentes às classes C1/C2/D (P.R. .68), que têm baixa mobilidade (P.R. .71) e em contexto de conversaço/leitura de jornal/depoimento (P.R. .58). No que concerne ao encaixamento linguístico, os resultados de Oushiro (2015) apontam

que o Retroflexo tende a ser mais realizado em verbo (raiz)/advérbio (P.R. .66), em sílaba ocupante da posição final da palavra (P.R. .60), em sílaba tônica (P.R. .55), quando o /r/ é precedido por vogal com o traço [-alto] (P.R. .63) e seguido de consoante [+coronal] (P.R. .55).

Diferentemente da capital, observa-se que, em cidades do interior paulista, a variante do (-r) predominante é a Aproximante retroflexa, o /r/ caipira, como comumente é caracterizada (Amaral, 1920). Leite (2010) pesquisou o uso dessa variável sociolinguística em Campinas (SP), município a 100 km de São Paulo (SP). A partir de uma amostra de fala de 12 campineiros, coletada numa sala com tratamento acústico<sup>23</sup>, sob a forma de entrevistas semidiretivas, de leitura de palavras e sentenças, Leite (2010) investigou a variação e as atitudes vinculadas à forma de realização do (-r) em posição de coda (medial e final). Os valores dos formantes, capturados com o auxílio do Praat, foram correlacionados estatisticamente, no programa Statistica, com as variáveis sociais Sexo (masculino; feminino), Faixa etária (I, 20-30 anos; II, 37-47 anos; III,  $\geq 54$  anos) e Nível de escolaridade (médio; universitário). Destacando apenas os resultados totais da análise de produção, evidencia-se a realização de três variantes no falar campineiro: o Retroflexo (90,6%), a Vogal rotacizada (9,0%) e o /r/ Vocalizado (0,4%). Vale dizer que, nessa amostra, não houve dados de Apagamento do rótico nem diferença estatisticamente significativa entre os níveis das variáveis sociais analisadas.

Em Louveira (SP), cidade a 28 km de Campinas (SP) e a 72 km de São Paulo (SP), tem-se notado, porém, o Tepe timidamente sendo utilizado em alternância com o Retroflexo (Carreão, 2019). No estudo de Carreão (2019), em que se observou a realização das oclusivas dentais [t, d] diante de [i] e o (-r) em posição de coda silábica, a hipótese de que variantes de grandes centros urbanos tendem a se espalhar para cidades menores foi acatada. Apoiando-se numa amostra de fala de 25 participantes, estratificada em Sexo/Gênero (masculino; feminino), Região de residência (bairro; centro) e Faixa etária (I, 18-29 anos; II, 30-45 anos; III, 46-60 anos), Carreão (2019) descreve, especialmente, o padrão dessas variáveis sociolinguísticas em relação à faixa etária e a sua relação com o mercado de trabalho. Particularmente quanto aos dados de (-r) em coda, os resultados para a Faixa etária apresentam diferença estatisticamente significativa ( $\chi^2 = 91,172(2)$ ,  $p = 0,001$ ) e indicam que os louveirenses com idade entre 30 e 45 anos tendem a empregar mais o Tepe (21%) em comparação com a faixa de idade entre 46 e 60 anos (11%) e com a faixa de idade entre 18 e 29 anos (8%). De acordo com Carreão (2019), o uso significativo do Tepe pela faixa etária II tem a ver com o crescimento econômico de Louveira (SP), que tem possibilitado aos louveirenses o desenvolvimento de atividades profissionais em empresas e não mais de atividades agrícolas.

Contrastando os dados linguísticos das variedades baianas com os das variedades paulistas, vê-se que elas têm diferentes padrões sociolinguísticos quanto à realização do (-r) em

<sup>23</sup> Especificamente no Estúdio de Gravação do Instituto de Estudos da Linguagem – IEL/UNICAMP.

coda silábica. Acerca das variedades baianas, com base nos dados de Callou, Moraes e Leite (2002 [1996]), Oliveira (1999) e Guedes (2014), depreendem-se duas principais informações sobre os padrões das variedades baianas. A primeira é a de que as variantes de (-r) em coda predominantes na Bahia (capital e interior) são o Zero fonético e as Fricativas velar e glotal. A segunda é a de que os resultados de Oliveira (1999) se assemelham ao padrão visto por Callou, Moraes e Leite (2002 [1996]) e se distanciam dos valores encontrados por Guedes (2014) quanto ao Apagamento do rótico nesse estado. Neste trabalho, ressalta-se que a queda do /r/ em coda, sobretudo em posição final, está num estágio mais avançado em comparação com os dados de Salvador (BA). Essas distinções talvez tenham a ver com as sincronias em que se levantaram os dados. As amostras de Callou, Moraes e Leite (2002 [1996]) e de Oliveira (1999) foram coletadas na década de 1970 e, portanto, refletem o comportamento linguístico dos falantes dessa sincronia, enquanto a amostra de Guedes (2014) foi recolhida na primeira década de 2000, quarenta anos depois das recolhas do Projeto NURC. Quanto às variedades paulistas, conforme os resultados de Callou, Moraes e Leite (2002 [1996]), Leite (2010), Oushiro e Mendes (2014), Oushiro (2015) e Carreão (2019), verifica-se que, embora se perceba o avanço da realização do Retroflexo na periferia de São Paulo (SP), o Tepe ainda é a variante prototípica dessa cidade, com destaque para a sua região central. A respeito de cidades do interior paulista, reconhece-se que o Retroflexo é proeminente, apesar de já se observarem taxas significativas de Tepe.

Dos resultados apresentados, merecem destaque os de Oushiro (2015), especificamente quanto à relação do uso do Retroflexo com o preditor social Origem dos pais. A correlação do (-r) em coda com esse preditor social demonstra a relevância de estudar os migrantes em situação de contato dialetal. A autora observa que os paulistanos filhos de migrantes nortistas, nordestinos e do interior do estado têm encabeçado o avanço do uso do Retroflexo na capital paulista. Assim, reconhecendo a relevância de estudar a produção dessa variável sociolinguística por essa população da qual os baianos fazem parte, vê-se que o confronto dos padrões de realização do (-r) em posição de coda entre as variedades baiana e paulista se faz indispensável para a tomada de decisões metodológicas, cujos critérios estão descritos na Subseção 5.2.

## 5.2 Manipulação dos dados

A partir dos critérios utilizados por Oliveira (1999) e Oushiro (2015), codificaram-se as ocorrências do (-r) em posição de coda silábica medial e final da fala dos migrantes baianos na RMSP. Dessa maneira, excluíram-se do cômputo as palavras estrangeiras, como *Big Brother*, por exemplo, por não poder distinguir se a pronúncia do /r/ Retroflexo seria influência do inglês ou do falar paulista (Oushiro, 2015); as palavras com o /r/ final seguido de vogal, por ele ser considerado pré-vocálico (Oliveira, 1999); os dados metalinguísticos; os duvidosos; os dados de permuta e os de silabação. Codificaram-se então 12.652 realizações do /r/ na posição de coda silábica em Apagamento, Fricativa velar/glotal, Retroflexo e Tepe, cujos valores absolutos (N)

e relativos (%) podem ser observados na Tabela 10.

**Tabela 10** – Distribuição das ocorrências do (-r) em coda silábica medial e final da fala dos migrantes baianos residentes na RMSP

<b>Codificação</b>	<b>N</b>	<b>%</b>
Apagamento	8.009	63,3
Fricativa velar/glotal	3.277	25,9
Retroflexo	850	6,7
Tepe	516	4,1
<b>Total</b>	12.652	100

Fonte: Elaboração própria.

Em seguida, com base em Callou, Moraes e Leite (2002 [1996]), Oliveira (1999), Guedes (2014) e Oushiro (2015), desprezaram-se as ocorrências de Apagamento, por ser um fenômeno que acontece tanto na variedade baiana quanto na variedade paulista; e porque as variantes do (-r) em coda que diferenciam esses dois dialetos são essencialmente a Fricativa velar/glotal, o Retroflexo e o Tepe (ver a Subseção 5.1). Na sequência, amalgamaram-se o Retroflexo e o Tepe, considerando que essas variantes são típicas de São Paulo (SP), sendo possivelmente a elas que os migrantes baianos se adaptam ao manter contato com os paulistas. Logo depois, testou-se, a partir de análises multivariadas de regressão logística de efeitos mistos com inclusão do Participante e do Item lexical como variáveis aleatórias, realizadas na plataforma R (R Core, 2023), a correlação do emprego do (-r) em posição de coda com os preditores sociais descritos na Seção 3 com os preditores linguísticos apresentados no Quadro 1 e com o preditor contextual de Estilo.

**Quadro 1** – Variáveis linguísticas testadas na análise do (-r) em coda

Variáveis	Níveis
Contexto fônico precedente	vogal anterior [i, e, ε] vogal central [a] vogal posterior [ɔ, o, u]
Contexto fônico seguinte	consoante labial [p, b, f, v, m] consoante coronal [t, d, d̥, ʃ, ʒ, tʃ, s, z, l, n ] consoante dorsal [k, g, h] pausa
Tonicidade da sílaba	átona tônica
Posição da sílaba	medial final
Classe morfológica	[+gramatical] (preposição e conjunção) [-gramatical] (verbo, substantivo, adjetivo, pronome, numeral, interjeição, advérbio)

Fonte: Elaboração própria.

As análises estatísticas sobre o uso do /r/ em posição de coda pelos migrantes baianos residentes na RMSP se respaldaram nas hipóteses aventadas para os preditores sociais apresentadas na Seção 3 e nas hipóteses listadas para os preditores linguísticos e contextual explicitadas a seguir:

(i) Segundo [Callou, Moraes e Leite \(2002 \[1996\], p. 472\)](#), em São Paulo (SP), o modelo estatístico reportou não haver relação do Tepe em posição medial com a vogal precedente, mas apontou haver desse preditor linguístico com o Apagamento em posição final. De acordo com esse resultado, a supressão do /r/ em posição final tende a ocorrer quando ele é precedido por vogal anterior ou por vogal central. Dito de outra maneira, na posição final, o Tepe tende a ser usado quando precedido por vogal posterior. Nos dados de [Oushiro \(2015, p. 107\)](#), o Retroflexo é levemente favorecido pelo contexto fônico precedente [+anterior]. Com base nesses resultados, aventa-se a hipótese de que os migrantes baianos residentes na RMSP utilizam mais o Retroflexo/Tepe quando o /r/ estiver precedido por vogal anterior e posterior.

(ii) O estudo de [Callou, Moraes e Leite \(2002 \[1996\]\)](#) mostra que, em posição medial, o Tepe tende a ser mais usado quando o segmento seguinte é uma consoante Fricativa alveolar. Quanto ao Apagamento em posição final, os resultados sugerem que esse fenômeno

tende a acontecer mais quando o /r/ é seguido de consoante. No estudo de Oushiro (2015, p. 107), o Retroflexo tende a ser relativamente mais realizado quando o /r/ é seguido de consoante [+coronal]. Assim, conjectura-se que os migrantes baianos na RMSP empregam mais o Retroflexo/Tepe quando o Contexto fônico seguinte é uma consoante coronal.

(iii) Os resultados de Callou, Moraes e Leite (2002 [1996]) sugerem que, nos dados de São Paulo (SP), não há relação entre o emprego do Tepe em posição medial e nem do Apagamento na posição final com a Tonicidade da sílaba. Já os dados de Oushiro (2015, p. 107) demonstram que o Retroflexo tende a ser mais empregado em sílabas tônicas. Sendo assim, levanta-se a hipótese de que o Retroflexo/Tepe, na fala dos migrantes baianos na RMSP, tende a ser mais usado em sílabas tônicas.

(iv) Na pesquisa de Oushiro (2015), verifica-se que o Retroflexo tende a ser mais empregado quando o /r/ está em posição final. Considerando esse resultado, conjectura-se que os migrantes baianos tendem a realizar mais o Retroflexo/Tepe em sílabas localizadas no final do vocábulo.

(v) Os dados de Callou, Moraes e Leite (2002 [1996]) indicam que, em São Paulo (SP), o emprego do Tepe em posição medial não tem diferença significativa quanto à Classe morfológica da palavra em que o /r/ é realizado. Por outro lado, tem relação com o Apagamento desse fonema na posição final, de maneira que este último fenômeno tende a ocorrer em verbos. Os dados de Oushiro (2015, p. 107) indicam que as classes dos verbos (raiz) e advérbios favorecem o emprego do Retroflexo. A partir desses resultados, conjectura-se que, na fala dos migrantes baianos, o Retroflexo/Tepe tende a ser mais recorrente em palavras [-gramatical] do que em palavras [+gramatical].

(vi) Para analisar o (-r) em diferentes graus de monitoramento, produziu-se, com base em Oushiro (2015, 2017), uma lista com 37 palavras (ver o Apêndice A) que tivessem combinações possíveis entre o contexto fônico precedente (as vogais [i, e, ε, a, o, u]), o contexto fônico seguinte (várias consoantes [b, d, g, p, t, k, dʒ, ʃ, ʒ, tʃ, f, v, s, z, l, m, n, h] ou pausa), a tonicidade da sílaba (átona e tônica) e a posição em que o /r/ se encontra (medial e final). Segundo os resultados de Oushiro (2015, p. 108), a conversação, a leitura de jornal e o depoimento favoreceram a realização do Retroflexo em oposição à lista de palavras. A expectativa para a análise do (-r) quanto ao Estilo então seria a de que os migrantes baianos tendem a realizar mais o Retroflexo/Tepe na leitura de palavras do que na conversação por considerar que a variante da comunidade anfitriã seria a de prestígio. Contudo como se pôde ver na pesquisa de Hora e Wetzels (2010) (ver a Seção 2), feita com uma amostra de fala de quatro paraibanos residentes em São Paulo (SP), em situação de fala mais monitorada, os migrantes paraibanos tenderam a empregar mais a forma Aspirada em relação ao Tepe. Oushiro (2021, 2024, no prelo[a]) encontrou padrões semelhantes aos de Hora e Wetzels (2010) na fala de 24 alagoanos residentes em Campinas (SP) quanto ao uso do (-r) e do (-s) em posição de

coda. Dessa maneira, considerando esses resultados, levanta-se a hipótese de que os migrantes baianos na RMSP tendem a realizar mais o Retroflexo/Tepe na conversação do que na leitura (texto e palavras).

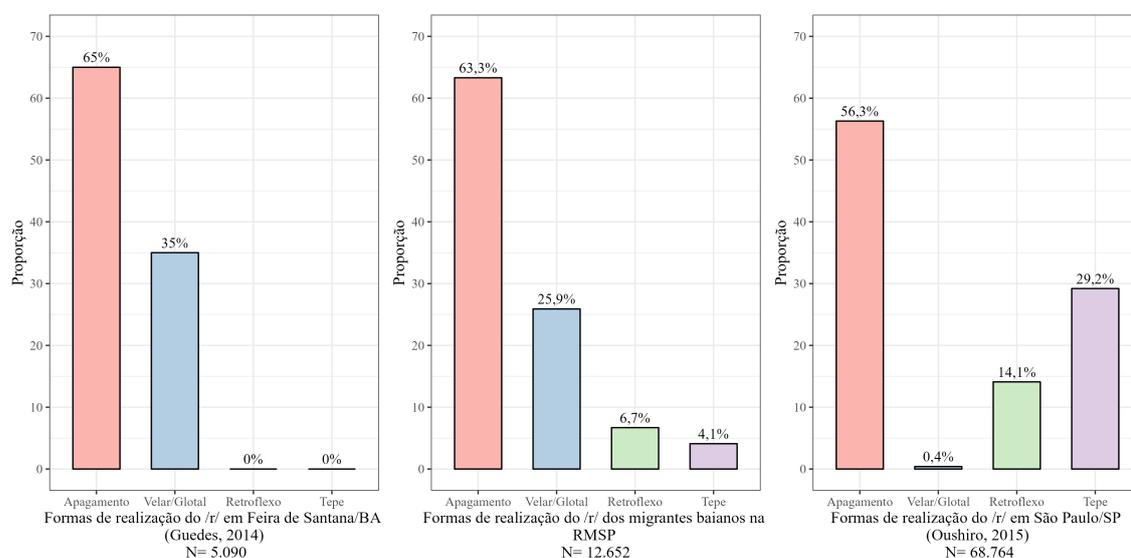
### 5.3 Resultados

Esta Subseção consta de duas partes. Na primeira, são delineadas as análises quantitativas do (-r) em coda da fala dos migrantes baianos na RMSP, enquanto a segunda parte aborda a interpretação dos resultados obtidos.

#### 5.3.1 Descrição dos achados

Antes da análise inferencial, comparam-se, na Figura 19, as taxas das variantes do (-r) em coda da fala dos migrantes baianos na RMSP com os dados de uso dessa variável sociolinguística encontrados em Feira de Santana (BA) por Guedes (2014) e com os dados dos róticos obtidos em São Paulo (SP) por Oushiro (2015). No primeiro gráfico da Figura 19, da esquerda para a direita, retomam-se os resultados da realização do (-r) no português popular e culto de Feira de Santana (BA) com base em Guedes (2014) e observa-se que ele é mais usado sob a forma de Apagamento (65%) e como Fricativa velar/glotal (35%). A autora não reportou a realização do /r/ em posição de coda como Retroflexo (0%) nem como Tepe (0%), o que indica que esses segmentos fonéticos nesse contexto são incharacterísticos desse dialeto. No segundo gráfico, apresentam-se os valores encontrados na fala dos migrantes e nota-se que, de 12.652 ocorrências de (-r) em coda, há 63,3% de Apagamento, 25,9% de Fricativa velar/glotal, 6,7% de Retroflexo e 4,1% de Tepe. No último gráfico, mostram-se os percentuais achados em São Paulo (SP) por Oushiro (2015) e verifica-se que há 56,3% de Apagamento, 0,4% de Fricativa velar/glotal, 14,1% de Retroflexo e 29,2% de Tepe.

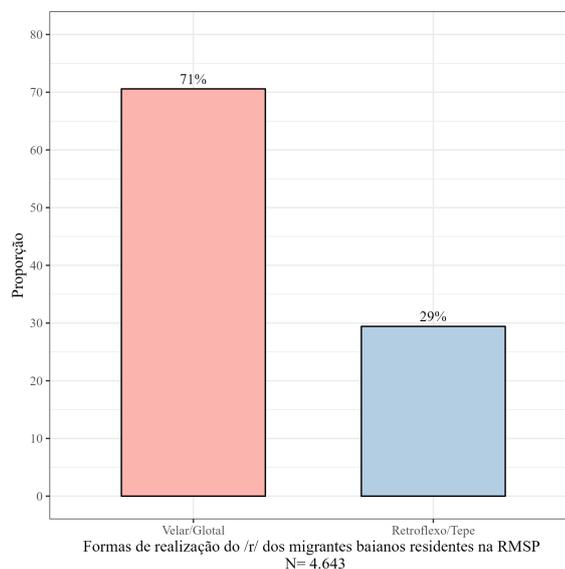
**Figura 19** – Comparação das proporções de uso do (-r) em coda silábica de baianos não migrantes, de migrantes baianos na RMSP e de paulistas não migrantes



Fonte: Elaboração própria.

Comparando os três gráficos da Figura 19, verifica-se que a população migrante se adapta quanto ao emprego do Retroflexo (6,7%) e do Tepe (4,1%) ao manter contato com a comunidade anfitriã. Para descrever o encaixamento social, linguístico e contextual dessas variantes na fala dos migrantes baianos residentes na RMSP, realizaram-se, na plataforma R (R Core, 2023), análises multivariadas de regressão logística de efeitos mistos cotejando a variação entre Retroflexo/Tepe e Fricativa velar/glotal. Sendo assim, após excluir os dados de Apagamento, analisaram-se 4.643 observações de /r/ em posição de coda medial e final, sendo 3.277 dados de Fricativa velar/glotal e 1.366 de Retroflexo/Tepe, cujas proporções reportam-se na Figura 20.

**Figura 20** – Proporção do (-r) em coda silábica da fala de migrantes baianos residentes na RMSP depois da exclusão dos dados de Apagamento e da junção das ocorrências do Retroflexo com o Tepe



Fonte: Elaboração própria.

A partir da Figura 20, nota-se que as Fricativas velar/glotal representam o total de 71%, enquanto o Retroflexo/Tepe totalizam 29%. Na Tabela 11, evidenciam-se os resultados do primeiro modelo de regressão logística de efeitos mistos. Antes de ler seus coeficientes, é pertinente destacar que não há ortogonalidade entre a variável Idade e a Idade de migração, entre a Idade e o Tempo de residência, assim como também não há ortogonalidade entre as variáveis Posição da sílaba e Contexto fônico seguinte. Por essa razão, nesse primeiro modelo, incluíram-se a Idade de migração, o Tempo de residência, o Contexto fônico seguinte junto às demais variáveis da rodada e excluíram-se a Idade e a Posição da sílaba. No segundo modelo, acrescentaram-se a Idade de migração, o Tempo de residência e a Posição da sílaba e descartaram-se a Idade e o Contexto fônico seguinte. No terceiro modelo, adicionaram-se a Idade e o Contexto fônico seguinte e eliminaram-se a Idade de migração, o Tempo de residência e a Posição da sílaba. Por fim, no quarto modelo, incluíram-se a Idade e a Posição da sílaba e foram excluídos a Idade de migração, o Tempo de residência e o Contexto fônico seguinte.

O primeiro modelo estatístico mostrou não haver correlação do uso do Retroflexo/Tepe com os preditores Sexo/Gênero ( $-0,028 \log odds, p = 0,956$ ), Nível de escolaridade ( $1,035 \log odds, p = 0,132$ ), Idade de migração ( $-0,002 \log odds, p = 0,955$ ), Grau de autoidentificação paulista declarada ( $0,062 \log odds, p = 0,482$ ), Tonicidade da sílaba ( $-0,007 \log odds, p = 0,963$ ) e Estilo ( $-0,258 \log odds, p = 0,124$ ). Em contrapartida, verificou-se a correlação do Retroflexo/Tepe com Tempo de residência ( $0,050 \log odds, p = 0,02$ ), Classe morfológica ( $-1,061 \log odds, p = 0,001$ ), a interação entre Rede de contato com baianos com Contexto fônico precedente ( $0,119 \log odds, p = 0,0006$ ) e a interação entre Rede de contato com baianos com Contexto

fônico seguinte (0,108 *logodds*,  $p = 0,0007$ ). Testou-se essa interação a partir da hipótese de que os migrantes baianos pertencentes a redes mais abertas utilizam regras diferentes para o emprego do Retroflexo/Tepe em relação àqueles que integram redes mais fechadas, como ocorreu na pesquisa de Guedes (2014), cujos resultados sugeriram que os migrantes alagoanos e paraibanos que tinham paulistas em suas redes tendiam a empregar as mesmas regras de uso do artigo definido diante de possessivos da comunidade anfitriã em relação àqueles que não tinham esse tipo de convivência (ver a Seção 2).

**Tabela 11** – Resultados do primeiro modelo de regressão logística de efeitos mistos com inclusão do Participante e do Item lexical como variáveis aleatórias quanto ao uso do /r/ Retroflexo/Tepe na fala dos migrantes baianos residentes na RMSP

*Intercept= -1,262.*

Variável	Apl./N	%	Est.	EP	Valor-z	<i>p</i>
<b>Sexo/Gênero</b>						
mulheres ( <i>ref.</i> )	866/3.071	28%				
homens	500/1.572	32%	-0,028	0,521	-0,054	0,956
<b>Escolaridade</b>						
Fundamental ( <i>ref.</i> )	758/2.340	32%				
Médio	608/2.303	26%	1,035	0,687	1,505	0,132
<b>Id. de migração</b>	–	–	-0,002	0,043	-0,056	0,955
<b>Autoident. paulista</b>	–	–	0,062	0,088	0,702	0,482
<b>Tonicidade</b>						
átona ( <i>ref.</i> )	978/3.055	32%				
tônica	388/1.588	24%	-0,007	0,155	-0,045	0,963
<b>Estilo</b>						
conversação ( <i>ref.</i> )	1.062/3.516	30%				
leitura	304/1.127	27%	-0,258	0,168	-1,536	0,124
<b>T. de residência</b>	–	–	0,050	0,022	2,286	0,02 *
<b>Rede</b>	–	–	-0,230	0,100	-2,296	0,02 *
<b>Cont. fôn. prec.</b>						
anterior ( <i>ref.</i> )	522/1.986	26%				
central	169/693	24%	0,169	0,328	0,516	0,605

Continua...

*Intercept= -1,262.*

Variável	Apl./N	%	Est.	EP	Valor-z	p
posterior	675/1.964	34%	-0,892	0,260	-3,433	0,0005 ***
<b>Cont. fôn. seg.</b>						
coronal ( <i>ref.</i> )	456/2.117	22%				
dorsal	453/1.056	43%	-0,132	0,291	-0,455	0,648
labial	367/1.204	30%	0,348	0,264	1,319	0,187
pausa	90/266	34%	0,290	0,432	0,671	0,502
<b>Rede*Cont. fôn. prec.</b>						
rede/anterior ( <i>ref.</i> )						
rede/central	–	–	-0,111	0,051	-2,142	0,032 *
rede/posterior	–	–	0,119	0,035	3,397	0,0006 ***
<b>Rede*Cont. fôn. seg.</b>						
rede/coronal ( <i>ref.</i> )						
rede/dorsal	–	–	0,108	0,040	2,670	0,007 **
rede/labial	–	–	0,033	0,040	0,830	0,406
rede/pausa	–	–	0,071	0,068	1,045	0,295
<b>Classe morf.</b>						
[+gramatical] ( <i>ref.</i> )	332/595	56%				
[-gramatical]	1.034/4.048	26%	-1,061	0,326	-3,250	0,001 **

Modelo1: glmer (VD ~ SEXO+NIVEL.ESCOLARIDADE+Estilo+IDADE.MIGRACAO+TEMPO.RESIDENCIA+  
 AUTOIDENTIFICACAO.PAULISTA+ CONT.FON.PREC\*REDE+REDE\*CONT.FON.SEG+  
 TONICIDADE+CLASSE.MORFOLOGICA+(1 | INFORMANTE)+ (1 | ITEM.LEXICAL), data = dados4)

Fonte: Elaboração própria.

No segundo modelo, em que se excluíram da análise os preditores Idade e Contexto fônico seguinte e incluíram-se a Idade de migração, o Tempo de residência e a Posição da sílaba, por não serem ortogonais entre si, notou-se que, quando se adicionou a Posição da sílaba, o preditor Estilo se correlacionou com a realização do Retroflexo/Tepe, sinalizando ser um caso interação entre esses dois preditores. A estimativa (-1,727 *logodds*,  $p = 0,001$ ) dessa interação, apresentada na Tabela 12<sup>24</sup>, sugere que o resultado de Estilo só pode ser explicado em face do preditor linguístico Posição da sílaba.

<sup>24</sup> Exibem-se apenas os coeficientes dos preditores não incluídos na Tabela 11.

**Tabela 12** – Resultados do segundo modelo de regressão logística de efeitos mistos com inclusão do Participante e do Item lexical como variáveis aleatórias quanto ao uso do /r/ Retroflexo/Tepe da fala dos migrantes baianos na RMSP

*Intercept= -1,370.*

Variável	Apl./N	%	Est.	EP	Valor-z	p
<b>Estilo</b>						
conversaço (ref.)	1.062/3.516	30%				
leitura	304/1.127	27%	0,066	0,187	0,355	0,722
<b>Posição</b>						
medial (ref.)	1.109/4.032	28%				
final	257/611	42%	2,125	0,300	7,066	< 0,001 ***
<b>Estilo*Posição</b>						
conv./med. (ref.)	–	–				
leit./fin.	–	–	-1,727	0,390	-4,422	< 0,001 ***

Modelo2: glmer (VD ~ SEXO+NIVEL.ESCOLARIDADE+Estilo\*POSICAO.R+IDADE.MIGRACAO+TEMPO.RESIDENCIA+AUTOIDENTIFICACAO.PAULISTA+CONT.FON.PREC\*REDE+TONICIDADE+CLASSE.MORFOLOGICA+(1 | INFORMANTE)+(1 | ITEM.LEXICAL), data = dados4)

Fonte: Elaboração própria.

No terceiro modelo, em que se adicionaram a Idade e o Contexto fônico seguinte na análise e descartaram-se a Idade de migração, o Tempo de residência e a Posição da sílaba, vê-se, a partir da Tabela 13, na qual se destacam apenas os coeficientes do preditor Idade, que há diferença significativa (0,046 *logodds*,  $p = 0,038$ ) para o uso do Retroflexo/Tepe entre os migrantes baianos mais jovens e mais velhos.

**Tabela 13** – Resultados do terceiro modelo de regressão logística de efeitos mistos com inclusão do Participante e do Item lexical como variáveis aleatórias quanto ao uso do /r/ Retroflexo/Tepe da fala dos migrantes baianos na RMSP

*Intercept= -2,416.*

Variável	Est.	EP	Valor-z	p
<b>Idade (contínua)</b>	0,046	0,022	2,067	0,038 *

Modelo3: glmer (VD ~ SEXO+NIVEL.ESCOLARIDADE+Estilo+IDADE+AUTOIDENTIFICACAO.PAULISTA+CONT.FON.PREC\*REDE+CONT.FON.SEG\*REDE+TONICIDADE+CLASSE.MORFOLOGICA+(1 | INFORMANTE)+(1 | ITEM.LEXICAL), data = dados4)

Fonte: Elaboração própria.

O quarto modelo estatístico<sup>25</sup>, no qual se acrescentaram a Idade e a Posição da sílaba e eliminaram-se a Idade de migração, o Tempo de residência e o Contexto fônico seguinte, não apresentou resultados diferentes dos expostos nos modelos anteriores, logo não houve a necessidade de exibir seus valores em outra tabela.

### 5.3.2 Discussão dos achados

As análises da realização do /r/ em posição de coda silábica na fala dos migrantes baianos residentes na RMSP revelam que essa população se adapta ao uso do Tepe (4,1%) e do Retroflexo (6,7%). Conforme se viu na Seção 2, a adaptação ao uso do Retroflexo também ocorreu com os baianos residentes em Bauru (SP), cuja amostra apresentou 75% da forma Aspirada e 25% da Retroflexa (Oliveira, 2020). Isso aconteceu também na fala de migrantes alagoanos e paraiibanos residentes na Região Metropolitana de Campinas (SP), em cuja amostra se verificou o percentual de 31,6% de Tepe/Retroflexo e 68,4% da forma Aspirada (Oushiro, 2020b).

Os processos de adaptação linguística para essas variantes fonéticas, porém, não são os mesmos encontrados nas análises dos itens lexicais. Se se considerar que os baianos não migrantes tendem a empregar o Tepe nos contextos em que o /r/ é seguido de vogal como, por exemplo, em “mar azul”, e tendem a utilizar as formas Velar ou Glotal em outros contextos Callou, Moraes e Leite (2002 [1996]), Oliveira (1999) e Guedes (2014), como em “cartão”, assume-se que o Tepe não é um fone incharacterístico no dialeto baiano. O que se pode classificar como atípico nesse dialeto é o seu uso seguido de consoante, por exemplo. Sendo assim, conclui-se que os migrantes baianos, ao manterem contato com os paulistanos, tendem a *ampliar o contexto* de uso de um fone existente em seu repertório linguístico. Em outras palavras, o uso deixa de ser em contexto restrito, conforme as regras fonológicas da variedade baiana, e passa a ser empregado em contextos mais amplos, seguindo as regras da gramática da comunidade anfitriã. O mesmo processo, entretanto, não se aplica na adaptação para o emprego do Retroflexo. Baseando-se em Callou, Moraes e Leite (2002 [1996]), vê-se que essa variante é proeminente no dialeto paulista, mas é atípica na variedade baiana. Dessa maneira, não se pode assumir que acontece a ampliação de contexto como se viu com o emprego do Tepe, mas sim que ocorre a *aquisição* de uma variante irregular no dialeto baiano e típica da comunidade anfitriã.

Esses casos de adaptação podem estar relacionados à saliência social das variantes paulistanas (Tepe e Retroflexo) na percepção dos migrantes baianos na RMSP. Depois da leitura da lista de palavras, perguntava-se aos participantes quais daquelas palavras os paulistas

<sup>25</sup> Modelo4: glmer (VD ~ SEXO+NIVEL.ESCOLARIDADE+IDADE+ AUTOIDENTIFICACAO.PAULISTA +CONT.FON.PREC\*REDE +Estilo\*POSICAO.R+TONICIDADE+CLASSE.MORFOLOGICA+(1 | INFORMANTE)+(1 | ITEM.LEXICAL), data = dados4)

falariam de modo diferente. Todos eles mencionaram vocábulos com o /r/ em coda, cuja pronúncia era exemplificada sempre com o Retroflexo [ɻ] ou algo próximo a esse segmento, como uma espécie de glide semialongado [iː], e com o Tepe [r]. Os comentários metalinguísticos de RodrigoC (10-a), de MaraA (10-b), de GleideS (10-c) e de MarioS (10-d) são alguns exemplos de que os baianos notam as diferenças entre os dialetos baiano e paulista a partir também do uso de variantes do (-r).

- (10) a. D1: *qual dessas palavras você acha que o paulista fala diferente?*  
 S1: *ih (Rodrigo C)*  
 D1: *tem alguma que ele fala diferente? que você acha?*  
 S1: *deixa eu ver... eu não tenho eu não acho muita diferença do paulista falar porque onde que eu morava na Bahia todo mundo fala tipo quase igual aqui então não tem muita*  
 D1: *uhum*  
 S1: *não vejo muita diferença / alguns puxam mais algumas letras*  
 D1: *quais letras você acha que/?*  
 S1: *“melho[ɻ]” / “mora[ɻ]” / tipo isso né*  
 D1: *uhum*  
 S1: *puxa alguma é puxa o R “mulhe[ɻ]” / só assim mesmo algumas alguns puxam uma o final da palavra ali um R um / “a[ɻ]gola” / eles falam*
- b. D1: *tem alguma diferença entre o paulista*  
 S1: *ah tem... tem / eles puxa muito o R né / tudo que eles falam vai o R (MaraA)*  
 D1: *como é que eles falariam (então?)*  
 S1: *e a gente não consegue... ah tipo “po[iː]ta” eh “ca[iː]ne” a maioria/ né tudo deles*
- c. D1: *o paulista fala diferente?*  
 S1: *ah meu filho esse R aqui eles treme até umas hora (GleideS)*  
 D1: *como é que eles falariam?*  
 S1: *ai eu não consigo falar igual aquele povo aquele “po[ɻ]ta”*  
 S1: *olha esse é do interior né*
- d. D1: *tem alguma que ele falaria diferente aí dessas palavras?*  
 S2: *ele fala diferente sobre o sotaque né (MarioS)*  
 D1: *qual delas você acha que ele falaria diferente da gente?*  
 S2: *quase quase todas velho*  
 D1: *ah é?*  
 S2: *porque nós nós que não somos daqui / querendo ou não nós tem aquele por*

*exemplo / “ve[r]dade” / nós não fala assim ve[r]dade né? e eles lá fala assim “ve[r]dade”... “não é ve[r]dade?”/ então já é diferente*

*D1: já é diferente né?*

*S2: é... a não ser que seja o paulistano que no caso nossos filhos / porque **o nossos filho vai na escola** / está na escola / mas **a primeira escola é a nossa** / então... **normalmente fala mais ou menos igual nós né** / mas quem é paulistano de paulistano / sempre coloca esse... esse sotaque diferente*

No Exemplo (10-a), RodrigoC diz que, da leitura que ele fez, os paulistas falariam de modo diferente as palavras com o /r/, de modo que haveria um “puxamento” da pronúncia quando esse fonema está no final delas e exemplifica, reproduzindo algumas palavras da lista, usando o Retroflexo. Em (10-b), MaraA também comenta sobre o “puxamento” do R, mas diferentemente de RodrigoC, ela não consegue reproduzir o Retroflexo. Na tentativa, ela reproduz um glide semialongado [iː], utilizado, inclusive, por alguns participantes na fala semiespontânea que, na codificação, considerou-se como Retroflexo. Em (10-c), GleideS menciona o /r/ “tremido”, fazendo alusão talvez ao Tepe ou à Vibrante múltipla, cuja imitação ela não consegue fazer. Por outro lado, ela reproduz um Retroflexo na palavra “porta”, chamando atenção para o fato de que aquele exemplo seria uma imitação de alguém do interior de São Paulo. Já em (10-d), MarioS aponta como diferença a pronúncia do /r/ sendo realizado como Tepe. MarioS explica que essa forma de falar seria de paulistanos e cita a fala de seus filhos, cujo aspecto é “mais ou menos” igual à maneira como ele fala, justificando, em linhas gerais, que seria uma mistura do que os filhos aprenderam na escola e na família. Segundo Trudgill (1986), os migrantes tendem a se adaptar linguisticamente à comunidade anfitriã principalmente quanto aos estereótipos presentes nela. Dessa maneira, é plausível que os migrantes baianos tenham se adaptado para o uso do Tepe e do Retroflexo por serem talvez os principais estereótipos do falar paulistano.

A percepção de MarioS no Exemplo (10-d) evidencia que pode haver adaptações na fala dos migrantes sob influência da variedade paulistana, mas também pode acontecer o contrário. Isto é, os migrantes podem influenciar a fala dos nativos da comunidade anfitriã, como possivelmente teria acontecido com os filhos de MarioS. Em conformidade com os dados de Oushiro (2015), o Retroflexo em São Paulo (SP) é mais usado por moradores da periferia, de classe social baixa e paulistanos filhos de migrantes nortistas/nordestinos e filhos de migrantes do interior de São Paulo, de Minas Gerais e do Paraná. Levando em conta que a periferia dessa cidade foi formada especialmente por migrantes do interior de São Paulo, nordestinos e mineiros (Fontes, 2008) e observando a proporção do Retroflexo dos baianos, infere-se que o emprego desse segmento fônico por parte dos paulistanos filhos de nordestinos tenha sido em decorrência de um *nivelamento linguístico* (Trudgill, 1986). Esse fenômeno teria ocorrido na periferia de

São Paulo (SP), entre a primeira geração de nordestinos migrantes e seus filhos nascidos nessa cidade, a partir da socialização de crianças com os familiares (migrantes nordestinos) e com os seus pares nas escolas, supostamente filhos de migrantes do interior desse estado.

Com os coeficientes dos modelos de regressão logística apresentados nas Tabelas 11, 12 e 13, pode-se estimar o padrão de uso do Retroflexo/Tepe na fala dos migrantes baianos na RMSP. Apesar de o preditor Idade de migração não ter apresentado correlação com o fenômeno, seu resultado deve ser acolhido com lucidez, pois ele pode acender luz para interpretações relevantes. A ausência de correlação da variável resposta com a Idade de migração possibilita conjecturar que o processo de adaptação linguística pode variar de acordo com o lugar onde o grupo de migrante reside, se na capital ou no interior. O resultado das análises desta amostra para este preditor foge ao que se tem visto na literatura sobre a fala de migrantes, sobretudo por ser uma variável fonético-fonológica. Com base em Chambers (1992), a expectativa para variáveis dessa natureza é a de que quanto mais cedo o migrante chega à comunidade anfitriã, mais aspectos linguísticos dela ele adquire. No estudo de Oushiro (2020b, p. 82) sobre os alagoanos e paraibanos em Campinas (SP), a Idade de migração se correlacionou com (-r), de modo que quem chegou mais cedo tendeu a usar mais o Retroflexo em comparação com quem chegou mais tarde. A pesquisa de Oliveira (2020) sobre baianos em Bauru (SP) também apresentou esse mesmo padrão.

Ao comparar esses resultados de Oushiro (2020b) e de Oliveira (2020) com os dados vistos nesta amostra, infere-se que os migrantes baianos tendem a se adaptar quanto à realização do Retroflexo/Tepe independentemente da idade com que chegam a RMSP, talvez porque a pressão social seja maior na capital em relação a cidades do interior. Numa comunicação pessoal com Ferraz (2023), mestrando em Linguística do IEL/UNICAMP que pesquisa a fala de baianos em Palmares Paulista (SP), ele relatou que, durante as entrevistas, poucos participantes comentaram ter sofrido algum tipo de preconceito por serem baianos nessa cidade, diferentemente do que os baianos na capital paulista comentaram (ver a Seção 3). Dessa forma, é provável que a pressão vivida pelos baianos residentes na Grande São Paulo esteja associada à xenofobia vivida por eles acima de tudo nos espaços de trabalho.

É plausível que essa população sofra a pressão do mercado de trabalho essencialmente por carregar o estigma de sua proveniência. De acordo com Fontes (2008), na década de 1950, quando explodiu a migração para São Paulo (SP), os paulistas utilizavam pejorativamente o termo “baiano” para todos os migrantes “nortistas” que demonstravam ser desprovidos de qualquer tipo de conhecimento, sobretudo aqueles relacionados ao ambiente profissional. Ainda segundo Fontes (2008), os migrantes de outras regiões do Nordeste não gostavam dessa generalização e, muitas vezes, faziam questão não só de dizer que não eram baianos como também de não se misturar com eles. Quase 70 anos depois, percebe-se que os baianos na RMSP continuam sofrendo preconceito no espaço de trabalho devido à sua origem, conforme ilustram

os depoimentos de AlanaS no exemplo (11-a), de SofiaB em (11-b) e de MaraA em (11-c).

- (11) a. D1:  *você já sofreu algum tipo de preconceito por ser baiana?*  
 S1:  *já... muitas (AlanaS)*  
 D1:  *pode dar pelo menos uma situação aqui?*  
 S1:  *uma das **mais fortes** foi no **meu primeiro trabalho** né? não sei se era pelo **meu jeito de falar** ou **meu jeito de se vestir** / uma das gerentes falou que eu não servia pra estar naquele cargo e sim para lavar banheiro*
- b. D1:  *e você já passaram... já perceberam algum / alguma cena de preconceito por ser baiano? / alguém já.. já teve uma atitude preconceituosa*  
 S2:  *não... comigo não / existe né... mas eu nunca / eu nunca tive não graças a Deus né (ArmandoM)*  
 S1:  *eu já / eu já? (SofiaS)*  
 D1:  *como foi isso aí?*  
 S1:  *eu **em entrevistas de trabalho** tipo assim / por ser baiana / e eles terem preferência /mo/ é antes de entrar nesse último emprego eu fiz várias entrevistas de emprego / teve lugares que tipo assim minha referência era boa mas **não me aceitaram por conta de não ser / por ser baiana** / entendeu? / então pra eles o pessoal certo é o do sul / baiano tem defeito / fofoqueira não sei o que não sei o que / então assim*  
 D1:  *nossa*  
 S1:  *eu passei por essas situações e também / eu escuto muito muito **no meu trabalho atual** né / e já escutei em outros preconceito tipo de baiano que baiano é preguiçoso que baiano vive de bolsa que baiano é isso que baiano é aquilo*
- c. D1:  *você acha que é baiano aqui sofre preconceito?*  
 S1:  *se sofre preconceito? / muito (MaraA)*  
 D1:  *você já sofreu algum preconceito por ser baiana?*  
 S1:  *já já*  
 D1:  *pode contar como foi?*  
 S1:  *então eu já eu já sofri muito preconceito **em serviço***

Desse modo, o estigma de ser baiano quiçá seja o motivo para que a Idade de migração dos baianos na RMSP não tenha apresentado diferença significativa quanto ao uso do Retroflexo/Tepe. Possivelmente quem mais cedo chegou a São Paulo (SP) adquiriu essa variante como previu Chambers (1992), porém aqueles que chegaram mais tarde, principalmente em idade de trabalho, tiveram que se adaptar para essa variante por uma necessidade de reduzir as características linguísticas que remetessem ao falar baiano, como a variante Aspirada. O

depoimento de GleideS, em (12), sustenta essa hipótese.

- (12) S1: *ai às vezes as pessoas têm que se / mudar o sotaque tal / eu nunca me submeti a isso / ah a não ser em uma vaga de emprego que falam que que tinha que falar o R / que eu não falava o R / é o R / o R / da quase né / enfim / mas eu engoli essa questão porque era uma palavra que ela falou fala mais o R / porque como que você vai ser / trabalhar com telemarketing / seria melhor que você falasse o R mais mais forte pra mostrar que o R / está saindo / não lembro que diacho de palavra que era / ah tá mas ela era um doce a menina eu vi que ela estava tentando me ajudar né mas é situações assim ridículas né / fala sério (GleideS)*

No Exemplo (12), a participante comenta que, certa vez, ela precisou mudar a forma de falar quanto à pronúncia do R. Segundo GleideS, sua supervisora pediu que ela falasse o R “mais forte”, porque, de acordo com a gestora, GleideS não o pronunciava. Talvez por causa dessa orientação no trabalho, a variante Fricativa velar/glotal tenha dado espaço para as variantes da comunidade anfitriã, tendo em vista que, na fala de GleideS, de 244 realizações de /r/ em coda, ela empregou 16% de Retroflexo/Tepe. De acordo com Lopes (1964, p. 69 *apud* Fontes, 2008), na década de 1950, na fábrica Nitro Química, localizada em São Miguel Paulista (SP), era muito comum ouvir comentários pejorativos de paulistas contra nordestinos sobre o “sotaque”, o “comportamento” e os “costumes”. Logo a vivência de GleideS não pode ser considerada um fato isolado e visto como consequência da função de *telemarketing*. Assim, se se assumir que muitas mulheres foram/vão para São Paulo (SP) trabalhar como babá, como SofiaB, não seria raro ver as mães das crianças, talvez preocupadas com a aquisição da linguagem dos filhos, cobrando das cuidadoras certas maneiras de falar. Também não seria raro encontrar migrantes baianos que lidam com o público sendo cobrados a falar com algumas características a fim de satisfazer os clientes.

O depoimento de GleideS chama atenção ainda para o fato de que, na adaptação linguística, os migrantes podem ser ativos na escolha do uso de determinada variante da comunidade anfitriã. Ou seja, o agenciamento estilístico (Eckert, 2012) pode ser praticado pelos migrantes a fim de construir uma *persona* que apresente características dialetais de um nativo a fim de garantir um espaço no mercado de trabalho. Para estudos futuros, sugere-se fortemente analisar as adaptações linguísticas ocasionadas pelo contato entre dialetos mutuamente inteligíveis na perspectiva da *Terceira Onda da Sociolinguística* (Eckert, 2012), uma vez que há indícios de que os migrantes baianos têm consciência dos significados sociais de determinadas variáveis na comunidade anfitriã, como o (-r) em coda, por exemplo.

No que diz respeito aos preditores que apresentaram diferença significativa nos mo-

delos estatísticos, verificou-se que há correlação da variável resposta com o Tempo de residência, de modo que os baianos residentes há mais tempo em São Paulo (SP) tendem a utilizar mais o Retroflexo/Tepe em comparação com quem está há menos tempo, conforme se previu. Esse também foi o resultado visto no estudo de Oliveira (2020) sobre os migrantes baianos em Bauru (SP) e no de Oushiro (2020b) acerca de alagoanos e paraibanos em Campinas (SP). Na correlação com o preditor Idade, viu-se que os migrantes baianos mais velhos tendem a realizar mais Retroflexo/Tepe do que os mais jovens. Esse resultado revela que os migrantes baianos mais jovens tendem a seguir um padrão oposto ao encontrado por Oushiro (2015) na fala dos paulistanos jovens quanto ao emprego do Retroflexo.

Conforme mencionado no início da descrição dos modelos estatísticos, não há ortogonalidade entre os preditores Idade e Tempo de residência, porque um jovem de 18 anos, por exemplo, não estaria em São Paulo (SP) há mais de 50 anos. Além disso, nota-se que esses dois preditores não são totalmente independentes, porque os migrantes baianos mais velhos são aqueles que estão em São Paulo (SP) há mais tempo. Sendo assim, é plausível que o resultado da correlação da variável resposta com a Idade do participante seja reflexo de algumas mudanças sociais, sobretudo do comportamento do migrante em relação ao empoderamento regionalista e ao seu posicionamento ideológico na comunidade anfitriã.

A primeira mudança a ser ponderada é a sanção da *Lei nº 9.459, de 13 de maio de 1997*, que altera a *Lei nº 7.716, de 05 de janeiro de 1989*, definindo crime qualquer prática, indução ou incitação de *discriminação ou preconceito de procedência nacional* (Brasil, 1997a). O conhecimento dessa Lei fez com que AlanaS, por exemplo, quando foi vítima de xenofobia no trabalho, procurasse uma delegacia para denunciar a agressora. De igual modo, MaraA ameaçou processar a vizinha por causa de preconceito. Nos últimos tempos, têm se escancarado nas mídias sociais casos de xenofobia, principalmente contra nordestinos. Muito provavelmente, os exemplos divulgados estão servindo de agentes reguladores desse tipo de crime e estão contribuindo para que os baianos mais jovens não se sintam intimidados com a maneira de falar. Vinculado a isso está o trabalho promovido pela Educação Brasileira desde os anos 1990 em combate ao preconceito social e linguístico feito com base nos *Parâmetros Curriculares Nacionais de Língua Portuguesa*, que nortearam o ensino de língua materna até quase os primeiros 20 anos do século XXI. O acesso à educação formal de adultos a partir da *Educação de Jovens e Adultos – EJA* é assegurado pela *Lei nº 9.394, de 20 de dezembro de 1996* (Brasil, 1996). Em paralelo a isso, estão os avanços das pesquisas na área da *Sociolinguística Educacional*, cujos resultados têm contribuído para a elaboração de material didático que respeite a diversidade linguística do Brasil (Bortoni-Ricardo, 2005). Outro fator é a ascensão da economia do Nordeste, que desacelerou a saída de naturais e se tornou atração, especialmente a Bahia, para pessoas de outras regiões (IBGE, 2000, 2010) e para muitos que buscam uma migração de retorno. Esse empoderamento econômico talvez coloque os nordestinos não mais como reféns de uma região

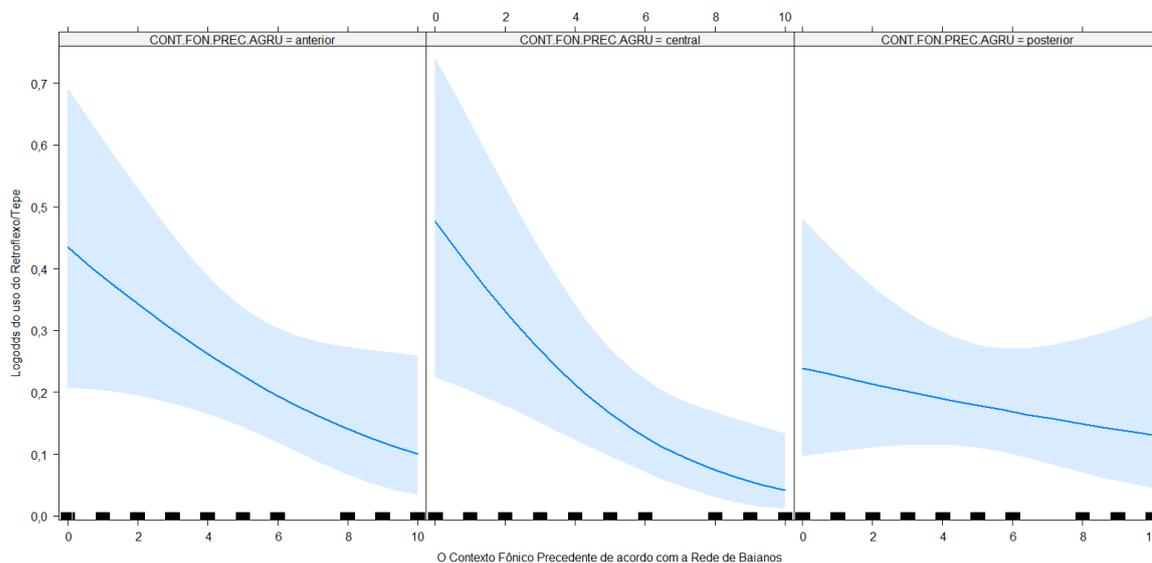
desenvolvida mas os coloque numa zona de confiança e segurança. Isso quiçá ocorra por sabermos que não lhes faltarão acolhimento e oportunidades no lugar de onde saíram, ou seja, talvez eles se sintam livres para transitar pelos espaços de origem e destino sem, possivelmente, muitos danos. Isso fica nítido no depoimento de JaneteA, em (13).

- (13) S1: *não vou dizer que eu vou falar mal do paulista / não vou dizer que eu vou falar mal de todos / não teve / eu encontrei paulistas bom entendeu que gosta de nordestino mas também encontrei muito paulista que não gosta do nordestino / eu tiro pelo fato de **eu trabalhar com o público** / tem uns que olha assim fala assim aí esse nordestino esse baiano / é esse cabeçudo / esse **tipo tem aquele preconceito** / não sei assim*  
 D1: *mas já você já teve alguma situação com você de preconceito? ter sofrido preconceito*  
 S1: *ah quando você chega num emprego a primeira coisa que você sofre é o seu sotaque / pode ver todo mundo fala / **tipo eles ficam imitando você dá risada** / às vezes você fala eles começa a dar risada aí você fica perguntando ué por que ele está rindo né? / tipo eles acham engraçado nosso sotaque / já tem gente que conversa comigo fala nossa é seu sotaque é lindo é lindo seu sotaque / tipo tem paulista que gosta / tem paulista que não gosta / então ah não conhece as nossas origem não conhece nosso estado / e fala tem mania de falar ah lá na Bahia não tem água / ah lá na Bahia não tem isso lá na Bahia não tem aquilo se desfaz da nossa / da nossa origem sem ao menos conhecer*  
 D1: *uhum*  
 S1: *e muitos eu tenho certeza que muitos não gostava / não curtia né / não sabia nem o que era chegar em Salvador (BA) olha que a Bahia é grande / o povo fala logo de Salvador (BA) / e de repente eles vão eles acaba gostando / tipo né fala assim nossa é uma cidade / maravilhosa é um estado maravilhoso / **eu não troco a minha Bahia por São Paulo***  
 D1: *uhum*  
 S1: *já disse os paulista que fique com o São Paulo deles / porque eu não troco minha Bahia por São Paulo*

JaneteA tem 27 anos e está em São Paulo (SP) desde os 17. Ela deixa claro, em seu depoimento, que já sofreu preconceito na capital paulista, principalmente dos colegas de trabalho e do público que ela atende, seja por causa do sotaque, seja pelas características físicas, ou pela escassez de bens de consumo em seu lugar de origem. Todavia ela tem a consciência dos aspectos positivos da Bahia e não se deixa levar pelos comentários depreciativos. Ela sente orgulho do lugar onde nasceu por saber que a realidade é outra. Todas essas mudanças sociais podem ter favorecido para que os migrantes mais jovens empreguem mais as variantes típicas da Bahia em relação ao Retroflexo/Tepe.

O primeiro modelo estatístico revelou haver interação entre o preditor social Rede de contato com baianos com duas variáveis linguísticas, o Contexto fônico precedente e o Contexto fônico seguinte. Chama atenção esse tipo de interação, porque se trata de um preditor social e dois linguísticos, cujos resultados evidenciam haver regras gramaticais distintas para o uso do Retroflexo/Tepe em mais de um grupo social, neste caso, aqueles baianos que têm menos contato com outros baianos (rede aberta) e aqueles que têm mais contato com outros baianos (rede fechada). A expectativa para esse preditor social era de que quanto menos contato com baianos, maior o uso de Retroflexo/Tepe; e de que quanto mais contato com baianos, menor a realização do Retroflexo/Tepe. O esperado para o Contexto fônico precedente era de que os migrantes baianos usassem o Retroflexo/Tepe quando o /r/ fosse precedido por vogal anterior e posterior em relação à vogal central. Esperava-se para o Contexto fônico seguinte que o Retroflexo/Tepe fosse favorecido por consoante coronal em relação a consoante labial, a dorsal ou a pausa. Na Figura 21, os três gráficos apresentam a aplicação do Retroflexo/Tepe no Contexto fônico precedente com vogal anterior (à esquerda), com vogal central (meio) e com vogal posterior (à direita), a depender do grau de contato que o migrante tem com outros baianos.

**Figura 21** – Interação das variáveis Contexto fônico precedente e Rede de contato com baianos em relação ao uso do /r/ Retroflexo/Tepe de migrantes baianos residentes na RMSP



Fonte: Elaboração própria.

Os dois primeiros gráficos, da esquerda para a direita, são semelhantes quanto à linha de regressão descendente. A partir dela constata-se que, quanto menos contato com outros baianos o migrante tem (à esquerda), maior é a realização do Retroflexo/Tepe quando o segmento fônico é precedido por vogal anterior e vogal central e quanto maior o contato com

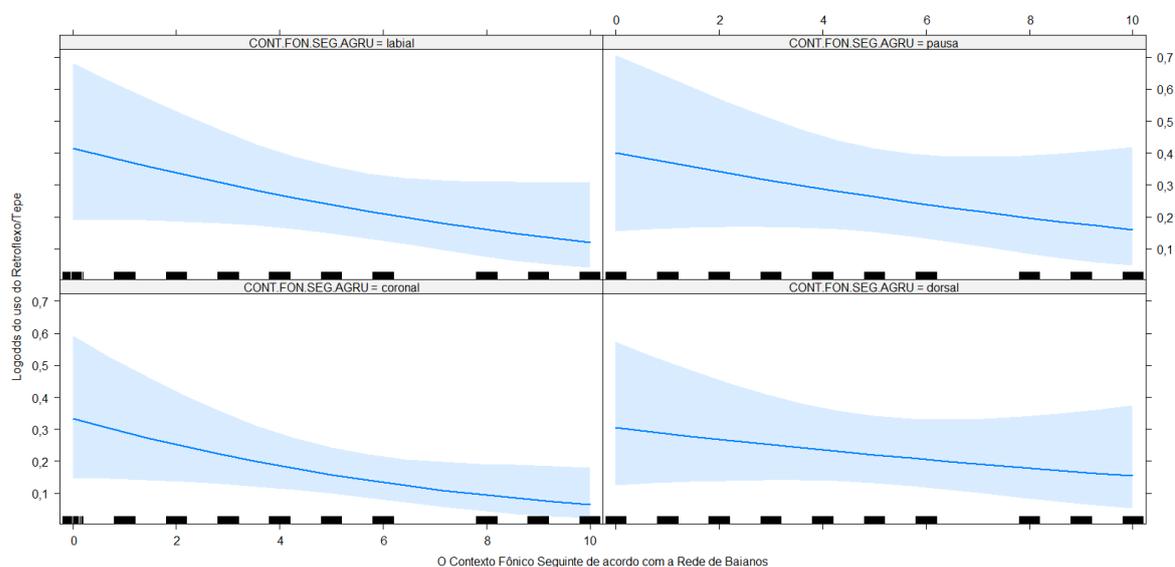
outros baianos (à direita), menor é o uso das variantes paulistas nesses contextos fônicos. No último gráfico consta a distribuição do Retroflexo/Tepe precedido por vogal posterior quando empregado pelos participantes nesses mesmos contextos sociais. A linha de regressão descendente apresenta uma leve diferença nos dados dos migrantes que têm mais contato com outros baianos em relação aos dois primeiros contextos linguísticos, sugerindo que eles empregam, mas em menor quantidade, essa mesma regra. Na amostra, houve 483 ocorrências da preposição “por” das quais 370 (77%) foram realizadas com o Retroflexo/Tepe e 113 (23%) como Fricativa velar/glotal. Já a conjunção “porque” apresentou 447 aplicações, das quais 247 (55%) foram empregadas com o Retroflexo/Tepe e 200 (45%) como Fricativa velar/glotal. Dessa maneira, é provável que os migrantes de redes mais fechadas realizem o Retroflexo/Tepe excepcionalmente em palavras mais frequentes como “porque” e “por” e os de rede mais aberta empreguem-no nessas e em outras palavras que tenham esse mesmo contexto linguístico.

Com esses resultados, verifica-se não só que os migrantes baianos tendem a empregar o Retroflexo/Tepe quando o /r/ é precedido por vogal anterior e a posterior, mas também pode-se afirmar que eles usam o Retroflexo/Tepe no contexto menos esperado, isto é, quando o /r/ é precedido por vogal central. Sendo assim, pode-se concluir que os baianos que convivem menos com outros baianos tendem a realizar mais o Retroflexo/Tepe em todos os contextos fônicos precedentes, mas aqueles que integram redes mais fechadas tendem a bloquear a aplicação dessa variante quando é precedida de vogal anterior e central, mas não se restringem quando o rótico é precedido por vogal posterior.

Na Figura 22, em que se ilustra a interação entre o Contexto fônico seguinte e a Rede de contato com baianos, observa-se que, quando o /r/ é seguido de pausa ou de consoante dorsal, praticamente não há diferença quanto à aplicação do Retroflexo/Tepe entre os migrantes de redes mais abertas e mais fechadas. O resultado do primeiro contexto talvez tenha uma relação com a saliência do som do /r/ quando é seguido de pausa. Já o resultado do segundo contexto, possivelmente está atrelado à limitação do uso do Retroflexo/Tepe nas palavras mais frequentes por parte dos baianos de redes mais fechadas. Por outro lado, verifica-se que essa variante tende a ocorrer quando o rótico é seguido de consoante labial e coronal na fala dos baianos que integram as redes abertas em comparação com os de redes fechadas.

Com base nesses dados, constata-se que o Retroflexo/Tepe tende a ser mais realizado quando são seguidos de consoante coronal, como também se nota o uso dessa variante em contextos cuja expectativa era menor, como quando o /r/ é seguido de consoante dorsal, labial e de pausa. Os resultados dessas duas interações permitem também acatar parcialmente a hipótese para o preditor Rede de contato com baianos de que os baianos que integram rede mais abertas tendem a aplicar mais o Retroflexo/Tepe em relação aos baianos que integram redes mais fechadas.

**Figura 22** – Interação das variáveis Contexto fônico seguinte e Rede de contato com baianos em relação ao uso do /r/ Retroflexo/Tepe na fala dos migrantes baianos residentes na RMSP



Fonte: Elaboração própria.

A diferença entre esses padrões das redes pode estar atrelada a avaliações negativas dos próprios migrantes baianos quanto às suas adaptações linguísticas na RMSP. Veja-se o comentário em (14) como exemplo disso.

- (14) S1: *então assim eu ensinei eu aprendi na Bahia / a a o português melhor do que aqui eu estudei mesmo naquela época / eles ensinavam bem o português / então assim eles fala eh por exemplo... eh eu falo assim ‘ô Z. C. cadê a A. G.?’ né / aí o lá na Bahia eles fala assim ‘ô Z. C. cadê A. G.?’ ele não fala ‘a A. G. ou o João’ mas porque ele não quer / preguiça de colocar mais uma letra / mas não é porque não sabe não porque eu aprendi foi na Bahia não foi aqui / o baiano fala assim ‘cadê o João?’ não fala ‘cadê João?’ ‘cadê João?’ ‘cadê João?’ ‘cadê Zé?’ / fala assim porque é preguiça de falar mas que sabe sabe até eu falei pra menina assim ‘ô N.’ eu falei assim ‘ô N. cadê a não ô L. cadê a N.?’ que é a irmã dela né ‘cadê a N.?’ aí ela ‘ó cadê a N.’ então achou que eu estava querendo falar difícil mas não é porque eu acostumei falar ‘a N.’ ‘a A. G.’ (MarieleF)*
- D1: *uhum*
- S1: *‘o João’ ‘o José’*

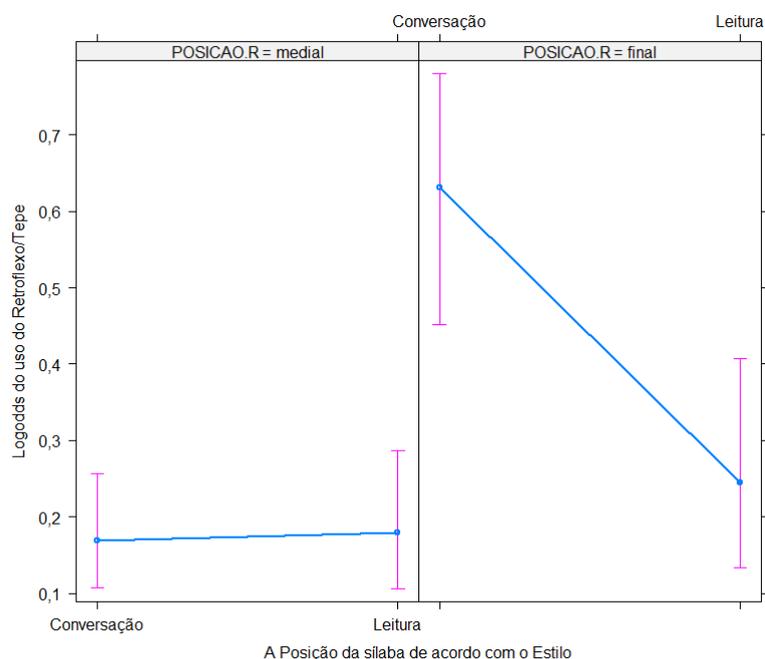
MarieleF descreve o emprego do artigo definido antes de antropônimos, um fenômeno variável que também diferencia as variedades baiana e paulista (Callou; Silva, 1997). Em seu comentário, a participante atribui a ausência do determinante a uma suposta “preguiça” que

os baianos teriam para colocar uma “letra”. Segundo MarieleF, essa regra foi aprendida na Bahia e não em São Paulo (SP), entretanto essa explicação parece uma justificativa para contrariar uma avaliação negativa que outra baiana teria feito quando ela usou o artigo antes de nomes de pessoas. É muito comum baianos não migrantes fazerem esse tipo de avaliação com baianos migrantes quando eles voltam para a Bahia, seja para passear ou para morar em definitivo. Com o comentário de MarieleF, percebe-se então que, mesmo em São Paulo (SP), os baianos avaliam de forma negativa quando outros baianos falam semelhantemente aos paulistas. Sendo assim, infere-se que os baianos que pertencem a redes mais abertas empregam o Retroflexo/Tepe por pressão da comunidade anfitriã e os que integram redes mais fechada usam essa variante em menor proporção pela pressão dos próprios baianos.

A respeito do resultado da correlação da variável resposta com a Classe morfológica, viu-se que as palavras [-gramatical] desfavorecem o uso do Retroflexo/Tepe em relação a palavras [+gramatical], cujo resultado contraria a hipótese aventada para esse preditor. Constata-se então que Retroflexo/Tepe, na fala dos migrantes baianos na RMSP, tendem a ser realizados em conjunções e em preposições em cotejo com substantivos, adjetivos, verbos, numerais, pronomes, interjeições e advérbios.

A interação do preditor Estilo com a Posição da sílaba pode ser visualizada nos gráficos da Figura 23. O resultado do gráfico da esquerda – referente à posição medial – indica que, tanto na conversação quanto na leitura, o Retroflexo/Tepe é empregado basicamente na mesma proporção. Por outro lado, no gráfico da direita – referente à posição final – esse segmento fônico é favorecido quando está mais nessa posição quando realizado na conversação em relação à leitura. Com esses resultados, aceitam-se parcialmente as hipóteses alternativas para os preditores Posição da sílaba e Estilo, de que o Retroflexo/Tepe tende a ser mais usado em posição final e de ser mais empregado em contexto de fala menos monitorada.

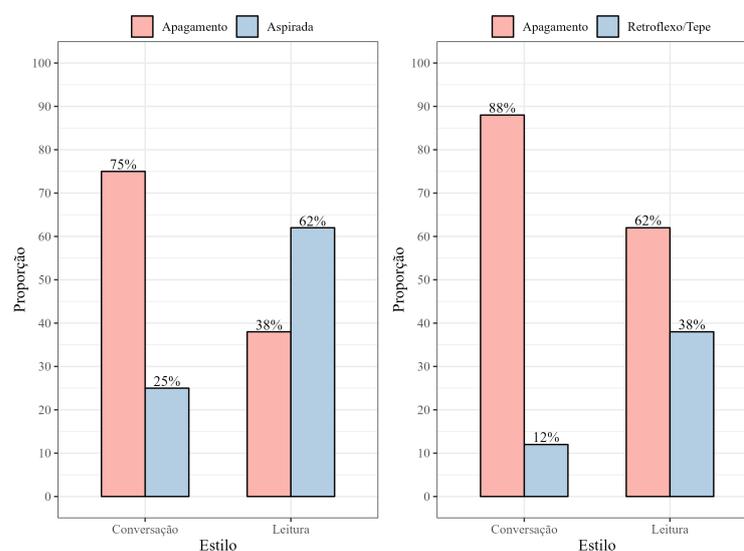
**Figura 23** – Interação das variáveis Estilo e Posição da sílaba quanto ao uso do /r/ Retroflexo/Tepe dos migrantes baianos residentes na RMSP



Fonte: Elaboração própria.

Esses resultados se assemelham aos de Oushiro (2024, no prelo[a]) e aos de Hora e Wetzels (2010). Em sua análise, Hora e Wetzels (2010) justificam que o fato de os migrantes paraibanos empregarem mais a variante nativa em relação à variante da comunidade anfitriã numa situação de fala monitorada se dá porque eles não têm total consciência da diferença entre os róticos utilizados na Paraíba e em São Paulo (SP). Essa interpretação, entretanto, não se aplica aos dados desta amostra, porque todos os participantes têm consciência da diferença da pronúncia do /r/ entre a Bahia e São Paulo. Esse resultado chama bastante atenção, porque, em termos de proporção, os dados de Retroflexo/Tepe na conversação (30%) e na leitura (27%) não são muito discrepantes. Com os resultados dessa interação, sobretudo do comportamento do rótico em cada posição da sílaba, com destaque para a posição final, é possível conjecturar que a noção de prestígio linguístico dos baianos não esteja ligada ao uso do Retroflexo/Tepe em relação à forma Aspirada, mas sim à noção de supressão e de preenchimento desse segmento fônico. Para testar essa hipótese, criaram-se dois subconjuntos de dados, o primeiro apenas com os dados de Apagamento vs. a forma Aspirada e o segundo só com os dados de Apagamento vs. de Retroflexo/Tepe, cujos valores demonstram-se na Figura 24.

**Figura 24** – Comparação de Estilo quanto ao uso do (-r) em posição de coda dos migrantes baianos residentes na RMSP



Fonte: Elaboração própria.

No primeiro gráfico (à esquerda), exibem-se as proporções de Apagamento e da forma Aspirada nos dois contextos de Estilo, cujos valores mostram que, na conversação, há 75% de Apagamento em relação a 25% de Aspirada; já na leitura, há 38% de Apagamento e 62% de Aspirada. No segundo gráfico (à direita), expõem-se os valores de Apagamento e de Retroflexo/Tepe nos mesmos contextos, sendo 88% de Apagamento e 12% de Retroflexo/Tepe na conversação e 62% de Apagamento e 38% de Retroflexo/Tepe na leitura. Esses resultados sugerem que, em ambos envelopes de variação, há redução das taxas de Apagamento, independentemente da variante utilizada para preencher a coda. Dessa maneira, assumindo que a queda do /r/ na Bahia tende a acontecer sobretudo em final de palavras, é plausível dizer que a variação de Estilo na fala dos migrantes baianos tem como referência padrão a realização do /r/ na coda em relação ao Apagamento. Ou seja, na fala mais monitorada, eles tendem a recuperar o /r/ com a variante Aspirada ou com o Retroflexo/Tepe. A elevada taxa de realização da variante Aspirada no contexto de leitura poderia ser um caso de acomodação a curto prazo (Giles; Taylor; Bourhis, 1973), já que o entrevistador também é baiano. Contudo, levando em consideração que esse mesmo padrão foi visto por Oushiro (2024, no prelo[a]), numa pesquisa cujas gravações foram feitas por campineiros, essa possibilidade logo pode ser descartada. Para esses resultados, é mais plausível assumir que, numa situação de fala monitorada, especialmente em leitura de texto e de palavras, os migrantes tendem a ativar a sua gramática nativa por uma questão de insegurança linguística. Ou seja, o fato de os participantes não saberem previamente quais palavras seriam lidas e/ou talvez por não serem palavras comuns em seu vocabulário, fez com que eles se sentissem mais seguros aplicando a variante Aspirada em vez do Retroflexo/Tepe.

#### 5.4 Síntese

Nesta Seção, explicitaram-se os resultados das análises estatísticas, realizadas no R (R Core, 2023), em que se testou a correlação do uso do (-r) em posição de coda silábica, especialmente quanto ao uso do Retroflexo/Tepe em relação às Fricativas velar/glotal, na fala de migrantes baianos residentes na RMSP. Os resultados globais dessa análise mostram que os baianos, ao migrar para essa capital, adaptam-se quanto ao uso do Retroflexo/Tepe (29%) ao manter contato com os paulistas. Essas variantes tendem a ocorrer em contexto de conversação, na fala de pessoas mais velhas, que estão há mais tempo residindo em São Paulo (SP) e integram redes mais abertas, isto é, convivem com poucos baianos. Quanto ao encaixamento linguístico, verifica-se que o Retroflexo/Tepe tende a ser mais realizado na posição final, em palavras [+gramatical], como preposição e conjunção. Quanto aos contextos fônicos, constatou-se que os migrantes de redes mais abertas utilizam Retroflexo/Tepe em todos os contextos precedentes e seguintes, enquanto os baianos integrantes de redes mais fechadas tendem a empregá-lo mais quando o /r/ é precedido por vogal anterior e posterior e seguido de pausa e consoante dorsal.

## 6

# Adaptação em nível morfosintático: a negação sentencial

Conforme visto na Introdução desta Tese, no português brasileiro, o advérbio “não” varia de posição em relação ao verbo em sentenças negativas, podendo ser realizado antes dele, como em “não vou” (Neg1); antes e depois dele, como em “não vou não” (Neg2); ou apenas posposto a ele, tal qual em “Vou não” (Neg3). Considera-se canônica a primeira forma e coloquiais a segunda e a terceira variantes (Barme, 2005; Noll, 2008), sendo estas mais recorrentes na oralidade em relação à escrita (Furtado da Cunha, 2001).

Além dessas características, a Negação sentencial (Neg) é uma variável sociolinguística cujas variantes não só distinguem o português brasileiro do português europeu mas também caracteriza diatópias do Brasil (Marroquim, 2008 [1934]; Barme, 2005; Noll, 2008; Nascimento, 2014), embora, aparentemente, não seja um fenômeno saliente socialmente comparado à pronúncia do /r/ em coda silábica. Essa dissimilaridade torna-se evidente, por exemplo, quando se comparam algumas variedades baianas com a variedade paulistana. Confrontando os dados de (Neg) na Bahia (Cavalcante, 2007; Santana; Nascimento, 2011) com a frequência de seu uso em São Paulo (SP) (Rocha, 2013), depreende-se que, em ambas as variedades, as taxas de Neg1 são maiores, porém Neg2 e Neg3 tendem a ser mais empregadas na Bahia do que em São Paulo (SP).

Esse contraste permite questionar se um baiano, ao migrar para São Paulo (SP), reduziria a frequência de realização de Neg2 e Neg3 conforme o padrão paulistano. Se sim, pode-se ainda perguntar quais preditores sociais e linguísticos estariam correlacionados com essa diminuição. Nesta Seção, mostram-se os resultados das análises multivariadas de regressão logística de efeitos mistos com a inclusão do Participante como variável aleatória as quais testaram a correlação de Neg2 e Neg3 em relação a Neg1 com os preditores sociais elencados na Seção 3 e com os preditores linguísticos Marcadores discursivos (presença; ausência); Ativação da proposição (direta; indireta); Tipo de sentença (independente; principal e subordinada) e

Manifestação do sujeito (lexicalizado; não lexicalizado).

Esta Seção está assim organizada: na Subseção 6.1, exploram-se os estudos de Furtado da Cunha (2001), Cavalcante (2007), Sousa (2007), Santana e Nascimento (2011), Rocha (2013), Goldnadel et al. (2013), Nunes (2014) e Pereira (2018) a fim de estimar, a partir de frequências relativas, uma possível isoglossa das variantes da negação sentencial do Nordeste ao Sul do Brasil, destacando especialmente os padrões da Bahia e de São Paulo; na Subseção 6.2, expõe-se a metodologia utilizada para manipular e analisar os dados de (Neg); na Subseção 6.3, descrevem-se e discutem-se os resultados das análises estatísticas realizadas na plataforma R (R Core, 2023) e; na Subseção 6.4, condensam-se os achados sobre as adaptações dos migrantes baianos quanto ao uso de (Neg).

### 6.1 A variação sintática do negador “não” em variedades do português brasileiro: Bahia e São Paulo em foco

A principal hipótese que fundamenta a alternância das estruturas de (Neg) no português brasileiro é a de que essa variável estaria passando pelo mesmo processo de gramaticalização ocorrido na negação do francês, conhecido como o *Ciclo de Jespersen*, em referência a seu postulador (Furtado da Cunha, 2001; Cavalcante, 2007; Rocha, 2013). Segundo Jespersen (1917, p. 6–7), no francês antigo, a negação era apenas pré-verbal e realizada pelo negador *ne*. Ao longo do tempo, esse advérbio teria sofrido perdas fonológicas, fazendo com que as sentenças negativas precisassem de elementos que reforçassem a negação, como *mie* (farelo/migalha), *point* (ponto), *pas* (passo). Com o enfraquecimento fonológico do *ne*, a negação nessa língua tornou-se usual na posição pós-verbal, de modo que, no francês moderno, apenas as construções [*ne* + verbo + *pas*] e [verbo + *pas*] coocorrem, como em (15).

- (15) a. Je *ne* parle *pas*. (“Eu não falo não.”)  
 b. Je parle *pas*. (“Eu falo não.”)

Como no francês, constata-se também no PB um enfraquecimento fonológico do “não” quando ele está na posição de Neg1 e na pré-verbal de Neg2, forte indício de que essa hipótese pode ser testada. Baseando-se em dados de 120 textos orais e escritos de 12 participantes que compõem o *Corpus Discurso & Gramática* – a língua falada e escrita na cidade de Natal (RN), Furtado da Cunha (2001) descreve algumas semelhanças entre o português brasileiro e o francês acerca da negação sentencial. Segundo a autora, no PB, Neg1 e o não pré-verbal de Neg2 perpassam uma erosão fonológica seguida de um desbotamento semântico (Furtado da Cunha, 2001, p. 17–18), assim como ocorreu ao longo do tempo no francês. Ao analisar a variação entre “não” e “num” em Neg1 e Neg2, Furtado da Cunha (2001, p. 18) observa que de 508 ocorrências de Neg1, 55% são de “não” e 45% de “num”; já em Neg2, das 52 sentenças analisadas, 19% são de “não” e 81% de “num”. Assim, Neg2 é mais produtiva quando o

primeiro negador é realizado como “num”.

A pesquisa sociofuncionalista de Pereira (2018) sobre a variação entre “não” e “num” no português falado de Vitória da Conquista (BA) também sustenta essa hipótese. Os dados obtidos de 24 entrevistas sociolinguísticas do *Corpus* do Português Popular e Culto de Vitória da Conquista (BA) revelam que “num” ocorre mais em Neg2 (75,7%) do que em Neg1 (51,7%); enquanto o uso do “não” em Neg3 é categórico (100%). Essa redução fonológica também foi notada por Sousa (2007), quando analisou acusticamente 918 ocorrências negativas de uma amostra de fala de 20 pessoas de Mariana (MG). Os resultados desse estudo mostram a existência de um *continuum* decrescente na duração da pronúncia do item “não”: “não” (150,4ms) – “num (130,9ms)” – “nu” (80,5ms) – “ũ” (55ms) – “N” (Sousa, 2007, p. 71). Na análise multivariada, feita no Goldvarb 2001, Sousa (2007, p. 81) observou que Neg2 é favorecida pelas formas reduzidas (P.R. .93) em comparação com a forma plena (P.R. .65).

O emprego de Neg2 e Neg3 delimita algumas áreas dialetais do Brasil. Na Tabela 14, exibem-se as taxas das variantes de (Neg) realizadas em variedades da Região Nordeste, como Feira de Santana, cujos dados foram levantados exclusivamente para este estudo, Matinha (Santana; Nascimento, 2011), Cinzento, Rio de Contas e Sapé (Cavalcante, 2007), na Bahia; Açuzinho e Itabaiana, em Sergipe (Santana, 2023); São Luís e Jamary dos Pretos, no Maranhão (Serra, 2018); Fortaleza, no Ceará (Roncarati, 1996) e Natal, no Rio Grande do Norte (Furtado da Cunha, 2001); em variedades da Região Sudeste como Pombal e Mariana, em Minas Gerais (Alkimim, 2001 *apud* Cavalcante, 2007); Vitória, no Espírito Santo (Nascimento, 2014); em São Paulo (SP) capital (Rocha, 2013) e na cidade do Rio de Janeiro (Nunes, 2014); e em variedades da Região Sul como Florianópolis, em Santa Catarina (Goldnadel et al., 2013); Curitiba, no Paraná (Goldnadel et al., 2013), e Porto Alegre, no Rio Grande do Sul (Goldnadel et al., 2013).

**Tabela 14** – Taxas de uso das variantes de (Neg) em variedades do PB

Variedade	Pesquisa	N	Neg1	Neg2	Neg3
Feira de Santana (BA)		642	79%	17%	4%
Matinha (BA)	Santana e Nascimento (2011)	541	78%	20%	2%
Cinzento (BA)	Cavalcante (2007)	2.026	66%	28%	6%
Rio de Contas (BA)					
Sapé (BA)					
Açuzinho e Itabaiana (SE)	Santana (2023)	434	79,3%	19,1%	1,6%
São Luís (MA)	Serra (2018)	770	80,5%	17,3%	2,2%
Jamary dos Pretos (MA)					
Fortaleza (CE)	Roncarati (1996)	813	77%	18%	5%
Natal (RN)	Furtado da Cunha (2001)	1.465	88,6%	10,8%	0,6%
Mariana (MG)	Alkimim (2001 <i>apud</i> Cavalcante, 2007)	2.322	77,1%	21,2%	1,7%
Pombal (MG)	Alkimim (2001 <i>apud</i> Cavalcante, 2007)	692	64,4%	31,3%	4,3%
Vitória (ES)	Nascimento (2014)	2.263	77,4%	21,1%	1,5%
São Paulo (SP)	Rocha (2013)	5.607	94%	5,8%	0,2%
Rio de Janeiro (RJ)	Nunes (2014)	843	73,1%	25,4%	1,5%
Florianópolis (SC)	Goldnadel et al. (2013)	1.065	95,6%	4,4%	–
Curitiba (PR)	Goldnadel et al. (2013)	1.408	97,4%	2,6%	–
Porto Alegre (RS)	Goldnadel et al. (2013)	1.410	99,4%	0,6%	–

Fonte: Elaboração própria.

A partir dos resultados apresentados na Tabela 14, verifica-se que Neg1 é a variante mais produtiva em todas as variedades estudadas; que Neg2 não é só mais frequente no Nordeste, como sugere Marroquim (2008 [1934]), mas também em comunidades do Sudeste, conforme as taxas de 31,3% e 21,2% encontradas em Minas Gerais Alkimim (2001 *apud* Cavalcante, 2007), 21,1% no Espírito Santo (Nascimento, 2014) e 25,4% no Rio de Janeiro (Nunes, 2014); e que Neg3 é mais recorrente no Nordeste, à exceção de Natal (RN) com 0,6% (Furtado da Cunha, 2001), e menos usual no Sudeste, exceto Pombal (MG) com 4,3% (Alkimim, 2001 *apud* Cavalcante, 2007) e muito menos comum na região Sul (Goldnadel et al., 2013).

Ainda sobre a realização de Neg2 e Neg3, constata-se que, quanto mais ao norte a comunidade está localizada, maior é a frequência relativa de seu uso. Muito provavelmente a constituição dessa isoglossa tenha a ver com a sócio-história desses lugares, cujas características remetem ao contato do português com línguas africanas durante o período da colonização,

como conjectura [Cavalcante \(2007\)](#). De acordo com [Pal e Araújo \(2015, p. 174\)](#), as línguas da família banto, cuja maioria contribuiu para a formação do PB, apresentam a negação com marcas múltiplas ou no final da frase, de maneira semelhante às línguas românicas, como o francês e o português brasileiro. Sendo assim, é possível que os fluxos migratórios de africanos escravizados do litoral pernambucano e baiano para as outras áreas do Nordeste, durante a cultivo da cana-de-açúcar e do algodão no século XVII, e para a Região Sudeste, ao longo da Corrida do Ouro em Minas Gerais, no século XVIII, e a chegada da Família Real ao Rio de Janeiro no século XIX ([Mattos e Silva, 2004](#)) tenham acelerado e descompassado o processo de variação e mudança da Negação sentencial em determinadas regiões do Brasil. O foco aqui, contudo, não é explicar como Neg2 e Neg3 foram incorporadas ao PB, mas reconhecer que há diferentes taxas de uso dessas variantes nas variedades do português brasileiro, chamando atenção para o contraste entre dialetos baianos e o paulistano.

Na Bahia, Neg2 tende a apresentar frequências de uso mais altas em variedades mais rurais em comparação com variedades mais urbanas; já Neg3 aparenta não ter essa distinção (ver a Tabela 14). Com o intuito de utilizar a mesma amostra controle baiana da análise do (-r) em coda, levantaram-se, para fins apenas exploratórios, os dados de (Neg) na cidade de Feira de Santana (BA). A partir de uma amostra de fala de 12 participantes, todos com baixo nível de escolarização (analfabeto/fundamental I ou II), estratificados em sexo (masculino; feminino) e faixa etária (I, 25-35 anos; II, 45-55 anos; III, 65+), pertencente ao Projeto A Língua Portuguesa no Semiárido Baiano ([Almeida; Carneiro, 2008](#)), verificou-se que, do total de 642 ocorrências de negação sentencial, 79% são de Neg1, 17% de Neg2 e 4% de Neg3. No distrito de Matinha (BA), localizado na zona rural de Feira de Santana (BA), a aproximadamente 18 km da área urbana, observou-se que a taxa de Neg2 é relativamente mais alta (20%) em contraste com os dados da área urbana de Feira de Santana (17%). Utilizando uma amostra de fala rural com a mesma estratificação social da amostra urbana, composta por seis participantes do banco de dados do Projeto A Língua Portuguesa no Semiárido Baiano, [Santana e Nascimento \(2011\)](#) constataram que, de 541 ocorrências de (Neg), 78% são de Neg1, 20% de Neg2 e 2% de Neg3. Com o auxílio do Goldvarb 2001, [Santana e Nascimento \(2011\)](#) testaram, através de análises multivariadas, a correlação entre as variáveis linguísticas Tipo de oração (absoluta; coordenada; principal; subordinada), Tipo de verbo (ação/movimento/processo/evento; estativo; cognitivo) e Tipo de sujeito (1ª pessoa; 2ª pessoa; 3ª pessoa) e as sociais Uso da TV (sempre; pouco/raramente), Sexo, Faixa etária e Escolaridade com o uso de Neg2 em comparação a Neg1 e de Neg3 em relação a Neg1. Os pesos relativos das rodadas mostraram que Neg2 tende a ser mais realizada em orações absolutas (P.R. .67), em contextos com verbos de ação/movimento/processo/evento (P.R. .58), com sujeito de 2ª pessoa (P.R. .94) e na fala de pessoas que usam pouco/raramente a TV (P.R. .70). Quanto às análises de Neg3, não houve nenhuma variável selecionada como estatisticamente significativa. Segundo os autores, isso ocorreu possivelmente

por causa da quantidade relativa dessa variante no *corpus*, apenas 3% de 431 ocorrências, somando os dados de Neg1 e Neg3.

Em Cinzento (BA), Rio de Contas (BA) e Sapé (BA), comunidades quilombolas do interior da Bahia com características mais rurais que Matinha (BA), notam-se proporções de Neg2 e Neg3 ainda maiores que os dados da zona urbana e da zona rural de Feira de Santana (BA). Com base em dados de uma amostra de fala do Projeto Vertentes do Português Rural da Bahia (Lucchesi, 2002), Cavalcante (2007) investigou a Negação sentencial realizada na fala de 18 participantes, todos com baixo nível de escolarização (analfabetos; semialfabetizados), estratificados em sexo (masculino; feminino) e faixa etária (I, 20-40 anos; II, 41-60 anos; III, 60+). Os resultados globais desse estudo revelam que, de 2.020 ocorrências de (Neg), Neg1 apresentou 66% de uso, Neg2 28% e Neg3 6%. Nas análises multivariadas, realizadas no Varbrul, Cavalcante (2007) testou a correlação das variáveis linguísticas Tipo de frase (respostas; não-respostas; perguntas), Tipo de oração (absoluta; principal; coordenada “livre”; coordenada introduzido por conjunção; substantiva; adverbial; causal/explicativa; relativa), Tipo de complemento verbal (realizado *in situ*; topicalizado ou anteposto; não realizado (nulo); não projetado), Tipo de realização do sujeito (realizado; realizado posposto; nulo referencial; nulo expletivo) e Tipo de realização fonética (não; num) e das variáveis sociais Sexo, Faixa etária, Escolaridade e Deslocamento da comunidade (descolamento; permanência) com o uso de Neg2 em relação a Neg1 e Neg3 em comparação com Neg1. Os resultados das análises multivariadas mostraram que Neg2 tende a ser mais usada em perguntas (P.R. .59) e respostas curtas (P.R. .68), em orações absolutas (P.R. .52), principais (P.R. .53), coordenadas ‘livres’ (P.R. .53) e coordenadas introduzidas por conjunção (P.R. .58), em contextos com complementos não realizados/nulos (P.R. .54) e não projetados (verbo intransitivo) (P.R. .56) e quando a negação é empregada sob a forma de “num” (P.R. .51), por pessoas que permanecem na comunidade (P.R. .56). Na análise de Neg3, os valores indicam que essa variante tende a ocorrer mais em perguntas (P.R. .62) e respostas curtas (P.R. .83), em orações absolutas (P.R. .89) e coordenadas ‘livres’ (P.R. .64), em contextos com sujeito nulo referencial (P.R. .66) e sujeito nulo expletivo (P.R. .66), na fala dos participantes com idade entre 20 a 40 anos (P.R. .56) e de +60 anos (P.R. .59).

Em contraste, o estudo das formas de negação no português paulistano, realizado por Rocha (2013), o único trabalho sobre a negação nessa variedade, apresentou certo distanciamento das taxas das variedades baianas, principalmente quanto ao uso de Neg2 e Neg3. Para essa pesquisa, o autor utilizou um *corpus* com 48 entrevistas sociolinguísticas realizadas pelo GESOL-USP. Diferentemente dos trabalhos realizados na Bahia, Rocha (2013) codificou seus dados a partir da ampliação do envelope de variação proposto por Schwenter (2005), isto é, considerou variantes intercambiáveis quando a proposição negada é ativada de forma direta ou indireta. Para as análises multivariadas de regressão logística feitas no Goldvarb, Rocha (2013) testou a correlação de (Neg) com as variáveis linguísticas Marcador conversacional (presença;

ausência), Outro termo negativo na sentença (presença; ausência), Ativação da proposição (direta; indireta), Adjuntos adverbiais (presente; presente deslocado à esquerda do verbo; ausente), Tipo de verbo (ligação; transitivo direto; transitivo indireto; intransitivo; impessoal/existencial), Tipo de sentença (absoluta; principal; subordinada; independente), Sujeito (lexicalizado; oculto/indeterminado; sem sujeito), Quantidade de elementos à direita do verbo, Tipo de constituinte pós-verbal (objeto direto; objeto indireto; predicativo do sujeito; cancelado) e com as variáveis sociais Sexo/Gênero (masculino; feminino), Faixa etária (18 a 35 anos; 36 a 55 anos; e +56 anos), Escolaridade (até o Ensino Médio; Ensino superior), Geração (1<sup>a</sup>; 2<sup>a</sup>) e Região da cidade (periférica; central). Os resultados globais desse estudo revelam que a taxa de uso de Neg1 é semicategórica (94%), tendo-se observado apenas 5,8% de Neg2 e 0,2% de Neg3. Os resultados das análises estatísticas inferenciais mostram que Neg2 tende a ser mais empregada em sentenças sem marcador conversacional (P.R. .53), quando no contexto não há outros termos negativos (P.R. .53) e quando a proposição é ativada diretamente (P.R. .73), na fala de paulistanos filhos de migrantes (P.R. .53), que estudaram até o Ensino Médio (P.R. .57), que têm idade entre 36 e 55 anos (P.R. .51) e +56 anos (P.R. .54) e que moram na região periférica (P.R. .52).

Cotejando as taxas da negação sentencial das variedades baianas com os valores encontrados em São Paulo (SP), vê-se uma diferença sobretudo entre o uso de Neg2 e Neg3. A partir da Tabela 14, depreende-se que, tanto na Bahia quanto em São Paulo, Neg1 é mais produtiva. Contudo chama a atenção o uso expressivo dessa variante na capital paulistana (94%) e, conseqüentemente, a baixa incidência de Neg2 (5,8%) e de Neg3 (0,2%) (Rocha, 2013). Na Bahia, principalmente as taxas de Neg2 apresentam certa diferença. Observa-se que quanto mais características rurais a comunidade tem, maior é a proporção de Neg2. Cinzento (BA), Rio de Contas (BA) e Sapé (BA), comunidades originárias de quilombos e classificadas como semi-isoladas, apresentaram 28% dela (Cavalcante, 2007). Em Matinha (BA), reconhecida como comunidade quilombola em 2016, encontraram-se 20% de Neg2 (Santana; Nascimento, 2011) e, na cidade de Feira de Santana (BA), registraram-se 17% dessa variante. O fato é que, independentemente da localização da comunidade, se rural ou urbana, o emprego de Neg2 e Neg3 é maior na Bahia do que em São Paulo (SP). Nos dados de São Paulo (SP), merece destaque o resultado da correlação de Neg2 com a ascendência dos paulistanos, segundo o qual os paulistanos filhos de migrantes tendem a empregar mais essa variante em comparação com os paulistanos filhos de paulistanos. Novamente se vê que os migrantes têm tido um importante papel na configuração linguística da cidade de São Paulo (SP), uma vez que, além de encabeçar o avanço do Retroflexo (Oushiro, 2015), parece que eles tendem a implementar o uso de Neg2 nessa comunidade (Rocha, 2013). As informações apresentadas nesta Subseção foram basilares para estabelecer os critérios, descritos na Subseção 6.2, que nortearam a manipulação e a análise dos dados de (Neg) capturados da fala dos migrantes baianos residentes na RMSP.

## 6.2 Manipulação dos dados

Os dados de (Neg) analisados neste estudo correspondem às sentenças negativas realizadas nos 20 primeiros minutos das entrevistas de 47 dos 50 migrantes baianos que compõem a amostra desta pesquisa. Conforme apresentado na Seção 3, a redução no número total de participantes em algumas análises ocorreu devido à presença de sobreposição de vozes em algumas gravações, que dificultou a identificação das ocorrências. Estabeleceu-se esse corte de tempo, porque é parte do roteiro que suscita do participante narrativas sobre sua vida, em comparação com o tempo restante, cujos questionamentos são de caráter metalinguístico; momento em que, supostamente, o participante daria mais atenção à fala.

Tais ocorrências foram codificadas com base no envelope de variação utilizado por Rocha (2013), conforme se viu na Subseção 6.1. Com base em Cavalcante (2007), não se consideraram variantes dessa mesma variável os advérbios negativos “nunca”, “ninguém”, “nada” e “nenhum” por terem sentidos mais amplos. Sendo assim, levou-se em conta apenas a posição variável do “não” em relação ao verbo, independentemente de suas realizações fonéticas, como “não”, “num”, “nu” e “ũ”. Além desses, outros critérios foram usados para refinar os dados, como a exclusão de ocorrências: (i) da negação seguida de pausa, como “Não, fiz isso”, cujo critério foi aplicado por Barthelemy (2005) e Cavalcante (2007); (ii) de expressões cristalizadas na língua como “Não sei o quê”, “Não sei você, mas...”, “Não vou dizer que...”; (iii) da leitura do texto, (iv) das orações adverbiais e perguntas-Qu, uma vez que nesses três últimos contextos não houve variação na amostra analisada e os dados foram categóricos para Neg1.

De acordo com os dados apresentados na Subseção 6.1, as taxas de Neg2 e Neg3 são menores em São Paulo (Rocha, 2013) do que na Bahia (Cavalcante, 2007; Santana; Nascimento, 2011). Por essa razão, testou-se, a partir de análises multivariadas de regressão logística de efeitos mistos, com inclusão do Participante como variável aleatória, feitas no R (R Core, 2023), a correlação de (Neg) com os preditores sociais reportados na Seção 3 e com os preditores linguísticos do Quadro 2, usando como referência a redução de Ne2/Neg3 em relação a Neg1:

**Quadro 2** – Variáveis linguísticas testadas na análise de (Neg)

Variáveis	Níveis
Marcadores conversacionais	presença ausência
Ativação da proposição	direta indireta
Tipo de sentença	oração principal oração subordinada oração independente
Manifestação do sujeito <sup>26</sup>	lexicalizado não lexicalizado

Fonte: Elaboração própria.

Tencionou-se, a partir dessas análises estatísticas, saber se os contextos de uso de Neg2/Neg3 dos migrantes baianos residentes na RMSP são os mesmos da variedade paulistana ou se mantiveram semelhantes aos dados de variedades baianas baseando-se nas hipóteses apresentadas a seguir. No estudo de Rocha (2013) e no de Santana (2023), Neg2 é mais produtiva quando não há marcadores conversacionais. Assim, conjectura-se que, na fala dos migrantes baianos residentes na RMSP, Neg2/Neg3 são menos usadas quando há presença dos marcadores, como em (16), em relação à ausência deles (17).

(16) Presença de marcadores conversacionais

- a. D1: *e aqui você não quis estudar?*  
S1: *não dava tempo né?* (EduardoS)
- b. D1: *a saudade dos filhos?*  
S1: *foi muita, né / nossa / Sofri muito mas Deus deu a vitória né? / que eu passei / né? / não foi fácil não né?* (RitaS)
- c. D1: *M.? até hoje eu lembro dela / nossa / A S. M. é maravilhosa /*  
S1: *é / foi a S.* (LetíciaC)  
D1: *e vocês?*  
S2: *tem muito assim não / sabe?* (SamiraS)

## (17) Ausência de marcadores conversacionais

- a. D1: *a sua adaptação aqui foi muito difícil / foi fácil? você teve um acolhimento de outras pessoas?*  
S1: **não** foi fácil (AlanaS)
- b. D1: *[...] tinha muita casa como antigamente?*  
S1: *não / antigamente era bem pouco / não era / não tinha asfalto não* (GeruzaA)
- c. D1: *[...] alguém assim na rua já parou pra perguntar a senhora por acaso é baiana?*  
S1: *perguntou não* (GeruzaA)

Na pesquisa de Rocha (2013) e na de Santana (2023), Neg2 é favorecida quando a ativação da proposição é direta. Assim, aventa-se a hipótese de que os baianos residentes na RMSP usam menos Neg2/Neg3 em sentenças cuja proposição é ativada de maneira indireta (19) em relação à direta (18).

## (18) Ativação direta

- a. D1: *se você tivesse oportunidade de ir pra outro bairro / qual bairro de São Paulo você iria? ou moraria né?*  
S1: *o bairro? não* tenho em mente assim o bairro / *mas eu gosto muito de sossego / eu gosto muito de interior / então se fosse pra escolher / ia pra um lugar bem interiorzão / bem sossegado* (IedaQ)
- b. D1: *seus pais eram muito rígidos quando você era criança? quando você era criança?*  
S1: *sim / se era o quê?*  
D1: *se eram rígidos*  
S1: *ah! Minha mãe era um pouco / meu pai era mais tranquilo / ele era bem tranquilão / não era rígido não / minha mãe era / minha mãe era* (JosafaF)  
D1: *se você fosse mudar de profissão / qual profissão você teria?*  
S1: *se eu fosse mudar de profissão... ah não sei não se... eu mudaria de profissão não* (CassiaC)

## (19) Ativação indireta

- a. D1: *o que você considera como ter infância?*  
S1: *ah eu acho que a gente não tinha maldade naquela época* (AlanaS)
- b. D1: *você conhece seus vizinhos?*  
S1: *eu fico menos aqui em casa que eu trabalho muito*  
D1: *uhum*  
S1 : *então ele eu já não sei se conhece todo mundo / eu não conheço todo mundo do prédio não* (SofiaB)
- c. S1: *tem que o homem ajudar em casa /eu acho que eu ajudo e eu acho legal quando os dois trabalha fora / ajudar e esse essa barreira que/ mulher tem que ficar em cozinha não* (NormandoA)

Nos exemplos em (18), a proposição negada foi ativada no discurso de maneira direta. Veja-se que, quando o participante responde, ele retoma (repete) trechos da fala do documentador. Em (18-b), por exemplo, o documentador questiona “Se fosse *mudar de profissão...*”; e CassiaC responde “eu *mudaria de profissão não*”. A ativação direta pode estar em forma de resposta curta também, uma resposta mais objetiva, deixando explícito o tópico discutido. Em (18-b), por exemplo, quando o documentador pergunta “Seus pais eram muito rígidos quando você era criança?”, JosafaF responde brevemente “Minha mãe era um pouco” / “Meu pai era mais tranquilo”. Depois ele replica a palavra “rígido” e diz “Não era rígido não.”

Opostamente, quando a proposição é ativada no discurso indiretamente, nota-se que o tópico fica subentendido, pressuposto. Há uma leve digressão até o interlocutor deixar evidente o que está sendo negado. Em (19-a), por exemplo, quando o documentador questiona “O que você considera como ter infância”, AlanaS responde não ter maldade naquela época. Nesse caso, ela reduz o período pueril a um comportamento. Igualmente, ao ser questionada se conhecia seus vizinhos, SofiaB, em (19-b), responde, justificando que ela fica pouco na casa dela e, por isso, não conhece todos os moradores do prédio.

Diferentemente dos resultados obtidos por Rocha (2013), a variável Tipo de sentença apresentou correlação com o uso de Neg2 no estudo de Santana e Nascimento (2011). Nele Neg2 é favorecida em orações absolutas. Esse resultado se aproxima do de Cavalcante (2007), cujos dados mostraram que Neg2 tende a ser mais realizada em orações absolutas, principais, coordenadas ‘livres’ e coordenadas introduzidas por conjunção. Nessa amostra, Neg3 tendeu a ser mais usada em orações absolutas e coordenadas ‘livres’. Dessa maneira, levanta-se a hipótese de que, na fala dos migrantes baianos residentes na RMSP, Neg2/Neg3 são menos recorrentes em orações principais (20) em comparação com orações subordinadas (21) e com

orações independentes<sup>27</sup>.

(20) Orações principais

- a. D1: *o pai de sua mãe?*  
S1: *isso o pai da minha mãe / mas teve uma época / tinha uma época assim olha / **não** sei se M. comentou também* (JaneteA)
- b. D1: *e você L.? como é sua relação com V.? S1: com minha irmã? quando a gente está longe / a gente se ama / quando está perto / **não** pode ficar muito perto **não** que briga* (LeticiaC)
- c. S1: *mulher tem **não** que ficar na cozinha / acabou / já era* (NormandoA)

(21) Orações subordinadas

- a. S1: *a gente vê que **não** é fácil se manter em São Paulo* (MiriamB)
- b. S1: *então **não** tenho o que falar mal deles **não*** (AndersonC)
- c. S1: *... quer mais saber de Bahia **não*** (MiraldaC)

(22) Orações independentes

- a. S1: *eu sou suspeita de falar porque eu amo meus irmão e são uns cara dez porque meu pai soube criar / eles **não** bebe* (AmandaS)
- b. D1: *[...] todas as ruas por exemplo eram asfaltadas? tinha muita casa? como era antigamente?*  
S1: *não / antigamente era bem pouco / não era... **não** tinha asfalto **não*** (GeruzaA)
- c. D1: *se você fosse mudar de profissão / qual profissão você teria?*  
S1: *se eu fosse mudar de profissão... ah não sei não se... eu mudaria de profissão **não*** (CassiaC)
- d. S1: *então todo mundo estudava na mesma escola / todo mundo brincava das mes-*

<sup>27</sup> Esse fator é a junção das orações coordenadas e absolutas, conforme o estudo de Rocha (2013).

*mas coisas / lógico que tinha um que tinha um brinquedo melhor... tal... mas não tinha muita* (AlanaS)

- e. S1: *eu acho que o que me sustenta aqui em São Paulo é as orações é... é verdade mas é as orações ... sabia? porque não é fácil não* (MirnaE)
- f. S1: *eu vim com a minha sobrinha e meu sobrinho que foram pra lá passear e eu pra cá mas eu vim pra morar não* (GeruzaA)

No estudo de Rocha (2013), o Tipo de Sujeito, independentemente de estar lexicalizado ou não, não teve relação com o uso de Neg2. No trabalho de Cavalcante (2007), essa variável linguística teve relação apenas com o uso de Neg3. Segundo os dados desse autor, Neg3 tende a ser mais realizada em orações com sujeito nulo referencial e nulo expletivo. A partir disso, conjectura-se que, na fala dos migrantes baianos, Neg2/Neg3 são menos recorrentes em orações cujo sujeito está lexicalizado (23) em relação a quando ele não está lexicalizado (24).

(23) Sujeito lexicalizado

- a. S1: *nós não tinha essa chance de brincar* (AmandaS)
- b. S1: *eu não sou tradicional não* (AndersonC)
- c. S1: *eu vou ficar aqui não* (MarioS)

(24) Sujeito não lexicalizado

- a. S1:  $\emptyset$  *não tinha nem experiência de cuidar de criança* (CassiaC)
- b. S1:  $\emptyset$  *não tinha muito brinquedo não* (GeruzaA)
- c. S1:  $\emptyset$  *vou lembrar não* (MarieleF)

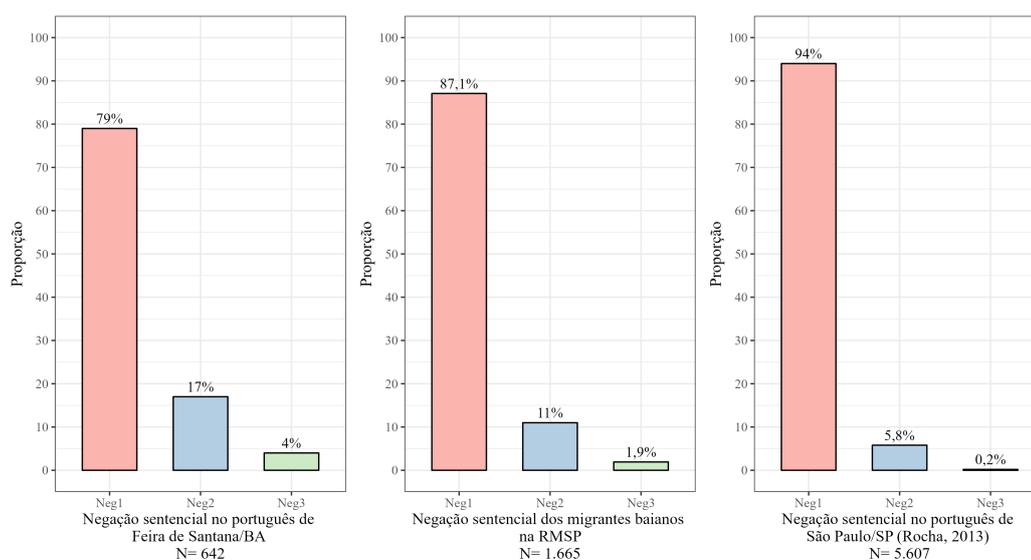
### 6.3 Resultados

Esta Subseção está dividida em duas partes. Na primeira, são explicitadas as análises quantitativas da Negação sentencial da fala dos migrantes baianos residentes na RMSP, enquanto a segunda parte trata da interpretação dos resultados obtidos.

### 6.3.1 Descrição dos achados

Assim como nas variedades baianas (Cavalcante, 2007; Santana; Nascimento, 2011) e na paulistana (Rocha, 2013), o uso da Negação sentencial na fala dos migrantes baianos na RMSP também é variável, com predominância para Neg1. Na primeiro gráfico da Figura 25, da esquerda para a direita, apresentam-se os resultados de (Neg) de baianos não migrantes a partir do *corpus* do Projeto A Língua Portuguesa no Semiárido Baiano (Almeida; Carneiro, 2008), cujos valores mostram que, do total de 642 ocorrências de negação, 79% são de Neg1, 17% de Neg2 e 4% de Neg3. No segundo gráfico, visualiza-se que, na fala dos migrantes baianos, de 1.665 ocorrências, há 87% de Neg1, 11% de Neg2 e 2% de Neg3. No último gráfico, explicitam-se os dados dessa variável sociolinguística na fala de paulistanos não migrantes, cujas taxas evidenciam que, de 5.607 realizações da negação, 94% são de Neg1, 5,8% são de Neg2 e 0,2% são de Neg3 (Rocha, 2013).

**Figura 25** – Comparação das proporções da Negação sentencial de baianos não migrantes, migrantes baianos residentes na RMSP e paulistas não migrantes



Fonte: Elaboração própria.

Comparando as proporções da Figura 25, observa-se que os dados de Neg2 e Neg3 dos baianos residentes na RMSP se aproximam dos valores encontrados na fala dos paulistanos e se distanciam dos valores da variedade feirense, de modo que Neg2 tem uma redução de 17% para 11% e Neg3 tem uma queda de 4% para 1,9% em relação a Neg1. Para estimar os contextos sociais e linguísticos em que os migrantes baianos residentes na RMSP reduzem o uso de Neg2/Neg3, fizeram-se análises multivariadas de regressão logística de efeitos mistos e testou-se a correlação de (Neg) com os preditores sociais listados na Seção 3 e os linguísticos

elencados no Quadro 2, cujos resultados apresentam-se na Tabela 15. Antes de ler os dados dessa tabela, é pertinente lembrar que não há ortogonalidade entre os preditores Idade e Tempo de residência nem entre Idade e Idade de migração; por esse motivo, realizaram-se dois modelos estatísticos de regressão logística, de maneira que, no primeiro incluíram-se os preditores Idade de migração e Tempo de residência e excluiu-se a Idade; no segundo modelo colocou-se a Idade e desprezaram-se a Idade de migração e o Tempo de residência (ver a Tabela 16<sup>28</sup>).

**Tabela 15** – Resultados do primeiro modelo de regressão logística de efeitos mistos com inclusão do Participante como variável aleatória quanto ao uso de Neg2/Neg3 na fala dos migrantes baianos residentes na RMSP

*Intercept= -4,882.*

Variável	Apl./N	%	Est.	EP	Valor-z	p
<b>Sexo/Gênero</b>						
mulheres ( <i>ref.</i> )	124/1.085	11,4%				
homens	91/580	15,7%	0,215	0,318	0,677	0,498
<b>Id. de migração</b>	–	–	0,008	0,029	0,280	0,779
<b>Autoident. paulista</b>	–	–	-0,050	0,052	-0,950	0,342
<b>T. de residência</b>	–	–	0,013	0,012	1,022	0,306
<b>Rede de baianos</b>	–	–	-0,015	0,056	-0,275	0,783
<b>Escolaridade</b>						
Fundamental ( <i>ref.</i> )	164/988	16,6%				
Médio	51/677	7,5%	-0,657	0,404	-1,628	0,103
<b>Marcadores</b>						
presença ( <i>ref.</i> )	8/188	4,2%				
ausência	207/1.477	14%	1,296	0,384	3,372	0,0007 ***
<b>Ativação</b>						
direta ( <i>ref.</i> )	32/627	5,1%				
indireta	183/1.038	17,6%	1,428	0,218	6,554	0,001 ***
<b>Sentença</b>						
independente ( <i>ref.</i> )	179/1.329	13,5%				
principal	13/255	5,1%	-0,986	0,311	-3,165	0,001 **
Continua...						

<sup>28</sup> Reportam-se na Tabela 16 apenas os coeficientes do preditor não incluído na Tabela 15 do primeiro modelo.

*Intercept= -4,882.*

Variável	Apl./N	%	Est.	EP	Valor-z	p
subordinada	23/81	28,4%	1,192	0,314	3,791	0,0001 ***
<b>Sujeito</b>						
lexicalizado ( <i>ref.</i> )	66/750	8,8%				
não lexicalizado	149/915	16,3%	0,654	0,179	3,642	0,0002 ***

Modelo1: glmer (VD SEXO + NIVEL.ESCOLARIDADE + IDADE.MIGRACAO + IDENTIDADE.PAULISTA + TEMPO.RESIDENCIA + REDE.BAIANO + MARCADORES.CONVERSACIONAIS + ATIVACAO.PROPOSICAO + TIPO.SENTENCA + SUJEITO + (1PSEUDONIMO), data: dados2)

Fonte: Elaboração própria.

**Tabela 16** – Resultados do segundo modelo de regressão logística de efeitos mistos com inclusão do Participante como variável aleatória quanto ao uso de Neg2/Neg3 na fala dos migrantes baianos residentes na RMSP

*Intercept= -4,980.*

Variável	Est.	EP	Valor-z	p
<b>Idade</b> ( <i>contínua</i> )	0,012	0,012	1,009	0,312

Modelo2: glmer (VD SEXO + NIVEL.ESCOLARIDADE + IDADE + IDENTIDADE.PAULISTA + REDE.BAIANO + MARCADORES.CONVERSACIONAIS + ATIVACAO.PROPOSICAO + TIPO.SENTENCA + SUJEITO + (1PSEUDONIMO), data: dados2)

Fonte: Elaboração própria.

A partir dos resultados das Tabelas 15 e 16, verifica-se que não há correlação de (Neg) com nenhum preditor social. Por outro lado, segundo os coeficientes da Tabela 15, todos os preditores linguísticos se correlacionaram com (Neg). Em relação ao preditor Marcadores conversacionais, verifica-se que a ausência deles favorece (1,296 *logodds*,  $p = 0,0007$ ) o uso de Neg2/Neg3 em relação ao *intercept*. Como o objetivo desta análise é observar a redução dessas variantes, nota-se então que, na fala dos migrantes baianos na RMSP, o uso de Neg2/Neg3 tende a ser reduzido quando há marcadores conversacionais nas sentenças, corroborando assim a hipótese alternativa aventada para esse preditor. Acerca do preditor Ativação da proposição, constata-se que a ativação indireta favorece (1,428 *logodds*,  $p = 0,001$ ) o uso de Neg2/Neg3 em comparação com o *intercept*, logo o uso dessas variantes, na fala dos migrantes baianos na RMSP, tende a ser reduzido quando a ativação é direta, cujo resultado contraria a expectativa para esse preditor. A respeito do preditor Tipo de sentença, observa-se que Neg2/Neg3 são favorecidas pelas orações subordinadas (1,192 *logodds*,  $p = 0,0001$ ) em relação ao *intercept* e desfavorecidas pelas orações principais (-0,986 *logodds*,  $p = 0,001$ ), em comparação com o

*intercept*. Sendo assim, percebe-se que a redução de Neg2/Neg3, na fala dos migrantes baianos na RMSP, tende a acontecer mais em orações principais, respaldando, portanto, a hipótese alternativa levantada para esse preditor. Quanto ao preditor Manifestação do sujeito, nota-se que as sentenças com sujeito não lexicalizado favorecem (0,654 *logodds*,  $p = 0,0002$ ) o uso de Neg2/Neg3 em relação ao *intercept*. Logo a redução de Neg2/Neg3 tende a acontecer em sentenças cujo sujeito está lexicalizado, consoante a hipótese aventada para esse preditor.

### 6.3.2 Discussão dos achados

A negação sentencial no português brasileiro tem apresentado diferentes padrões de uso quanto à realização de suas variantes. De acordo com os dados apresentados na Subseção 6.1, viu-se que as proporções de uso de Neg1, Neg2 e Neg3 distinguem algumas áreas dialetais brasileiras bem como diferenciam variedades rurais de urbanas. No padrão de comunidades mais ao norte do país, constatou-se que essa variação está mais acelerada em comparação com algumas comunidades mais ao sul. Ao contrastar algumas variedades baianas com a variedade paulistana, por exemplo, verificou-se que Neg2 e Neg3 são mais produtivas naquelas comunidades em relação a esta. Com base nessa discrepância, questionou-se se os baianos, ao migrarem para a RMSP, tenderiam a reduzir o uso de Neg2 e Neg3 sob influência do padrão paulistano.

Segundo a comparação de dados apresentada na Figura 25, essa população tende a reduzir o uso dessas variantes, uma vez que os valores que estimam as taxas de Neg2 e Neg3 de antes de os baianos migrarem para a RMSP caíram de 17% e 4%, respectivamente, para 11% e 1,9% nesta mesma ordem. Os dados de Santana (2023) apresentados na Seção 2 mostram que, em relação à amostra controle formada por sergipanos não migrantes, a taxa de Neg2 da fala de migrantes sergipanos residentes na Grande São Paulo também diminuiu, saindo de 19% para 12,71%. Essa diminuição fez elevar apenas a frequência de uso de Neg1, uma vez que o seu valor de aplicação subiu de 79% para 85,79% e que a taxa de realização de Neg3 praticamente se manteve a mesma, de 1,6% para 1,5%. Nos resultados encontrados por Oushiro (2020b) nas análises feitas a partir da amostra da fala de alagoanos e paraibanos residentes em Campinas (SP), detectou-se o aumento de Neg1, cujo valor saltou de 82% (amostra controle) para 87% e, conseqüentemente, a diminuição das taxas de Neg2 e Neg3.

Os resultados globais dessas pesquisas evidenciam que o processo de adaptação linguística dos migrantes para a Negação sentencial não é o mesmo identificado nas análises das variáveis lexicais nem nas análises do (-r) em coda. Numa descrição à primeira vista, poderia-se dizer que o uso de Neg1 é potencializado, porque as taxas dessa variante tendem a aumentar. Contudo, se se considerar que Neg1 apresenta taxas mais elevadas em todas as variedades observadas e que a diferença das taxas de Neg2 e Neg3 é o que distingue algumas variedades, pode-se inferir que o efeito dessa adaptação está na *contenção* destas duas variantes.

Dito de outra forma, Neg2 e Neg3, que estão em processos mais acelerados de

variação em variedades baianas, tendem a ficar contidas quando os baianos migram para a RMSP. Ao que parece, o processo de contenção não está diretamente ligado à saliência social das variantes, já que não houve nenhum comentário metalinguístico sobre as diferenças dialetais das formas de negação e também os modelos estatísticos não mostraram nenhuma correlação com os preditores sociais.

Entretanto se se considerar que muitas vezes o participante não sabe explicar os fenômenos linguísticos com termos técnicos, pode-se conjecturar que, para alguns casos de variação, como o da Negação sentencial, o migrante identifique diferenças dialetais na forma como são pronunciadas algumas sentenças. Para muitos migrantes, em linhas gerais, o “sotaque” baiano diz respeito a um conjunto de pronúncias linguísticas “erradas” e a do paulista a um conjunto de pronúncias “corretas”, como comenta OrlandoM em (25).

- (25) D1: *como é que o baiano fala então?*  
 S1: *é... baiano fala palavra va/ arrastado / negócio assim fala erra/ meio errado (OrlandoM)*  
 D1: *e como é que o paulista fala?*  
 S1: *o paulista ele fala mais diferente né / ele sabe falar mais coisa **completamente** né / não não fala tanto errado igual o baiano*

No Exemplo (25), foi questionado a OrlandoM como os baianos e os paulistas falam. Para esse participante, os baianos falam “errado” em relação aos paulistas, porque falam de maneira “incompleta”. Possivelmente o participante estava aludindo aos processos fonológicos de supressão que são comuns na fala, como a síncope e a aférese que ocorrem na negação sentencial quando há a realização de Neg2 (“não” – “num” – “nu” – “ũ”) (Furtado da Cunha, 2001; Cavalcante, 2007; Sousa, 2007; Pereira, 2018) e em outros contextos linguísticos. Dessa maneira, por mais que os migrantes baianos não tenham comentado sobre a variação dialetal de (Neg), é possível que a saliência social dessa variável esteja vinculada à forma como ela é pronunciada. Sendo isso verdade, é provável que a redução de Neg2/Neg3 na fala dos migrantes baianos esteja associada à pronúncia suprimida do negador que remete ao falar baiano. Para trabalhos futuros, é relevante analisar o uso de (Neg) considerando as formas como o “não” é pronunciado.

Em relação aos preditores testados nas análises estatísticas, viu-se que apenas os preditores linguísticos se correlacionaram com (Neg). O preditor Marcadores conversacionais demonstrou ter diferença significativa, de modo que a redução de Neg2/Neg3 tende a acontecer mais em sentenças em que há a presença deles. Esse resultado também foi observado por Santana (2023) no estudo da fala de migrantes sergipanos em São Paulo (SP). Ambos os resultados, desse modo, seguem o padrão da comunidade anfitriã, conforme os dados de Rocha (2013). Cavalcante (2007) e Santana e Nascimento (2011) não testaram essa variável em suas pesquisas,

por isso não se pôde fazer essa comparação com as variedades baianas. Sendo assim, os migrantes baianos residentes na RMSP, em princípio, tendem a usar as mesmas regras de aplicação de Neg2/Neg3 de acordo com a presença e a ausência de marcadores conversacionais.

A respeito do preditor Ativação da proposição, constatou-se que os migrantes baianos residentes na RMSP tendem a reduzir a aplicação de Neg2/Neg3 em contextos cuja proposição tenha sido ativada de forma direta. Esse resultado chama bastante atenção, porque a expectativa para esse preditor era o contrário, conforme se viu nos resultados de Rocha (2013). Nos resultados de Santana (2023, p. 116), nota-se que o padrão de uso de Neg2/Neg3 nesse contexto linguístico foi semelhante ao encontrado por Rocha (2013). Contudo cabe destacar que, em relação à amostra controle, a redução de Neg2/Neg3 na fala dos migrantes sergipanos em São Paulo (SP) é maior quando a ativação é direta (de 30% para 21%) em comparação com a indireta (de 12% para 10%). Esses dados permitem inferir que a redução das formas de Neg2/Neg3 tendem a ocorrer em sentenças negativas cuja proposição foi ativada diretamente e que, embora estejam na mesma comunidade anfitriã, o processo de contenção está mais acelerado na fala dos migrantes baianos em relação aos migrantes sergipanos.

Quanto ao Tipo de sentença, constatou-se que a redução de Neg2/Neg3 tende a acontecer em orações principais. Nos trabalhos de Cavalcante (2007) e Santana e Nascimento (2011), observou-se que esse contexto linguístico não favoreceu o emprego de Neg2 e de Neg3 (especificamente nos dados de Cavalcante (2007)). Na pesquisa de Rocha (2013, p. 74), esse preditor linguístico não teve correlação com a aplicação de Neg2 mas, com base nos resultados de rodada descartada, observa-se que, em termos de proporção, essa variante é menos recorrente em orações principais (10%) em comparação com as orações subordinadas (11%) e com as orações independentes (15%). Logo se vê que o padrão da comunidade de origem tende a ser mantido, por ser o mesmo utilizado na comunidade anfitriã.

Acerca da Manifestação do sujeito, depreendeu-se que os migrantes baianos tendem a reduzir o emprego de Neg2/Neg3 em sentenças cujo sujeito está lexicalizado. Esse padrão tem sido encontrado em variedades baianas especialmente para o uso de Neg3, conforme os dados de Cavalcante (2007). Nos dados do português paulistano, esse preditor não teve relação com o uso de Neg2 (Rocha, 2013). Dessa forma, observa-se que aparentemente os migrantes baianos residentes na RMSP tendem a manter a regra de uso de Neg2/Neg3 nesse contexto linguístico. Santana (2023) não testou a correlação de (Neg) com o Tipo de sentença nem com a Manifestação do sujeito, portanto não foi possível fazer comparações desses resultados com a amostra de sergipanos em São Paulo (SP).

#### 6.4 Síntese

Os resultados gerais dessas análises mostram que os migrantes baianos residentes na RMSP tendem a reduzir as taxas de Neg2/Neg3 seguindo o padrão paulistano. Essa redução parece

---

não ter relação com a saliência social de Neg2 e Neg3 uma vez que os coeficientes de regressão indicaram não haver correlação com nenhum preditor social. Por outro lado, a redução de Neg2/Neg3 teve relação com todos os preditores linguísticos, de maneira que a regra mais afetada é a Ativação da proposição.

## 7

## A plasticidade dialetal

Na presente Seção, sistematizam-se os resultados das análises feitas com base na amostra de fala dos baianos residentes na RMSP e apresentam-se generalizações acerca das adaptações linguísticas na fala de migrantes em situação de contato dialetal. Esta Seção está organizada da seguinte forma: na Subseção 7.1, descrevem-se os processos de adaptação linguística identificados na fala dos baianos residentes na RMSP; na Subseção 7.2, discorre-se sobre a relação das adaptações linguísticas dos migrantes baianos com o contexto social e estilístico nos quais eles estão inseridos; na Subseção 7.3, listam-se outros fenômenos variáveis que diferenciam o falar baiano do falar paulista percebidos por este pesquisador durante a sua estadia em São Paulo.

### 7.1 Processos de adaptação linguística de migrantes em situação de contato dialetal

Os migrantes em situação de contato dialetal, conforme demonstrado nas Seções 4, 5 e 6, tendem a se adaptar à variedade linguística da comunidade anfitriã por meio de diferentes processos. Tais adaptações não estão vinculadas somente à natureza da variável, mas ao descompasso de uso das variantes dos dialetos em contato devido aos casos de regularidade ou irregularidade linguística. Considera-se um caso de regularidade linguística entre os dialetos em contato quando eles compartilham de uma mesma variável sociolinguística, mas se diferenciam quanto às taxas de uso de suas variantes. Como se viu na Seção 6, a Negação sentencial, no português brasileiro, apresenta variantes cujas proporções, por exemplo, diferenciam a variedade baiana (17% de Neg2 e 4% de Neg3) da variedade paulistana (5,8% de Neg2 e 0,2% de Neg3).

No que diz respeito aos casos de irregularidade, como visto nas seções mencionadas, os dialetos em contato podem divergir quanto: (i) à presença de determinada variante no dialeto da comunidade anfitriã e a sua ausência na variedade nativa do migrante como, por exemplo, o Retroflexo, que é um estereótipo do falar paulistano e é totalmente incaracterístico da variedade baiana (ver a Seção 5); (ii) à versatilidade de contextos de uso de certa variante na comunidade anfitriã e a sua restrição no dialeto nativo do migrante, como, por exemplo, o Tepe, que tende a ser empregado em variados contextos no falar paulistano, mas, na variedade baiana, ele tende

a se restringir ao /r/ em *onset* (ver a Seção 5); (iii) à diferença de significados de significantes partilhados. Conforme se descreveu na Seção 4, na Bahia, emprega-se o item “Mandioca” para nomear a raiz tóxica utilizada na produção de farinha ou de ração para animais; já em São Paulo, esse mesmo significante é usado como lexia que nomeia a raiz que se cozinha para comer ou que serve para fazer bolos.

Tendo em mente essas ponderações, classificam-se os processos de adaptação linguística encontrados na fala dos migrantes baianos na RMSP. A categorização sugerida toma como referência a perspectiva da comunidade anfitriã, i.e., analisam-se as adaptações dos migrantes quanto ao uso das variantes típicas da nova comunidade e não em relação ao emprego das variantes típicas de sua gramática nativa. A proposta taxonômica, no entanto, pode ser usada sob essa outra perspectiva, uma vez que os processos de adaptação linguística estão interligados. Aliás, é oportuno destacar que a análise da fala de migrantes na perspectiva da comunidade anfitriã parece ser um caminho menos complexo de percorrer, principalmente se os participantes da amostra tiverem residido em mais de dois lugares. Essa questão metodológica, contudo, é tópico para discussões futuras.

Na amostra da fala dos migrantes baianos na RMSP, foram identificados os processos de *aquisição*, *ampliação de contexto*, *ampliação semântica*, *potencialização* e *contenção*. Isso não quer dizer, é claro, que eles são os únicos tipos de adaptação que podem ocorrer na fala de migrantes. Deve haver muitos outros que ainda carecem de ser descritos.

Ao verificar que os migrantes baianos apresentaram 6,7% de uso do /r/ Retroflexo em posição de coda, cuja variante não é típica do falar baiano, assume-se que eles a adquiriram ao manterem contato com os paulistanos. Com base nesse caso de variação, classifica-se como *aquisição* o processo de adaptação em que os migrantes adquirem variantes da comunidade anfitriã que são totalmente atípicas em sua gramática nativa.

Esse processo de adaptação, porém, não se aplica ao uso do Tepe, porque, no dialeto baiano, esse fone é realizado em posição de coda quando o /r/ é seguido de vogal, conforme verificado na pesquisa de Oliveira (1999), revisada na Seção 5. Em casos como esse, portanto, ocorre o processo de *ampliação de contexto*, isto é, os migrantes expandem os contextos de uso de variantes que têm contextos restritos em sua gramática nativa. O processo de *ampliação*, todavia, não se limita apenas a contextos de uso. Na Seção 4, foi observado que as lexias “Mandioca” e “Marmita” são usadas nas variedades baiana e paulista com significados distintos. Muito provavelmente por causa do contato com os paulistas, os migrantes baianos incorporaram a essas palavras os sentidos utilizados na comunidade anfitriã. Para esse tipo de adaptação, verifica-se então que houve a *ampliação semântica*.

Na Seção 4, mostrou-se que os migrantes baianos aumentaram as proporções de uso das variantes “Mexerica” e “Serviço”. Conforme se observou na Seção 2, apenas a título de exemplo, o aumento também aconteceu no emprego do Imperativo com morfologia de in-

dicativo por essa população migrante em São Paulo (Souza, 2019); na Realização do artigo diante de possessivos por migrantes alagoanos e paraibanos em Campinas (SP) (Guedes, 2019 [2017]); na Palatalização de /t, d/ diante de [i] na fala de alagoanos e paraibanos residentes em Campinas (SP) (Oushiro, 2020c) e em São Paulo (Santana, 2023). Sendo assim, partindo desses resultados, denomina-se *potencialização* o processo de adaptação em que as taxas das variantes de menor uso na gramática nativa dos migrantes são elevadas devido a sua alta frequência na comunidade anfitriã.

Na Seção 6, demonstrou-se que as taxas de uso de Neg2 e Neg3 são maiores na Bahia e menores em São Paulo (SP) e evidenciou-se que, na fala dos migrantes baianos residentes na RMSP, as taxas dessas variantes foram menores em comparação com os valores observados na Bahia. Esse mesmo resultado foi constatado nas análises da Negação sentencial dos migrantes alagoanos e paraibanos em Campinas (SP) (Oushiro, 2020c) e dos sergipanos residentes em São Paulo (SP) (Santana, 2023). Também se notou a diminuição na Palatalização do /s/ em coda de migrantes cariocas residentes em João Pessoa (PB) (Possatti, 2020). Com base nesse padrão, classifica-se como *contenção* o processo de adaptação linguística no qual os migrantes diminuem as taxas das variantes de maior uso em sua gramática nativa por influência das baixas taxas de seu emprego na comunidade anfitriã.

A dificuldade para identificar o processo de adaptação linguística do item Marmitex, devido à ausência de parâmetros de dados da comunidade de origem e da comunidade anfitriã, permite advogar pelo uso do método comparativo nos estudos sobre o contato entre dialetos mutuamente inteligíveis, como o empregado nas pesquisas de Adant (1989), Guedes (2014), Guirelli (2018), Oushiro (2020c) e Santana (2023) (apenas para citar alguns) e nas análises explicitadas nas Seções 4, 5 e 6.

Os dados encontrados na amostra de fala de baianos residentes na RMSP, em comparação com outros estudos sobre a fala de migrantes, revelaram que o encaixamento social das adaptações linguísticas na fala de migrantes está estritamente ligado aos aspectos (sociais e linguísticos) intrínsecos das comunidades em contato, conforme se verificou na Subseção 7.2

## **7.2 As adaptações linguísticas de migrantes em situação de contato dialetal no contexto social e estilístico**

Em uma análise sobre as adaptações ocorridas na fala de migrantes oriundos de uma mesma área geográfica e que residem numa mesma comunidade, espera-se que as variantes para as quais eles se adaptam, independentemente do nível da língua a que pertencem, apresentem padrões de encaixamento social semelhantes. No entanto isso não é o que tem sido observado. Segundo os dados de Oushiro (2020c), revisados na Seção 2, os padrões sociais das variáveis fonético-fonológicas e morfossintáticas dos migrantes alagoanos e paraibanos residentes em Campinas (SP) não foram os mesmos. As variáveis sociolinguísticas de mesma natureza lin-

guística também não mostraram os mesmos padrões sociais. Os resultados de Oushiro (2020c) demonstraram que as vogais médias pretônicas /e/ e /o/ se correlacionaram apenas apenas com a Idade de migração; a variável /t, d/ diante de [i] se correlacionou com a Faixa etária e a Idade de migração; enquanto o (-r) em coda se correlacionou com Escolaridade, Motivo da migração, Idade de migração, Tempo em SP e Proporção de vida em SP. Na amostra dos baianos residentes na RMSP, conforme os dados apresentados no Quadro 3, apenas as variáveis sociolinguísticas de nível lexical e fonético-fonológico se correlacionaram (células pintadas de cinza) com os preditores sociais, ao contrário da variável morfossintática. Além disso, os padrões das variáveis que se correlacionaram se mostraram diferentes, exceto para Tempo de residência.

Essa informação é bastante auspiciosa para os estudos de contato entre dialetos mutuamente inteligíveis, porque, mesmo não havendo correlação com os principais preditores sociais examinados nesse tipo de pesquisa, como foi o caso de (Neg) na amostra de fala aqui investigada, a adaptação linguística para as regras da comunidade anfitriã pode acontecer. Ao que tudo indica, as pressões sociais das quais resultam as adaptações linguísticas na fala de migrantes têm relação com a saliência social das variantes, conforme já sinalizaram Trudgill (1986), Bortoni-Ricardo (2011 [1985]), Barbosa (2022) e Oushiro (2024, no prelo[a]). Veja-se, por exemplo, que, na amostra de fala dos baianos residentes na RMSP, só se correlacionaram com os preditores sociais os Itens lexicais e o (-r) em coda, que também receberam comentários metalinguísticos sobre as diferenças dialetais de suas variantes. Um ponto a se considerar, porém, é que a saliência social das variantes precisa ser examinada caso a caso, ponderando, sobretudo, a sua relação com as comunidades em contato. No estudo de Hora e Wetzels (2010) (ver a Seção 2), os autores reportaram que os paraibanos residentes em São Paulo (SP) não tinham consciência das diferenças dialetais entre a Paraíba e São Paulo para o uso do /r/ em coda. Na amostra aqui analisada, entretanto, viu-se que todos os seus participantes têm consciência da diferença no emprego dessa variável sociolinguística entre Bahia e São Paulo.

**Quadro 3** – Comparação das correlações das variáveis sociolinguísticas Lexias, (-r) em coda e (Neg) identificadas na fala dos migrantes baianos na RMSP com os preditores sociais

Variáveis sociais	Variáveis sociolinguísticas		
	Lexias	(-r)	(Neg)
Sexo/Gênero			
Idade			
Nível de escolaridade			
Tempo de residência			
Idade de migração			
Rede de contatos com baianos			
Grau de autoidentificação paulista declarada			

Fonte: Elaboração própria.

Ao comparar os resultados das análises do (-r) em coda dos migrantes baianos na RMSP com os dados obtidos por [Oliveira \(2020\)](#) na análise da fala dos baianos residentes em Bauru quanto a essa mesma variável resposta (ver a Seção 5), constata-se que migrantes da mesma origem residentes em diferentes pontos de uma mesma região, como interior e capital, pode apresentar padrões sociais distintos quanto ao uso de uma mesma variável sociolinguística. Os padrões sociais da realização do (-r) nessas duas amostras de baianos, em certa medida, distanciam-se (ver a Seção 2). Com base nas informações do Quadro 4, verifica-se que as correlações de (-r) com os preditores sociais em ambas as amostras se assemelham quanto aos preditores Sexo/Gênero, Idade, Tempo de residência e Rede de contato com baianos e se diferenciam em relação ao Nível de escolaridade e à Idade de migração. O preditor Grau de autoidentificação paulista declarada não foi testado (N.T.) por [Oliveira \(2020\)](#) nem o preditor Atitudes em relação à comunidade anfitriã foi testado aqui neste estudo, logo não foi possível fazer a comparação entre eles.

**Quadro 4** – Comparação das correlações do uso do (-r) em coda na fala dos migrantes baianos residentes em Bauru (SP) e na RMSP com os preditores sociais

Preditores sociais	Comunidade anfitriã	
	Bauru (SP)	RMSP
Sexo/Gênero		
Idade		
Nível de escolaridade		
Tempo de residência		
Idade de migração		
Rede de contato com baianos		
Grau de autoidentificação paulista declarada	N.T.	
Atitudes em relação à comunidade anfitriã		N.T.

Fonte: Elaboração própria com base nos dados de [Oliveira \(2020\)](#).

Contudo, ao contrastar as tendências sociais das adaptações linguísticas dos baianos residentes no interior (Bauru (SP)) com as dos baianos residentes na RMSP, nota-se que essa população migrante, em ambas as comunidades anfitriãs, apresentou a mesma tendência para os preditores Sexo/Gênero e Tempo de residência, mas divergiu em relação aos preditores Idade, Rede de contato, Nível de escolaridade e Idade de migração. Em Bauru (SP), os migrantes baianos de redes fechadas, com idade entre 38 e 44 anos, que migraram com idade entre 14 e 26 anos, que estudaram até o Ensino Médio, que residem há mais de 20 anos tendem a empregar mais o Retroflexo. Por outro lado, na RMSP, observou-se que os mais velhos, que integram redes abertas (em todos os contextos fonológicos precedentes e seguintes) e que estão há mais tempo em São Paulo tendem a usar mais as variantes típicas da comunidade anfitriã.

Focalizando apenas nos resultados do preditor Rede de contato com baianos das duas amostras, é plausível inferir que a diferença detectada entre eles seja um efeito da maneira como a população anfitriã age em relação aos migrantes. Os bauruenses possivelmente conscientes de que aqueles provenientes do interior de São Paulo enfrentam preconceitos linguísticos na capital paulista, quiçá sejam mais solidários com os migrantes baianos. Assim, a adaptação deles tenha se estabelecido nas redes fechadas devido a atitudes positivas em relação à população da comunidade anfitriã, como sugere a correlação do (-r) com a variável Atitudes. Na capital, por outro lado, a ausência de solidariedade dos paulistanos para com os baianos, conforme se viu nos exemplos da Seção 5, possivelmente tenha feito com que apenas os migrantes de redes mais frouxas tenham se adaptado ao Retroflexo/Tepe não por questões atitudinais, mas por terem sido alvo de preconceito.

Ao contrapor os resultados de (Neg) da amostra aqui investigada com os dados de Santana (2023) sobre o uso dessa mesma variável sociolinguística na fala de sergipanos residentes na capital paulista (ver a Seção 6), verifica-se que migrantes de origens distintas, ainda que de áreas geográficas próximas, e que residem numa mesma comunidade anfitriã podem apresentar padrões diferentes para o uso de uma mesma variável sociolinguística. No Quadro 5, contrapõem-se os padrões sociais de (Neg) encontrados nessas duas amostras. Nesse cotejo, observa-se que, dos preditores testados em comum, apenas a Rede divergiu entre os resultados de suas análises. Em Santana (2023), viu-se que os migrantes sergipanos integrantes de redes abertas tendem a diminuir o uso de Neg2/Neg3, seguindo o padrão paulistano. Em contrapartida, não se verificou diferença significativa entre os migrantes baianos na RMSP de redes fechadas e abertas quanto à redução de Neg2/Neg3.

A diferença entre os padrões sociais dos migrantes baianos e dos migrantes sergipanos pode estar relacionada às regras de uso dessas variantes nas comunidades de origem. Por exemplo, ao confrontar as taxas de (Neg) da amostra de Sergipe com a de Feira de Santana (BA), que possuem características sociais mais afins em relação à amostra de Matinha (BA) (Santana; Nascimento, 2011) e de Cinzento (BA), Rio de Contas (BA) e Sapé (BA) (Cavalcante, 2007), percebe-se que Neg1 tem a mesma proporção em ambas as comunidades (79%). Em compensação, se se assumir como verdadeira a hipótese de Jespersen (1917) (ver a Seção 6), nota-se que a variação entre Neg2 e Neg3 está mais acelerada na Bahia (17% e 4%, respectivamente) do que em Sergipe (19% e 1,6%, nesta mesma ordem) (ver a Seção 6).

**Quadro 5** – Comparação das correlações do uso de (Neg) na fala dos migrantes baianos e sergipanos residentes na RMSP com os preditores sociais

Preditores sociais	Comunidade de origem	
	BA	SE
Sexo/Gênero		
Idade		N.T.
Nível de escolaridade		
Tempo de residência		
Idade de migração		
Rede		
Grau de autoidentificação paulista declarada		N.T.
Índice de integração	N.T.	
Identificação com São Paulo	N.T.	

Fonte: Elaboração própria com base nos dados de Santana (2023).

Com base nos dados apresentados na Seção 4, descobriu-se que as adaptações linguísticas podem obliterar ao longo do tempo. À exceção do item Mandioca, cujos dados se distribuíram de maneira ascendente entre as Faixas do Tempo de residência em São Paulo (Faixa I, 17%; Faixa II, 21%; Faixa III, 24%), os itens Mexerica (Faixa I, 12%; Faixa II, 26%; Faixa III, 22%), Marmitex (Faixa I, 7%; Faixa II, 23%; Faixa III, 20%) e Serviço (Faixa I, 4%; Faixa II, 11%; Faixa III, 6%) apresentaram uma relativa queda nas taxas de uso entre a Faixa II e III. Um declínio semelhante a esse foi visto no estudo de Souza (2019) sobre o Imperativo com morfologia de indicativo (ver a Seção 2), cujo resultado mostrou que os migrantes baianos que residem entre 0-7 anos em São Paulo tenderam a usar mais a variante da comunidade anfitriã, em comparação com quem está entre 8-15 anos e +16 anos. Presumivelmente, as variantes típicas da comunidade anfitriã para as quais os migrantes tenham se adaptado tendam a obliterar por efeito do relaxamento linguístico, conforme conjecturou Souza (2019). Nesse caso, em específico, por consequência da redução da pressão do mercado de trabalho, visto que a obliteração aconteceu entre os participantes mais velhos, já aposentados, que estão há mais de 32 anos residindo na RMSP.

A obliteração das variantes típicas da comunidade anfitriã na fala de migrantes parece se efetivar nos casos de migração de retorno, exceto em contextos de fala monitorada. Embora não haja dados de baianos regressos da capital paulista com os quais se possa comparar com os resultados desta amostra, a pesquisa de Oliveira (2023, inédito) sobre o retorno de alagoanos de São Paulo para a sua comunidade nativa (ver a Seção 2) permite estimar isso. Em sua análise, Oliveira (2023, inédito) identificou diferença significativa entre os níveis da variável Estilo (entrevista, leitura) quanto à realização das variantes típicas de São Paulo, de maneira que elas, com exceção para (-r), foram favorecidas em contextos de fala [+monitorada]. Em outras palavras, em dados de conversação, os regressos tendem a empregar mais variantes de sua gramática nativa.

Na Seção 5, na análise do uso de (-r) em coda em contexto de Estilo, verificou-se que, conforme já havia sido destacado por Oushiro (2024, no prelo[a]), os migrantes tendem a empregar mais as variantes de sua gramática nativa em contextos de fala [+monitorada] em comparação à fala [-monitorada]. Hipoteticamente, ao confrontar os resultados das análises descritivas nas quais se considerou a variação Apagamento vs. Aspirada e Apagamento vs. Retroflexo/Tepe (ver a Seção 5), parece que os migrantes baianos residentes na RMSP têm como prestígio a forma preenchida, independentemente da variante empregada, seja a típica da comunidade de origem ou da comunidade anfitriã. No entanto o fato é que, recursivamente, os migrantes, em situação de contato dialetal, têm usado variantes de sua gramática nativa em contextos de fala monitorada, salvo, com base em Oliveira (2023, inédito), para o uso de (-r) em coda em contexto de migração de regresso, conforme se observa no Quadro 6.

**Quadro 6** – Comparação dos padrões de Estilo (modelo laboviano) de migrantes em situação de contato dialetal quanto ao emprego de variantes típicas da comunidade de origem (VTCO), da comunidade anfitriã (VTCA) e da comunidade de onde os migrantes regressaram (VTR)

Pesquisa	Migrantes	Com. anf.	Variável	[-monitorada]	[+monitorada]
	BA	São Paulo (SP)	(-r) em coda	VTCA	VTCO
Oliveira (2020)	BA	Bauru (SP)	(-r) em coda	VTCA	VTCO
Hora e Wetzels (2010)	PB	São Paulo (SP)	(-r) em coda	VTCA	VTCO
Oushiro (2024, no prelo[c])	AL/PB	Campinas (SP)	(-r) em coda	VTCA	VTCO
Oushiro (2024, no prelo[c])	AL/PB	Campinas (SP)	(-s) em coda	VTCA	VTCO
Chacon (2012)	SP	João Pessoa (PB)	/s/ diante de [t, d]	VTCA	VTCO
Oliveira (2023, inédito)	AL reg. SP	Alagoas	/s/ diante de [t, d]	VTCO	VTR
Oliveira (2023, inédito)	AL reg. SP	Alagoas	/t, d/ diante de [i]	VTCO	VTR
Oliveira (2023, inédito)	AL reg. SP	Alagoas	(-r) em coda	VTR	VTCO

Fonte: Elaboração própria com base nos dados de Hora e Wetzels (2010), Chacon (2012), Oliveira (2020), Oushiro (2024, no prelo[c]) e Oliveira (2023, inédito).

No Quadro 6, resumem-se os resultados das correlações de variáveis sociolinguísticas com o preditor Estilo encontrados nas análises das amostras de fala: de migrantes baianos residentes na RMSP (ver a Seção 5) quanto ao uso de (-r) em coda; de baianos residentes em Bauru (SP) também para o emprego de (-r) em coda (Oliveira, 2020); de paraibanos residentes em São Paulo (SP) (Hora; Wetzels, 2010) para o uso de (-r); de alagoanos e paraibanos residentes em Campinas (SP) para a realização de (-r) e (-s) em posição de coda (Oushiro, 2024, no prelo[c]); de paulistas residentes em João Pessoa (PB) quanto à utilização de /s/ diante de [t, d] (Chacon, 2012), e de migrantes alagoanos regressos de São Paulo para a comunidade de origem em relação ao uso de /s/ seguido de [t, d], /t, d/ diante de [i] e de (-r) em coda (Oliveira, 2023, inédito).

No cotejo específico do (-r) em coda (marcado de azul), verifica-se que, em todos os contextos de fala [-monitorada], os migrantes, em situação de contato ou recontato, tenderam a utilizar mais a variante típica da comunidade anfitriã (VTCA) ou de onde regressaram, no caso dos alagoanos que voltaram de São Paulo. Por outro lado, em contextos de fala [+monitorada], ocorreu o oposto, os migrantes tenderam a empregar as variantes típicas da comunidade de origem (VTCO). Os dados Oushiro (2024, no prelo[c]) para o uso do (-s) em coda e dos de Chacon (2012) se alinham com os resultados de (-r) em ambos os contextos de Estilo. Entretanto os padrões do uso de /s/ diante de [t, d] e de /t, d/ diante de [i] na fala dos alagoanos regressos diferem dos resultados das demais pesquisas, inclusive para o emprego do (-r) na mesma amostra, pois eles tendem a usar mais a variante típica da comunidade de onde regressaram (VTR).

No estudo de Hora e Wetzels (2010), os autores avaliam que a concepção de estilo

laboviana não conseguiu explicar seus resultados. Eles indicam que as concepções propostas por Bell (1984) e Eckert (2005) talvez sejam mais adequadas, mas não desenvolvem essa argumentação. Oushiro (2021) interpreta seus resultados com base na consciência que os falantes têm sobre a norma local e nacional, considerando questões de identidade e convergência em relação à variedade da comunidade anfitriã. Por outro lado, Chacon (2012) e Oliveira (2020) não detalharam os resultados obtidos para esse preditor. Oliveira (2023, inédito) justifica que os migrantes alagoanos que retornam de São Paulo tendem a usar as variantes típicas da comunidade de contato no que diz respeito ao emprego de /s/ diante de [t, d] e /t, d/ diante de [i] em contextos de fala monitorada, porque eles as consideram variantes prestigiosas. Em contrapartida, eles tendem a empregar a variante típica de sua gramática nativa na realização do /r/ em coda, nos contextos de fala [+monitorada], porque os migrantes não consideram as variantes Retroflexo e Tepe como prestigiosas.

Confrontando apenas os dados de (-r) do Quadro 6 e fundamentando-se na sugestão de Hora e Wetzels (2010) e na explicação de Oushiro (2021), é possível assumir que o padrão para a fala [+monitorada] resulta da adaptação linguística ocorrida entre o cruzamento dos fatores *lugar (abstrato) da interlocução – consciência da norma local – interlocutor*.

A título de ilustração, os migrantes baianos residentes na RMSP reconhecem que essa comunidade se diferencia linguisticamente do seu lugar de origem no que diz respeito ao uso do /r/ em final de sílaba. Em contexto de fala [+monitorada], presumivelmente devido à insegurança linguística, pois na percepção do migrante pode existir a expectativa de que o entrevistador espera um bom desempenho, eles tendem a adotar a variante de sua gramática nativa e não a variante da comunidade anfitriã.

No caso dos migrantes que retornam, a título de exemplo, supõe-se que os alagoanos que voltam de São Paulo para Alagoas também reconhecem as diferenças diatópicas no uso do /r/ em coda entre o local de origem e o local para onde retornaram. Em um contexto de fala monitorada, possivelmente devido ao receio de o entrevistador achar o uso do Retroflexo ou Tepe caricato, já que eles estão num lugar onde tais variantes são atípicas, eles tendem a utilizar a variante da sua gramática nativa em relação à da comunidade de onde retornaram.

Essas discussões sobre os resultados obtidos das análises do uso de (-r) na fala de baianos residentes na RMSP em contextos de estilo corroboram a avaliação de Hora e Wetzels (2010) de que as concepções de estilo de Bell (1984) e de Eckert (2005) dariam conta de explicar os padrões de estilo dos migrantes em situação de contato dialetal, já que, como se viu, é preciso considerar o significado social da variação.

Conforme se constata no decorrer desta Seção, os processos de adaptação linguística identificados na fala de migrantes em situação de contato dialetal não podem ser reduzidos à aquisição de um segundo dialeto, como defendido por Chambers (1992) e Siegel (2010). Dessa maneira, reconhecendo que (i) os dialetos mutuamente inteligíveis, por vezes, divergem apenas

nas taxas de uso de variantes de uma mesma variável; (ii) há outros processos de adaptação linguística para além da aquisição; (iii) as adaptações linguísticas na fala de migrantes tendem a não acontecer de maneira uniforme, mas considerando as particularidades linguísticas e sociais dos dialetos em contato; (iv) os migrantes tendem a intercambiar variantes da comunidade anfitriã e de sua gramática nativa em contextos de fala [+monitorada], propõe-se o uso do termo *plasticidade dialetal* para nomear as adaptações linguísticas ocorridas na fala de migrantes em situação de contato dialetal.

### 7.3 Sugestões de análises

Outros casos de adaptação linguística percebidos na fala dos migrantes baianos residentes na RMSP precisam ser analisados empiricamente. Alguns desses fenômenos aparecem nas gravações, outros foram notados na convivência deste pesquisador com os participantes de sua amostra.

Em relação às adaptações no nível lexical, observou-se, por exemplo, que os migrantes baianos usam:

1. “canjica” com o sentido “mungunzá”;
2. “Cândida” no lugar de “Qboa”, metonímias para água sanitária;
3. “guia” com o sentido de “meio-fio”;
4. “farol” com o sentido de “sinal/sinaleira/semáforo”;
5. “lombada” no lugar de “quebra-mola”;
6. “bolacha” com o sentido de “biscoito”;
7. “papel sulfite” em vez de “papel de ofício”.

No que diz respeito às adaptações no nível fonético-fonológico, perceberam-se, por exemplo:

1. a potencialização das Vogais médias pretônicas /e/ e /o/ mais altas em relação às mais baixas;
2. a contenção do alçamento das Vogais médias pretônicas /e/ e /o/, como em “minino” e “buneca”;
3. a contenção da Ditongação do /e/ seguido de [s], como em “inglês” e “fez”, em relação à forma monotongada;

4. a potencialização do /s/ alveolar seguido de /t/ e /d/, como em “pasta” e “desde” em comparação com a variante palatal;
5. a aquisição da Ditongação do /e/ nasal, como em “entendo” e “fazendo”, em relação à forma não ditongada;
6. a aquisição do Alongamento do /a/ nasal, como em “amando” e Unicamp, em comparação com a forma não alongada.

Acerca das variáveis morfossintáticas, notaram-se:

1. a potencialização do uso do Artigo definido antes de nomes de pessoas, como em “A Maria viajou”, em relação à forma suprimida “ $\emptyset$  Maria viajou”;
2. a potencialização do uso Artigo definido diante de possessivos, como em e “A minha mãe deixa”, em comparação com a sua ausência, como em “ $\emptyset$  Minha mãe deixa” ;
3. a potencialização do Imperativo com morfologia de indicativo, como em “Faz o exercício”, em relação à forma com morfologia de subjuntivo, como em “Faça o exercício”;
4. a potencialização do Apagamento de clíticos, como “João  $\emptyset$  operou”, em relação à forma preenchida, como em “João *se* operou”;
5. a aquisição do Vocativo “meu”;
6. a potencialização do Pronome de tratamento “moço/moça”, em relação a “senhor/senhora”;
7. a realização expressiva de Diminutivos, como em “clubinho”, “amorzinho”;
8. a contenção das Interjeições “oxe” e “oxente”;
9. a potencialização da Interjeição “nossa!”.

## 8

# Considerações finais

A presente pesquisa analisou uma amostra de fala de 50 migrantes baianos residentes na Região Metropolitana de São Paulo a fim de: (i) analisar se variáveis sociolinguísticas de diferentes naturezas (lexical, fonético-fonológica e morfossintática) que perpassam os mesmos processos de adaptação linguística durante o contato dialetal; (ii) avaliar como se dá o encaixamento social e linguístico de variáveis sociolinguísticas de distintas naturezas na fala de migrantes em situação de contato dialetal. Para tanto, foram analisadas, por meio de modelos estatísticos realizados na plataforma R (R Core, 2023), as variáveis lexicais *Tangerina/Mexerica*, *Aipim/Mandioca*, *Quentinha/Marmita*, *Trabalho/Serviço*; a variável fonético-fonológica (-r) em coda; e a variável morfossintática *Negação sentencial*; assumindo que os migrantes baianos tenderiam a se adaptar às variantes tipicamente paulistas (*Mexerica*, *Mandioca*, *Marmita*, *Serviço*, /r/ *Retroflexo/Tepe* e a baixa frequência de uso de *Neg2/Neg3*) em relação às variantes tipicamente baianas (*Tangerina*, *Aipim*, *Quentinha*, *Trabalho*, /r/ *Velar/Glotal* e a alta frequência de *Neg2/Neg3*, respectivamente).

Os resultados das análises dos Itens lexicais, apresentados na Seção 4, mostraram que os baianos residentes na RMSP tendem a se adaptar às variantes da comunidade anfitriã. Quanto ao item “*Mexerica*”, observou-se que, em relação ao valor encontrado na variedade baiana (3%), os baianos residentes na RMSP aumentaram a sua realização (72%), aproximando-se da taxa encontrada em São Paulo (92%). No que diz respeito à lexia “*Mandioca*”, notou-se que os participantes da amostra analisada tenderam a aumentar a sua taxa de uso, saindo de 11%, valor encontrado na variedade baiana, para 76%, acompanhando a frequência relativa identificada em São Paulo (100%). Esse resultado chamou bastante atenção, porque se verifica que os baianos residentes na RMSP ampliaram o significado de “*Mandioca*”. Conforme se demonstrou na Seção 4, na Bahia, essa lexia nomeia a raiz tóxica utilizada para fazer farinha ou ração para animais, enquanto em São Paulo ela denomina tanto essa raiz quanto a que se cozinha para comer.

Nas análises da variável *Quentinha/Marmita*, notou-se a presença de outra variante,

“Marmitex”, cuja proporção, por sinal, foi superior (60,4%) em relação ao item “Marmita” (33,3%) e ao item “Quentinha” (6,2%). Na adaptação linguística de “Marmita”, observou-se o seu aumento de uso em comparação ao valor encontrado no dialeto baiano (13%). Esse crescimento, no entanto, reflete a ampliação de significado dessa variante, uma vez que, na Bahia, ela é empregada para nomear um vasilha com compartimentos utilizada para transportar comida ou então é a porção de comida levada de um lugar para outro, como de casa para o trabalho, por exemplo. Já em São Paulo, tal lexia nomeia a porção de comida comprada num recipiente laminado ou plástico em restaurantes ou lugares desse gênero. Devido à ausência de dados comparáveis de uso da lexia “Marmitex” nas variedades baiana e paulista, não foi possível descrever como ocorreu o seu processo de adaptação. Com relação ao item “Serviço”, verificou-se o seu aumento de uso (33%) na fala dos migrantes baianos em comparação com o valor visto na variedade baiana (2%). Esse crescimento, inclusive, ultrapassou a taxa encontrada na variedade paulista (17%).

Nos modelos estatísticos de regressão logística, nos quais se testou a correlação dos Itens lexicais (baianos e paulistas) com os preditores sociais (ver a Seção 3, observou-se que os Itens lexicais tipicamente paulistas tendem a ser utilizados por migrantes baianos que estudaram até o Ensino Fundamental, que migraram mais cedo e que estão há mais tempo residindo na comunidade anfitriã.

Na análise da distribuição de cada item lexical tipicamente paulista em duas Faixas de idade de migração (Faixa I, até 19 anos; II, +19 anos) e em três Faixas de tempo de residência (Faixa I, 1-15 anos; II, 16-31 anos; III, +32 anos), notou-se que essas variantes apresentaram tendências de uso diferentes. Em relação à Faixa de idade de migração, viu-se que os itens “Mexerica”, “Mandioca”, “Marmita” e “Marmitex” apresentaram taxas de uso maiores na Faixa I (23%, 21%, 10% e 16%, respectivamente) em comparação à Faixa II (15%, 19%, 8% e 15%, nessa mesma ordem), conforme o esperado. Por outro lado, o item “Serviço” apresentou valor baixo (5%) na Faixa I em relação à Faixa II (9%), contrariando as expectativas.

Quanto à Faixa de tempo de residência, observou-se que apenas o item “Mandioca” teve o resultado ascendente entre as Faixas I (17%), II (21%) e III (24%). As lexias “Mexerica”, “Marmitex” e “Serviço” apresentaram resultados ascendentes somente da Faixa I (12%, 7% e 4%, respectivamente) para a Faixa II (26%, 23% e 11%, nesta ordem), mas descendentes desta para a Faixa III (22%, 20% e 6%, respectivamente). O item “Marmita”, por outro lado, teve a maior taxa de uso na Faixa I (14%), mas caiu desta para Faixa II (5%) e, em seguida, apresentou um tímido crescimento para a Faixa III (6%).

Os resultados das análises de (-r) em coda, explicitados na Seção 5, mostraram que, na fala dos migrantes baianos residentes na RMSP, identificaram-se quatro principais variantes de /r/: Apagamento (63,3%); Velar/Glotal (25,9%); Retroflexo (6,7%) e Tepe (4,1%). Ao comparar esses dados com os valores das variantes de /r/ encontradas na Bahia, especificamente nos

dados de Feira de Santana (BA) (65% de Apagamento e 35% de Velar/Glotal (Guedes, 2014)) e com as proporções dos fones de /r/ detectados em São Paulo (SP) (56,3% de Apagamento; 0,4% de Velar/Glotal; 14,1% de Retroflexo; 29,2% de Tepe), percebeu-se que os migrantes baianos residentes na RMSP possivelmente adquirem o Retroflexo e o Tepe ao manterem contato com os paulistas.

Nas análises de regressão logística de efeitos mistos, nas quais se testou a correlação de (-r) em coda com preditores sociais e linguísticos, verificou-se que o Retroflexo/Tepe tende a ser mais usado na fala de migrantes baianos mais velhos, que estão há mais tempo em São Paulo, que pertencem a redes mais abertas e quando realizam o /r/ em todos os contextos fônicos precedentes e seguintes, em contexto de conversação e quando o /r/ está na posição final da palavra.

Os resultados das análises da Negação sentencial, reportados na Seção 6, evidenciam que as taxas de Neg2 (11%) e Neg3 (1,9%) dos migrantes baianos residentes na RMSP diminuíram em relação aos valores encontrados na Bahia (17% de Neg2 e 4% de Neg3, nos dados de Feira de Santana (BA)). Essa redução sugere que os migrantes baianos estão seguindo na direção dos valores vistos em São Paulo (SP): 5,8% de Neg2 e 0,2% de Neg3 (Rocha, 2013).

Os modelos estatísticos, nos quais se testou a correlação de (Neg) com os preditores sociais e linguísticos, revelaram não haver correlação com nenhum preditor social. Por outro lado, eles mostraram que a redução de Neg2 e Neg3 tende a acontecer em orações principais, com a presença de marcadores conversacionais e em sentenças cuja proposição é ativada de maneira direta.

Na Seção 7, observou-se que as variáveis sociolinguísticas de distintas natureza passam processos diferentes de adaptação e eles não estão relacionados à natureza delas, mas aos casos de regularidade e irregularidade linguística dos dialetos em contato. Nos casos de regularidade, isto é, quando os dialetos em contato compartilham a mesma variável sociolinguística, mas se diferenciam apenas quanto às taxas de uso de suas variantes, identificaram-se os processos de *potencialização* e *contenção*. Já nos casos de irregularidade, i.e., quando os dialetos em contato se diferenciam basicamente quanto à presença ou à ausência de determinadas variantes, identificaram-se os processos de *aquisição*, *ampliação de contexto* e *ampliação semântica*.

Ademais, ao comparar os dados da amostra aqui investigada com os resultados encontrados nas análises da amostra de fala de migrantes baianos em Bauru (SP) (Oliveira, 2020) e com os resultados de Santana (2023) sobre a fala de migrantes sergipanos em São Paulo (SP), depreendeu-se que:

- (i) Conforme já notado por Oushiro (2024, no prelo[a]), variáveis de distintas natureza não têm o mesmo padrão de encaixamento social;
- (ii) Migrantes da mesma origem residentes em diferentes pontos de uma região,

como capital e interior, podem apresentar resultados diferentes para uma mesma variável sociolinguística;

(iii) Migrantes de origens distintas, ainda que de áreas geográficas próximas, e residentes numa mesma comunidade anfitriã podem apresentar distintos padrões sociais para uma mesma variável sociolinguística;

(iv) O uso das variantes típicas da comunidade anfitriã pode diminuir ao longo do tempo, conforme visto no trabalho de Souza (2019);

(v) Como já observado por Oushiro (2024, no prelo[a]), na comunidade anfitriã, em contexto de fala monitorada, os migrantes tendem a usar regras de sua gramática nativa.

O presente estudo também chegou à conclusão de que o termo “acomodação a longo prazo”, utilizado por Trudgill (1986), e o termo “aquisição de segundo dialeto”, proposto por Chambers (1992) e Siegel (2010), não abarcam as particularidades dos resultados obtidos da amostra aqui investigada e, em substituição a ele, sugere-se o termo “plasticidade dialetal”.

Por fim, os achados deste estudo permitem advogar por:

(i) Uma metodologia capaz de identificar e classificar os processos de adaptação linguística de migrantes em situação de contato dialetal;

(ii) Estudos sobre as adaptações linguísticas de crianças e adolescentes migrantes e a sua relação com o ensino e o ambiente escolar;

(iii) Investigações sobre a fala de migrantes em contexto de retorno para a comunidade de origem;

(iv) Pesquisas acerca da interferência da fala adaptada dos migrantes em situação de retorno na fala da comunidade para a qual regressaram;

(v) Estudos sobre as atitudes linguísticas dos migrantes quanto à fala adaptada dos próprios migrantes;

(vi) Pesquisas sobre a avaliação e a percepção da comunidade de origem a respeito da fala adaptada dos que migraram;

(vii) Estudos acerca dos significados sociais das variantes para as quais os migrantes em situação de contato dialetal se adaptam.

## Referências

- Adant, J. A. d. S. Difusão dialetal: o caso dos alagoanos em Brasília. In: Tarallo, F. (Ed.). **Fotografias sociolinguísticas**. Campinas: Unicamp, 1989. P. 181–197.
- Aguilera, V. d. A. Léxico e áreas dialetais: o que podem demonstrar os dados do ALiB. **Anais ABRALIN 2009**, Universidade Federal da Paraíba, p. 4219–4233, 2009. Disponível em: <[http://www.leffa.pro.br/tela4/Textos/Textos/Anais/ABRALIN\\_2009\\_vol\\_2/PDF-VOL2/Microsoft%20Word%20-%20Vanderci%20de%20Andrade%20Aguilera.pdf](http://www.leffa.pro.br/tela4/Textos/Textos/Anais/ABRALIN_2009_vol_2/PDF-VOL2/Microsoft%20Word%20-%20Vanderci%20de%20Andrade%20Aguilera.pdf)>. Acesso em: 15 ago. 2023.
- Alkimim, M. G. R. d. **As negativas sentenciais no dialeto mineiro: uma abordagem variacionista**. 2001. 267 f. Tese (Doutorado) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte. 2001.
- Almeida, N. L. F. d.; Carneiro, Z. O. N. **Coleção de amostras da língua falada no Semiárido Baiano**. Feira de Santana: UEFS, 2008.
- Alves, M. I. P. M. **Atitudes linguísticas de nordestinos em São Paulo: abordagem prévia**. 1979. 226 f. Diss. (Mestrado) – Programa de Pós-Graduação em Linguística, Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas. 1979.
- Amaral, A. **O dialeto caipira**. São Paulo: O livro, 1920. Disponível em: <<http://www.dominiopublico.gov.br/download/texto/bi000004.pdf>>. Acesso em: 7 abr. 2023.
- Amorim, T. d. S.; Costa, C. d. S. S. M. d. As vogais médias pretônicas na fala de gaúchos em situação de contato dialetal com teresinenses. **Entreletras**, Universidade Federal do Tocantins, v. 10, p. 267–287, jun. 2019. Disponível em: <<https://sistemas.uft.edu.br/periodicos/index.php/entreletras/article/view/6695/15105>>. Acesso em: 3 ago. 2023.
- Barbosa, S. P. **Análise do /s/ em coda na fala de migrantes alagoanos e paraibanos em Campinas**. 2022. 85 f. Monografia (Graduação) – Instituto de Estudos da Linguagem da Universidade Estadual de Campinas, Campinas. 2022.

- Barme, S. A negação no brasileiro falado informal. **Zeitschrift für Romanische Philologie**, v. 121, p. 405–425, 3 2005. Disponível em:  
<<https://ur.booksc.eu/book/42113581/da4103>>. Acesso em: 12 jun. 2021.
- Barnes, J. A. Network and political process. In: Mitchell, J. C. (Ed.). **Social networks in urban situations**. Manchester: Manchester University Press, 1969.
- Becker, K. The sociolinguistic interview. In: Mallinson, C.; Becky, C.; Herk, G. V. (Ed.). **Data collection in Sociolinguistics: Methods and Applications**. New York: Routledge, 2013. P. 91–100.
- Bell, A. Language style as audience design. **Language in Society**, v. 13, p. 145–204, 1984.
- Biderman, M. T. C. A face quantitativa da linguagem: um dicionário de frequências do português. **Alfa**, v. 42, Especial, p. 161–181, 1998. Disponível em:  
<<https://periodicos.fclar.unesp.br/alfa/article/view/4049>>. Acesso em: 20 out. 2023.
- Bieler da Silva, M. E. **Entre duas metrópoles: (-r) em Itanhandu**. 2015. 170 f. Diss. (Mestrado) – Programa de Pós-Graduação em Semiótica e Linguística Geral do DL–FFLCH–USP, São Paulo. 2015.
- Bortoni-Ricardo, S. M. **Do campo para a cidade: estudo sociolinguístico de migração e redes sociais**. São Paulo: Parábola, 2011 [1985].
- \_\_\_\_\_. **Nós cheguem na escola, e agora?: Sociolinguística & Educação**. São Paulo: Parábola, 2005.
- Brandão, S. **A geografia linguística no Brasil**. São Paulo: Ática, 1991.
- Brasil. Lei nº 9.394, de 20 de dezembro de 1996. **Diário Oficial da República Federativa do Brasil**, Brasília, DF, 23 dez. 1996. Disponível em:  
<[http://www.planalto.gov.br/ccivil\\_03/Leis/L9394.htm](http://www.planalto.gov.br/ccivil_03/Leis/L9394.htm)>. Acesso em: 9 mar. 2023.
- \_\_\_\_\_. Lei nº 9.459, de 13 de maio de 1997. **Diário Oficial da República Federativa do Brasil**, Brasília, DF, 14 mai. 1997. Disponível em:  
<[https://www.planalto.gov.br/ccivil\\_03/Leis/L9459.htm#art1](https://www.planalto.gov.br/ccivil_03/Leis/L9459.htm#art1)>. Acesso em: 9 mar. 2023.
- \_\_\_\_\_. **Parâmetros Curriculares Nacionais: Língua Portuguesa**. Brasília: Secretaria de Educação Fundamental, 1997.
- Callou, D.; Moraes, J. a.; Leite, Y. Variação e diferenciação dialetal: a pronúncia do /r/ no português do Brasil. In: Koch, I. V. (Ed.). **Gramática do português falado, vol. VI**. 2. ed. Campinas/SP: Editora da Unicamp, 2002 [1996]. P. 463–490.

- Callou, D.; Silva, G. M. O uso do artigo definido em contextos específicos. In: Hora, D. (Ed.). **Diversidade linguística no Brasil**. João Pessoa: Ideia, 1997. P. 11–26.
- Cardoso, S. A. M. d. S. et al. **Atlas linguístico do Brasil**. Londrina: Eduel, 2014.
- Carreão, V. Transformações econômicas e mudança linguística: a língua em Louveira/SP. **Estudos linguísticos (São Paulo. 1978)**, Grupo de Estudos Linguísticos do Estado de São Paulo, v. 48, p. 709–727, jul. 2019. Disponível em: <<https://revistas.gel.org.br/estudos-linguisticos/article/view/2318>>. Acesso em: 22 abr. 2023.
- Castilho, A. T. (Ed.). **Gramática do português falado**. Campinas: Editora da Unicamp, 1990.
- Cavalcante, R. **A negação pós-verbal no português brasileiro: análise descritiva e teórica de dialetos rurais de afro-descendentes**. 2007. 160 f. Diss. (Mestrado) – Instituto de Letras da Universidade Federal da Bahia, Salvador. 2007.
- Chacon, K. d. A. **Contato dialetal: análise do falar paulista em João Pessoa**. 2012. 118 f. Diss. (Mestrado) – Programa de Pós-Graduação em Linguística da Universidade Federal da Paraíba, João Pessoa. 2012.
- Chambers, J. K. Dialect acquisition. **Language**, v. 68, n. 4, p. 673–705, 1992.
- Chaves, M. d. F. G. Mulheres que migram solteiras: aspectos da migração interna feminina no Brasil, 1981/1991. In: Teixeira, P. E.; Braga, A. M. d. C.; Baeninger, R. (Ed.). **Migrações: implicações passadas, presentes e futuras**. Marília: Cultura acadêmica, 2012.
- Eckert, P. Three waves of variation study: the emergence of meaning in the study of variation. **Annual Review of Anthropology**, v. 41, p. 87–100, 2012. Disponível em: <<http://www.stanford.edu/~eckert/PDF/ThreeWavesofVariation.pdf>>. Acesso em: 4 nov. 2023.
- \_\_\_\_\_. Variation, convention, and social meaning. **Annual Meeting of the Linguistic Society of America**, p. 1–32, 2005.
- Ferraz, L. **Comunicação pessoal**. [S.l.: s.n.], jun. 2023.
- Ferreira, A. B. d. H. **Miniaurélio Século XXI: o minidicionário da língua portuguesa**. 5. ed. Rio de Janeiro: Nova Fronteira, 2001.
- \_\_\_\_\_. **Novo Dicionário da Língua Portuguesa**. 2. ed. Rio de Janeiro: Nova Fronteira, 1986.

- Figueiredo, C. R. d. S. **Topodinâmica da variação do português gaúcho em áreas de contato intervarietal no Mato Grosso**. 2014. 299 f. Tese (Doutorado) – Programa de Pós-Graduação em Letras, Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre. 2014.
- Figuereido, J. G. d. S.; Souza, E. S. d. O uso do imperativo por migrantes baianos em São Paulo: um estudo comparativo. In: VII ENCONTRO DE SOCIOLINGÜÍSTICA, 2017, Aracaju.
- Fontes, P. **Um Nordeste em São Paulo**. Rio de Janeiro: FGV, 2008.
- Furtado da Cunha, M. A. O modelo das motivações competidoras no domínio funcional da negação. **D.E.L.T.A.**, n. 17, 1 2001.
- Gasiorek, J. Theoretical Perspectives on Interpersonal Adjustments in Language and Communication. In: Giles, H. (Ed.). **Communication Accommodation Theory: Negotiating Personal Relationships and Social Identities Across Contexts**. United Kingdom: Cambridge University Press, 2016. P. 13–35.
- Giles, H.; Taylor, D.; Bourhis, R. Toward a theory of interpersonal accommodation through speech: some Canadian data. **Language in Society**, v. 2, p. 177–192, 1973.
- Goldnadel, M. et al. Estratégias alternativas de negação sentencial na região sul do Brasil: análise da influência de fatores pragmáticos a partir de dados do projeto VARSUL. **Revista de Estudos da Linguagem**, 21.n2, p. 35–74, 2013.
- Gomes da Silva, F. **Alagoanos em São Paulo e a concordância nominal de número**. 2014. 103 f. Diss. (Mestrado) – Programa de Pós-Graduação em Semiótica e Linguística Geral do DL–FFLCH–USP, São Paulo. 2014.
- Gries, S. T. **Quantitative Corpus Linguistics with R: a practical introduction**. New York e London: Routledge, 2009.
- Guedes, S. Emprego do artigo definido em situação de contato dialetal: um estudo da fala de migrantes paraibanos em São Paulo. **Domínios da lingu@gem**, 14.n4, p. 1401–1432, 2019 [2017]. Disponível em: <<https://seer.ufu.br/index.php/dominiosdelinguagem/article/view/46873>>. Acesso em: 20 ago. 2022.
- Guedes, S. C. **Apagamento do /r/ nas falas popular e culta de Feira de Santana**. 2014. 100 f. Diss. (Mestrado) – Universidade Estadual de Feira de Santana, Feira de Santana. 2014.
- Guirelli, F. d. O. **O uso variável do artigo definido diante de pronomes possessivos no contato dialetal de migrantes paraibanos em Campinas**. 2018. 45 f. Monografia (Graduação) – Instituto de Estudos da Linguagem da Universidade Estadual de Campinas, Campinas. 2018.

- Guy, G. R.; Zilles, A. **Sociolinguística quantitativa: instrumental de análise**. São Paulo: Parábola, 2007.
- Hellwig Birgit; Geerts, J. **ELAN – Linguistic Annotator. Versão 6.0**. Nijmegen: Max Planck Institute for Psycholinguistics, 2020. Disponível em: <<https://archive.mpi.nl/tla/elan>>. Acesso em: 13 fev. 2020.
- Hoffman, M.; Walker, J. A. Ethnolects in the city: ethnic orientation and linguistic variation in Toronto English. **Language Variation and Change**, v. 22, p. 37–67, 2010.
- Hora, D. d. **Projeto Variação Linguística no Estado da Paraíba**. 1993. Disponível em: <<https://projetoalpb.com.br/>>. Acesso em: 29 jul. 2021.
- Hora, D. d.; Negrão, E. V. (Ed.). **Estudos da Linguagem: casamento entre temas e perspectivas**. João Pessoa: Ideia, 2011.
- Hora, D. d.; Wetzels, L. Róticos: uma “po[h, r, ʃ]ta” entre paraibanos e paulistanos. **Linguística**, v. 24, p. 51–76, dez. 2010.
- Houaiss, A. **Minidicionário Houaiss da língua portuguesa**. 4. ed. Rio de Janeiro: Objetiva, 2012.
- IBGE. **Censo demográfico 2000 Migração e deslocamento**. Rio de Janeiro: IBGE, 2000.
- \_\_\_\_\_. **Censo demográfico 2010 Nupcialidade, fecundidade e migração**. Rio de Janeiro: IBGE, 2010.
- IPEA. **Comunicados do IPEA no. 115 – Perfil dos migrantes em São Paulo**. 2011. Disponível em: <[http://www.ipea.gov.br/portal/images/stories/PDFs/comunicado/111006\\_comunicadoipea115.pdf](http://www.ipea.gov.br/portal/images/stories/PDFs/comunicado/111006_comunicadoipea115.pdf)>. Acesso em: 23 set. 2019.
- Jespersen, O. **Negation in English and other languages**. Kobenhan: AF.Host, 1917.
- Labov, W. **Padrões sociolinguísticos**. Tradução: Marcos Bagno, Maria Marta Pereira Scherre e Caroline Rodrigues Cardoso. São Paulo: Parábola, 2008 [1972].
- \_\_\_\_\_. **Principles of linguistic change: social factors**. Oxford & Cambridge: Blackwell, 2001.
- \_\_\_\_\_. Some Principles of Linguistic Methodology. **Language in Society**, v. 1, n. 1, p. 97–120, 1972.
- \_\_\_\_\_. **The social stratification of English in New York City**. Cambridge: Cambridge University Press, 2006 [1966].
- Leite, C. M. B. **O /R/ em posição de coda silábica no falar campineiro**. 2010. 227 f. Tese (Doutorado) – Instituto de Estudos da Linguagem da Universidade Estadual de Campinas – UNICAMP, Campinas. 2010.

- Levshina, N. **How to do Linguistics with R**. Amsterdam/Philadelphia: John Benjamins, 2015.
- Lima, N. et al. Plasticidade fenotípica. **Revista Ciência Elementar**, v. 5, n. 2, p. 1–7, jun. 2017. Disponível em: <<https://rce.casadasciencias.org/rceapp/art/2017/017/>>. Acesso em: 17 set. 2023.
- Lopes, J. B. **Sociedade industrial no Brasil**. São Paulo: Difel, 1964.
- Lucchesi, D. **Projeto Vertentes do Português Rural da Bahia**. 2002. Disponível em: <<http://www.vertentes.ufba.br/home>>. Acesso em: 16 nov. 2020.
- Marques, S. M. O. **As vogais médias pretônicas em situação de contato dialetal**. 2006. 161 f. Tese (Doutorado) – Programa de Pós-Graduação em Letras Vernáculas da Faculdade de Letras da Universidade Federal do Rio de Janeiro, Rio de Janeiro. 2006.
- Marroquim, M. **A língua do nordeste: Alagoas e Pernambuco**. Maceió: edUFAL, 2008 [1934].
- Martins, M. d. S. **A palatalização de oclusivas dentais em contato dialetal**. 2008. 145 f. Diss. (Mestrado) – Faculdade de Letras – Universidade Federal do Rio de Janeiro, Rio de Janeiro. 2008.
- Mattos e Silva, R. V. **Ensaio para uma sócio-história do português brasileiro**. São Paulo: Parábola, 2004.
- Mendes, R. B. Relatório Científico Final apresentado à FAPESP. São Paulo, 2012. Disponível em: <<http://projetosp2010.fflch.usp.br/producao-bibliografica>>. Acesso em: 26 nov. 2023.
- Mendes, R. B.; Oushiro, L. **Documentação do Projeto SP2010 – Construção de uma amostra da fala paulistana**. 2013. Disponível em: <<http://projetosp2010.fflch.usp.br/producao-bibliografica>>. Acesso em: 26 nov. 2023.
- Menezes, M. A. d. Migrações e Mobilidades: Repensando Teorias, Tipologias e Conceitos. In: Teixeira, P. E.; Braga, A. M. d. C.; Baeninger, R. (Ed.). **Migrações: implicações passadas, presentes e futuras**. Marília: Cultura acadêmica, 2012.
- Meurer, M. **O português de migrantes sulistas no nordeste do Brasil: variação e mudança de marcas regionais no contato intervarietal**. 2022. 344 f. Tese (Doutorado) – Programa de Pós-Graduação em Letras, Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre. 2022.
- Milroy, L. **Language and social networks**. 2. ed. Oxford: Basil Blackwell, 1987 [1980].

- Mitchell, J. C. The concept and use of social networks. In \_\_\_\_\_. **Social networks in urban situations**. Edição: James Clyde Mitchell. Manchester: Manchester University Press, 1969.
- Mourão, N. R. Mobilidade e difusão linguística no interior de São Paulo: uma análise de Campinas e Jundiaí. **Projeto de Mestrado**. Ms. Campinas, 2018.
- Nascimento, C. A. R. d. **A negação no português falado em Vitória/ES**. 2014. 98 f. Diss. (Mestrado) – Programa de Pós-Graduação *Stricto Sensu* em Linguística da Universidade Federal do Espírito Santo, Vitória. 2014.
- Noll, V. **O português brasileiro: formação e contrastes**. São Paulo: Globo, 2008.
- Nunes, E. S. d. O. A negação no português falado do Rio de Janeiro: um estudo baseado em *corpus*. **Revista do SELL**, 04.n1, p. 1–19, 2014.
- Oliveira, A. J. **Projeto PORTAL: variação linguística no português alagoano**. Maceió, 2017. Disponível em: <<http://www.portuguesalagoano.com.br/>>. Acesso em: 26 nov. 2023.
- Oliveira, A. A. d. Migração de retorno. **Relatório final de estágio de Pós-Doutorado**. Ms. Campinas, 2023, inédito.
- Oliveira, J. M. **O apagamento do /R/ implosivo na norma culta de Salvador**. 1999. 80 f. Diss. (Mestrado) – Instituto de Letras da Universidade Federal da Bahia, Salvador. 1999.
- Oliveira, M. A. J. d. **Dialetos em contato: acomodação dialetal por migrantes baianos habitantes da cidade de Bauru**. 2020. 88 f. Diss. (Mestrado) – Universidade Estadual Paulista, Araraquara. 2020.
- Oushiro, L. A fala de migrantes internos: uma agenda de estudos. In: ABRALIN50, 2019, Maceió. Pôster. Disponível em: <<https://zenodo.org/records/2658839>>. Acesso em: 20 nov. 2023.
- \_\_\_\_\_. As variáveis sexo/gênero e indivíduo em situação de contato dialetal. In: Carvalho, D.; Brito, D. (Ed.). **Gênero e lingua(gem): formas e uso**. Salvador: EDUFBA, 2020.
- \_\_\_\_\_. Contrasting Age of Arrival and Length of Residence in Dialect Contact. **University of Pennsylvania Working Papers in Linguistics**, v. 25, n. 2, 2020.
- \_\_\_\_\_. Converging to local and supralocal norms: stylistic variation in migrants speech. In: NWAV49, 2021, Austin. Comunicação oral.

- Oushiro, L. Dialect contact in lusophone communities. In: Carvalho, A. M.; Oushiro, L. (Ed.). **The Oxford Handbook of the Portuguese Language**. Oxford: Oxford University Press, 2024, no prelo.
- \_\_\_\_\_. **Identidade na pluralidade: Avaliação, produção e percepção linguística na cidade de São Paulo**. 2015. 394 f. Tese (Doutorado) – Programa de Pós-Graduação em Semiótica e Linguística Geral, Departamento de Linguística, Universidade de São Paulo, São Paulo. 2015.
- \_\_\_\_\_. Interaction, confounding effect, and collinearity in the analysis of Brazilian internal migrants speech. In: Fernández-Mallat, V.; Nycz, J. (Ed.). **Dialect Contact: Speaker-level and Dialect-level Phenomena**. Washington DC: Georgetown University Press, 2024, no prelo.
- \_\_\_\_\_. **Introdução à Estatística para Linguistas**. Campinas: Editora da ABRALIN, 2022. Disponível em: <<https://editora.abralin.org/publicacoes/introducao-a-estatistica-para-linguistas/>>. Acesso em: 2 set. 2023.
- \_\_\_\_\_. Múltiplas variáveis na fala de nordestinos residentes em São Paulo. In: Brescansini, C. R.; Monaretto, V. N. d. O. (Ed.). **Sociolinguística no Brasil: textos selecionados**. Porto Alegre: EDIPUCRS, 2020. P. 121–154.
- \_\_\_\_\_. **O estudo da fala de migrantes internos: desafios, procedimentos e resultados do Projeto Acomodação**. Ms. Campinas, 2024, no prelo.
- \_\_\_\_\_. **Processos de acomodação dialetal na fala de nordestinos residentes em São Paulo**. 2017. Disponível em: <<http://www.bv.fapesp.br/pt/auxilios/93662/processos-de-acomodacao-dialetal-na-fala-de-nordestinos-residentes-em-sao-paulo/>>. Acesso em: 2 set. 2023.
- \_\_\_\_\_. Transcrição de entrevistas sociolinguísticas com o ELAN. In: Freitag, R. M. K. (Ed.). **Metologia de coleta e manipulação de dados em Sociolinguística**. São Paulo: Blucher, 2014. Disponível em: <<http://blucheropenaccess.com.br/articles/download/290>>. Acesso em: 15 out. 2023.
- Oushiro, L.; Mendes, R. B. O apagamento de (-r) em coda nos limites da variação. **Veredas - Revistas de Estudos Linguísticos**, Universidade Federal de Juiz de Fora, v. 18, p. 251–266, 2014. Disponível em: <<https://periodicos.ufjf.br/index.php/veredas/article/view/24963>>. Acesso em: 8 set. 2023.

- Oushiro, L. et al. Estudos sociolinguísticos sobre contato dialetal: contribuições do VARIEM e agenda de pesquisa. **Cadernos de Estudos Linguísticos**, v. 65, Dossiê comemorativo dos 45 anos da CEL, p. 1–18, 2023. Disponível em: <<https://periodicos.sbu.unicamp.br/ojs/index.php/cel/article/view/8673331>>. Acesso em: 26 nov. 2023.
- Pal, D. C.; Araújo, P. P. Sintaxe e Semântica. In: Petter, M. (Ed.). **Introdução à linguística Africana**. São Paulo: Contexto, 2015. P. 159–192.
- Pereira, S. S. d. C. **Um estudo sociofuncionalista das variantes negativas não e num no português falado de Vitória da Conquista**. 2018. 130 f. Diss. (Mestrado) – Universidade Estadual do Sudoeste da Bahia, Vitória da Conquista. 2018.
- Possatti, L. **Acomodação dialetal de cariocas residentes em João Pessoa: uma análise sociolinguística**. 2020. 123 f. Diss. (Mestrado) – Universidade Federal da Paraíba, João Pessoa. 2020.
- R Core, T. **R: A language and environment for statistical computing**. Vienna: R-Project, 2023. Disponível em: <<http://www.R-project.org/>>. Acesso em: 1 mai. 2023.
- Rocha, R. S. **A negação dupla no português paulistano**. 2013. 97 f. Diss. (Mestrado) – Programa de Pós-Graduação em Semiótica e Linguística Geral, Departamento de Linguística, Universidade de São Paulo, São Paulo. 2013.
- Romano, V. P.; Cá, J. F. Mandioca, macaxeira e aipim na Região Sudeste do Brasil: distribuição diatópica e comentários geolinguísticos dos informantes. **MOARA Revista Eletrônica do Programa de Pós-Graduação em Letras**, v. 0, n. 55, p. 109–128, 2020. Disponível em: <<https://www.periodicos.ufpa.br/index.php/moara/article/view/9171>>. Acesso em: 16 ago. 2023.
- Roncarati, C. A negação no português falado. In: Macedo, A.; Roncarati, C.; Mollica, M. C. (Ed.). **Variação e Discurso**. Rio de Janeiro: Tempo Brasileiro, 1996. P. 97–112.
- Sadler, J. **Introduction to network analysis with R**. 2017. Disponível em: <<https://www.jessesadler.com/post/network-analysis-with-r/>>. Acesso em: 28 mai. 2020.
- Santana, A. d. L. A pronúncia variável de /t/, /d/ diante de [i] na fala de migrantes sergipanos em São Paulo. **Estudos linguísticos (São Paulo. 1978)**, Grupo de Estudos Linguísticos do Estado de São Paulo, v. 50, p. 1283–1304, dez. 2021. Acesso em: 13 jul. 2023.

- Santana, A. d. L. **Acomodação dialetal e covariação na fala de migrantes sergipanos em São Paulo**. 2023. 181 f. Tese (Doutorado) – Programa de Pós-Graduação em Semiótica e Linguística Geral, Departamento de Linguística, Universidade de São Paulo, São Paulo. 2023.
- \_\_\_\_\_. **As vogais médias pretônicas na fala de sergipanos em São Paulo**. 2018. 162f. Diss. (Mestrado) – Departamento de Linguística, Programa de Pós-Graduação em Semiótica e Linguística Geral, Universidade de São Paulo, São Paulo. 2018.
- Santana, J. C. D.; Nascimento, P. B. S. A negação no português falado de Matinha/BA: um estudo sociolinguístico. **Letra Magna**, 07.n14, p. 1–17, 2011.
- Schwenter, S. A. The pragmatics of negation in Brazilian Portuguese. **Lingua**, v. 115, p. 1427–1456, 2005. Disponível em: <[www.elsevier.com/locate/lingua](http://www.elsevier.com/locate/lingua)>. Acesso em: 17 mai. 2020.
- Serra, F. P. **“Eu não digo ‘não’ duas vezes não”**: usos e percepções avaliativas sobre a dupla negação no português falado no Maranhão. 2018. 190 f. Diss. (Mestrado) – Universidade Federal do Maranhão, São Luís/MA. 2018.
- Siegel, J. **Second dialect acquisition**. Cambridge: Cambridge University Press, 2010.
- Sousa, L. T. d. **Formas reduzidas de itens negativos no português brasileiro**. 2007. 113 f. Diss. (Mestrado) – Programa de Pós-Graduação em Estudos Linguísticos da Faculdade de Letras da Universidade Federal de Minas Gerais, Belo Horizonte. 2007.
- Souza, E. S. d. O uso variável do imperativo de migrantes baianos em São Paulo. **Domínios da Linguagem**, v. 13, p. 1433–1464, 2019.
- Tagliamonte, S. **Variationist sociolinguistics: change, observation, interpretation**. Malden, MA: Wiley-Blackwell, 2011.
- Trudgill, P. **Dialects in contact**. New York: Basil-Blackwell, 1986.
- \_\_\_\_\_. **New-dialect Formation. The Inevitability of Colonial Englishes**. Edinburgh: Edinburgh University Press, 2004.
- Weinreich, U.; Labov, W.; Herzog, M. I. **Fundamentos empíricos para uma teoria da mudança linguística**. Tradução: Marcos Bagno. São Paulo: Parábola, 2006 [1968].
- Woods, A.; Fletcher, P.; Hughes, A. **Statistics in language studies**. Cambridge: Cambridge University Press, 1986.

## Apêndice - A

### Roteiro para as entrevistas sociolinguísticas

Esta é uma adaptação do roteiro utilizado nas entrevistas sociolinguísticas do Projeto Acomodação (Oushiro, 2017).

#### Primeira parte

##### **Bairro** (aprox. 10 min.)

Objetivos: descobrir o “grau de enraizamento” do participante no bairro onde vive/outros bairros; descobrir padrões de sociabilidade nos diferentes bairros; descobrir o “grau de mobilidade” da pessoa.

- 1 Há quanto tempo você mora nesse bairro?
- 2 Você gosta de morar nesse bairro?
  - a. (Se o participante mora há bastante tempo) Como era o bairro antes/quando você se mudou pra cá ? Mudou muito?
  - b. (Se o participante não mora lá há muito tempo, ou se nem sempre viveu ali) Em que outros lugares você já morou? Como era lá comparado com esse bairro aqui? Onde você preferia morar?
- 3 O que esse bairro tem de diferente de outros bairros em x (qualquer cidade da Região Metropolitana de São Paulo)?
- 4 Aqui costuma haver festas do bairro? Existe algum lugar no bairro em que as pessoas se reúnem?
- 5 Você conhece seus vizinhos?
- 6 Se você precisar de ajuda, a quem você recorre? Se você ficar doente, a quem você pode pedir para tomar conta de sua família?
- 7 Já aconteceu alguma coisa aqui que te fez pensar em se mudar?
- 8 Hoje, tem algum outro bairro em que você gostaria de morar?

##### **Infância** (aprox. 5 min.)

Objetivos: falar sobre tópicos menos formais (em geral, as pessoas gostam de falar de sua infância); obter informações sobre grau de mobilidade do participante; obter informações sobre

escolaridade.

9. E como foi a sua infância na Bahia? Você pode contar um pouco de como foi, o que você fazia...?
  - a. Brincava na rua/dentro de casa? Do que vocês brincavam?
  - b. Como eram os seus pais? Eram rígidoss...? Você tinha horário para estar em casa?
  - c. Vocês tinham alguma tradição de família?
10. O que mais te marcou na sua época de escola? Como era a escola? Você sempre estudou na mesma escola? Você gostava de ir pra escola? Tem algum professor que te marcou? Até que série você estudou?
11. Você acha que a escola fornece aquilo que uma pessoa precisa para encontrar um emprego?

### **Família** (aprox. 5 min.)

Objetivo: obter informações sobre rede social do participante, grau de enraizamento no bairro/cidade.

(Com vistas a coletar o artigo antes de antropônimos)

12. Você tem irmãos? (é possível que esta informação já tenha aparecido na parte sobre infância; neste caso, falar algo como: “Você disse que tem uma irmã ... Você tem outros irmãos?”) Onde eles estão? O que fazem?
13. Onde seus pais nasceram?
14. Sua família também vive aqui em São Paulo (ou outra cidade da Região Metropolitana de São Paulo em que, porventura, morem)? (Se sim, em que bairros? Moram próximo? Se não, onde? Tem contato sempre?)
15. Com quem você mora? (Para os adultos) Eles trabalham? (Para crianças e adolescentes) Todos estudam?
16. Muitas pessoas falam que as crianças hoje são mal-educadas. O que você acha disso? Você acha que as crianças hoje têm menos respeito pelos adultos?

17. No passado, esperava-se que as mulheres ficassem em casa para tomar conta das crianças. E hoje, como é isso? Mudou?
  - a. Na sua casa, os homens ajudam nos afazeres domésticos? O que você acha de um homem ficar em casa e cuidar dos filhos?

**Motivos da migração e avaliação que o informante faz da comunidade anfitriã** (aprox. 10 min.)

Objetivos: identificar os motivos que levaram o informante a migrar para São Paulo; saber como ocorreu o processo de instalação dele na comunidade anfitriã e observar as avaliações que ele faz dela.

18. Qual o motivo de você ter vindo morar em São Paulo?
19. Foi a primeira vez que você saiu da cidade em que você nasceu para morar em outra? (Se não) Quais foram as outras cidades e quanto tempo ficou por lá?
20. Alguma situação específica facilitou a sua vinda para cá? (família que já estava em São Paulo, trabalho já definido, estudos)
21. Você gosta de morar em São Paulo (Ou outra cidade da Região Metropolitana de São Paulo em que, porventura, more)? (Se não) Em que lugar preferiria morar? Por quê?
22. O que você acha que caracteriza a cidade (tanto as coisas boas quanto ruins)?
23. O que você acha que caracteriza o paulista (tanto as coisas boas quanto ruins)?
24. O que você mais gosta aqui?
25. O que você não gosta aqui? (a depender do tópico mencionado pelo participante - violência, trânsito, poluição etc. – procurar explorar mais o assunto. P.ex.: você já foi assaltado? O que aconteceu? (para obter narrativa pessoal) O que o governo deve fazer pra solucionar esse problema? (para obter uma fala mais distanciada))
26. Para as pessoas que não vivem aqui, como você acha que elas imaginam que seja São Paulo? Qual é a imagem que as pessoas de fora têm daqui?

**Trabalho/ocupação** (aprox. 5 min.)

Objetivos: obter informações sobre rede social do participante; características socioeconômicas.

27. Em que você trabalha? É aqui por perto? (se não souber onde a pessoa trabalha)
28. Como você faz para chegar até lá? Quais meios de transporte você utiliza?
29. O que você faz? Faz tempo que você trabalha lá?
30. O que você mais gosta lá? E o que menos gosta?
31. Como é a sua relação com seus colegas?
32. Que outro cargo você deseja alcançar?
33. Qual é a profissão dos seus sonhos?
34. Se você ganhasse na Mega Sena, o que você faria?

**Lazer** (aprox. 5 min.)

Objetivos: obter informações sobre rede social do participante; mobilidade na cidade; características socioeconômicas.

35. E nas horas de lazer, o que você e sua família gostam de fazer? (Se saem) Vão pra que lugares?
36. Você acha que São Paulo tem boas opções de lazer? Quais?
37. A maioria dos seus amigos mora aqui nesse bairro mesmo? (Se não) Onde?
38. Quais s o seus amigos mais antigos?
39. Você costuma viajar? Pra que lugares já viajou? Que lugares gostaria de conhecer?

**Segunda parte**

**Estímulos sociolinguísticos** (aprox. 10 min)

Objetivo: fornecer estímulos através dos quais algumas variáveis sociolinguísticas possam ser coletadas.

**Lista de palavras**

40. Quais dessas palavras você acha que o paulista falaria diferente?

---

abajur	discoteca	mulher
açúcar	doido	necessidade
alma	erguer	normal
aniversário	falar	oito
antigo	fêmur	órgão
argola	ferver	perdido
atitude	firme	perto
bandido	flúor	pesado
biscoito	furgão	porca
caráter	geladeira	porto
carteiro	gérmen de trigo	relevante
cerca	gordo	sabor
cerveja	indireta	soberba
chácara	irmã	sorrir
circo	justiça	sorvete
cisne	mártir	teatro
colega	melhor	trabalho
crítica	morar	turco
curto	mortadela marba	urgente
diferente	muito	verdade

## Leitura de texto

### As estrelas

O frio ar da noite não desencorajava José de ficar ao relento. Aquela hora, após o jantar, era um de seus momentos preferidos do dia. Enquanto o pai já se preparava para dormir, ele ia até a parte de trás da casa, numa pequena colina, olhava para o alto e contemplava a beleza dos pontos luminosos na escuridão.

Jovem, mas quase um homem, ele via as estrelas e imaginava. Para cada um daqueles pontinhos prateados no céu, criava uma história. Assim, José viajava até o lugar mais distante que alguém poderia ir. E, ao descer para sua casa, deitava-se e aquela história continuava até se transformar em sonho.

Certa vez, no meio de uma delas, um velho índio lhe disse:

“Você não precisa ficar só olhando. Pode também criar mais estrelas”.

E por alguma razão, José lembrou a primeira história que ouviu, a mais antiga de todas, sobre a luz e a escuridão. “No começo, havia s trevas”, pensou. E desde então, ficou determinado a fazer algo para que o território escuro do céu fosse semeado com novas estrelas. Para isso, teria que fazer da própria vida uma história a ser lembrada, contada e recontada.

(Adaptação do texto de L.F.Rieseberg)

(Disponível em: < <http://www.rieseberg.com/2014/03/as-estrelas-lfrieberg.html> > Acesso em: 09 mar. 2018)

### **Questões semântico-lexicais**

- a. Como se chamam as frutas menores que a laranja, que se descascam com a mão e, normalmente, deixam um cheiro na mão? Como elas são?
- b. Como se chama aquela raiz branca por dentro, coberta por uma casca marrom, que se cozinha para comer?
- c. Como se chama aquela por o de comida que compramos, normalmente num recipiente laminado ou plástico, num restaurante ou lugares afins quando não temos tempo de cozinhar?
- d. Quando você vai para a sua ocupação diária, da qual você obtém sua renda, você diz que vai para ...?
- e. Você nota outras palavras que diferenciam a fala baiana da paulista? Quais?

### **Percepção/avaliação/identidade baiana (aprox. 5 min)**

Objetivos: saber como os migrantes avaliam a percepção o que os paulistas têm a respeito da identidade baiana do migrante; obter informações sobre a avaliação que os baianos fazem quando o migrante volta à Bahia.

41. Você falou que já foi pra (X, Y, Z). Quando você foi pra esses lugares, as pessoas percebiam que você era baiano? (Se sim) como elas percebiam?
42. Você gosta de ser identificado como baiano?
43. Já sofreu algum preconceito por ser baiano?
44. Olhando pra mim, você diria que eu sou de onde? Por quê?

- 
45. Quando você conhece alguém, você percebe se a pessoa é baiana?
- a. (Se sim) como você percebe? (Se o participante mencionar o modo de falar, seguir nessa linha e tentar conseguir informações mais precisas. É normal as pessoas não saberem definir o porquê, mas devemos tentar obter descrições mais detalhadas).
  - b. (Se não) Quando você ouve uma pessoa falando, por exemplo, você percebe que ela é de lá pelo modo de falar?
46. Qual o sotaque do Brasil que você mais gosta? E tem algum que te irrita? Como é que (o gaúcho / o carioca / o mineiro / o caipira etc.(a depender dos sotaques mencionados) fala(m)?
47. Como é que fala o paulista? Você gosta do modo como as pessoas falam aqui?
48. O que tem de diferente no modo de falar da Bahia e daqui?
49. Quando você vai Bahia, as pessoas dizem que você está falando como paulista? (Se sim) O que eles percebem de diferente?
50. Você gosta de que eles digam que você está falando como paulista? Por quê?
51. Quem são as 10 pessoas com as quais você mais conversa?
52. Dessas 10 pessoas, há baianos? Quem são?

## Apêndice - B

### Questionário *Modus vivendi*

1. Numa escala de 0 a 10, quanto você se considera baiano?
2. Numa escala de 0 a 10, quanto você já se considera paulista?
3. Numa escala de 0 a 10, quanto você já se considera nordestino?
4. Numa escala de 0 a 10, quanto você acha que baiano é diferente de paulista?
5. Numa escala de 0 a 10, quanto você acha que seu jeito de falar é diferente das pessoas que nasceram em São Paulo?
6. Numa escala de 0 a 10, quanto você acha que baiano sofre preconceito em São Paulo?
7. A maioria dos seus amigos aqui em São Paulo é da Bahia?
  - (a) nenhum amigo é;
  - (b) poucos são;
  - (c) metade é;
  - (d) a maioria é.
8. Você tem vizinhos baianos?
  - (a) não;
  - (b) sim, alguns;
  - (c) sim, vários.
9. As pessoas com quem você trabalha são baianas?
  - (a) não;
  - (b) sim, alguns;
  - (c) sim, vários.
10. As pessoas com quem você trabalha são de outro lugar do Nordeste que não seja da Bahia?
  - (a) não;
  - (b) sim, alguns;

(c) sim, vários.

11. Com que frequência você come cuscuz?

- (a) nunca, porque não gosto;
- (b) quase nunca;
- (c) pelo menos uma vez por mês;
- (d) pelo menos uma vez por semana;
- (e) todo dia ou quase todo dia.

12. Com que frequência você come farinha de mandioca?

- (a) nunca, porque não gosto;
- (b) quase nunca;
- (c) pelo menos uma vez por mês;
- (d) pelo menos uma vez por semana;
- (e) todo dia ou quase todo dia.

13. Com que frequência você come carne do sol?

- (a) nunca, porque não gosto;
- (b) quase nunca;
- (c) pelo menos uma vez por mês;
- (d) pelo menos uma vez por semana;
- (e) todo dia ou quase todo dia.

14. Com que frequência você come acarajé?

- (a) nunca, porque não gosto;
- (b) quase nunca;
- (c) pelo menos uma vez por mês;
- (d) pelo menos uma vez por semana;
- (e) todo dia ou quase todo dia.

15. Com que frequência você forró?

- 
- (a) nunca, porque não gosto;
- (b) quase nunca;
- (c) às vezes;
- (d) quase sempre.
16. Com que frequência você axé?
- (a) nunca, porque não gosto;
- (b) quase nunca;
- (c) às vezes;
- (d) quase sempre.
17. Qual dos temperos você mais gosta de usar em suas comidas?
- (a) coentro;
- (b) salsa;
- (c) nenhum dos dois.
18. Com que frequência você fala com seus parentes e amigos que ficaram na Bahia?
- (a) nunca ou quase nunca;
- (b) às vezes;
- (c) quase sempre.
19. Com que frequência você vai visitar seus amigos na Bahia?
- (a) nunca ou quase nunca;
- (b) com relativa frequência;
- (c) uma vez por ano.
20. Há alguma tradição baiana que você faz questão de manter em São Paulo?

## Apêndice - C

### Questionário *on-line* para coletar os dados dos itens lexicais

O objetivo desta pesquisa é analisar a língua em situações cotidianas. Não há respostas certas ou erradas. Você pode responder do modo como acha que as sentenças soam mais naturais. Os dados coletados serão utilizados em análises estatísticas, como na construção de gráficos, por exemplo, e serão reportados em trabalhos científicos. Sua identificação não será divulgada. Assim, ao passar para a próxima seção, você concorda em participar deste estudo. Desde já, agradecemos a sua participação.

#### Informações pessoais

1. Nome:
2. Sexo/Gênero:  
 Masculino  
 Feminino
3. Idade:
4. Nível de escolaridade:  
 Ensino Fundamental  
 Ensino Médio  
 Ensino Superior
5. Cidade e estado onde você mora:  
Ex.: João Pessoa/PB
6. Cidade e estado onde você nasceu:  
Ex.: João Pessoa/PB
7. Cidade e estado onde seu pai nasceu:  
Ex.: João Pessoa/PB
8. Cidade e estado onde sua mãe nasceu:  
Ex.: João Pessoa/PB

#### Cenas do dia a dia

As cenas adiante têm o objetivo de simular uma situação real de fala. Você deve marcar a opção que soaria mais natural para o modo como você diria as sentenças.

1. Você foi ao horto comprar algumas plantas e ficou encantado (a) por uma delas, mas não lembrava o seu nome. Como explicaria como é a planta?  
 Tem uma planta feito um coqueiro que eu não sei o nome.  
 Tem uma planta feito um coqueiro cujo nome eu não sei.
  
2. Você vai emprestar um livro ao seu amigo. Que recomendação você daria?  
 Vou te emprestar, mas toma cuidado.  
 Vou lhe emprestar, mas toma cuidado.
  
3. O pai percebe que o filho está mexendo no celular em vez de fazer as atividades escolares. Se você fosse o pai, o que diria para seu filho?  
 Faça as atividades ou vou tomar seu celular.  
 Faz as atividades ou vou tomar seu celular.
  
4. Você está comprando frutas na feira e sente falta de uma fruta específica na barraca. Como perguntaria ao vendedor?  
 Você não tem mexerica?  
 Você não tem tangerina?
  
5. Seu namorado/namorada pergunta que horas vão ao cinema. Como responderia?  
 Vamos às 9h.  
 Iremos às 9h.
  
6. Você é um(a) professor(a) que solicitou aos alunos uma maquete como atividade avaliativa da unidade, mas um aluno esqueceu de trazê-la.  
 Traz a maquete na próxima semana ou vai ficar com zero.  
 Traga a maquete na próxima semana ou vai ficar com zero.
  
7. Você está escolhendo um presente para sua amiga e fica em dúvida se compra uma calça jeans ou uma saia. Como perguntaria?  
 Amiga, você prefere usar calça jeans ou saia?  
 Amiga, tu prefere usar calça jeans ou saia?

8. Você foi ao salão cortar as pontas do cabelo, mas a cabeleireira é famosa por cortar mais do que as clientes pedem. Como você pediria para ela não fazer isso com seu cabelo?
- Não corte muito, apenas as pontas.
  - Não corta muito, apenas as pontas.
9. Você está prestes a viajar e não está com tempo de arrumar a sua mala. Como você diria para sua filha fazer isso por você?
- Filha, por favor, arrume minha mala.
  - Filha, por favor, arruma minha mala.
10. Você se esqueceu de comprar os ovos para fazer o bolo e lembra que seu filho está na rua. Ao ligar pra ele, como você pediria pra ele trazer?
- Quando vier pra casa, compre ovos pro bolo.
  - Quando vier pra casa, compra ovos pro bolo.
11. Você levou seu computador para formatar, mas só poderá pegar dois dias depois. Você está com receio de o funcionário se esquecer de fazer o serviço. Como você diria para ele?
- Por favor, não se esqueça de formatar o meu computador.
  - Por favor, não se esquece de formatar o meu computador.
12. Uma amiga acaba de realizar um sonho: está grávida. O que você diria para ela?
- Vou lhe dizer que estou muito feliz.
  - Vou dizer-lhe que estou muito feliz.
13. Seu filho vê você chegar do supermercado e pergunta o nome de uma raiz que está dentro da sacola. Como você responderia?
- Isso é aipim.
  - Isso é mandioca.
14. Você está ensinando seu irmão a conectar a TV pela primeira vez. Como você o instruiria?

- Não liga a TV na tomada antes de montar ela.
- Não ligue a TV na tomada antes de montar ela.

15. O telefone da sua casa tocou e estão procurando por seu pai. Como você responderia?
- Ele está no trabalho.
  - Ele está no serviço.

16. Maria está nervosa, pois vai fazer um prato especial pra o maridão para comemorar o primeiro ano de casados. Se você fosse amiga de Maria, o que lhe diria?
- Vai ser lindo! Não fique nervosa!
  - Vai ser lindo! Não fica nervosa!

17. Você está estudando com seu filho para uma prova amanhã. Como você diria para ele não se esquecer de preencher o cabeçalho?
- Não deixe o cabeçalho em branco.
  - Não deixa o cabeçalho em branco.

18. Você mora com duas amigas e hoje é o seu dia de cozinhar, mas não está com disposição para ir para cozinha, então resolve comprar comida. Como pediria no restaurante?
- Eu quero três quentinhas.
  - Eu quero três marmitas.

19. Você reuniu toda a família em confraternização por ter passado no vestibular. Mas um dos seus tios não sabe em qual universidade você vai estudar. O que diria para ele?
- Vou estudar na Unicamp.
  - Irei estudar na Unicamp.

20. Sua amiga tirou carteira de habilitação e está prestes a viajar, mas tem muito medo de dirigir em uma BR. Que conselho lhe daria?

- 
- Pense que está em uma via calma e tudo vai ficar bem.
- Pensa que está em uma via calma e tudo vai ficar bem.
21. Sua amiga se preocupa muito com a opinião das outras pessoas e, por isso, evita fazer coisas que a sociedade condene. Que conselho você daria para ela?
- A vida é curta! Dá menos importância pra essas pessoas!
- A vida é curta! Dê menos importância pra essas pessoas!
22. Sua amiga te chama para passear, mas você não está a fim. Como diria isso para ela?
- Amiga, não estou a fim não.
- Amiga, não estou a fim.
23. Você está aplicando uma prova de matemática e percebe que um aluno está bastante agitado em sua carteira, tentando olhar para prova do colega ao lado. O que lhe diria?
- Não olhe a prova do colega pra que você não seja reprovado!
- Não olha a prova do colega pra que você não seja reprovado!
24. Algum comentário sobre o questionário?