

Cristina dos Santos Carvalho
Silvana Silva de Farias Araújo
Valter de Carvalho Dias
(Organização)

A SOCIOLINGUÍSTICA NO NORDESTE

Diferentes abordagens

PÁDIPALAVRA

CONSELHO EDITORIAL:
Alexandre Cadilhe [UFJF]
Ana Cristina Ostermann [Unisinos/CNPq]
Ana Elisa Ribeiro [CEFET-MG]
Carlos Alberto Faraco [UFPR]
Cleber Ataíde [UFRPE]
Clécio Bunzen [UFPE]
Francisco Eduardo Vieira [UFPB]
Irandé Antunes [UFPE]
José Ribamar Lopes Batista Júnior [LPT-CTF/UFPI]
Luiz Gonzaga Godoi Trigo [EACH-USP]
Márcia Mendonça [IEL-UNICAMP]
Marcos Marcionilo [editor]
Vera Menezes [UFMG]

Cristina dos Santos Carvalho
Silvana Silva de Farias Araújo
Valter de Carvalho Dias
(Organização)

**A SOCIOLINGUÍSTICA
NO NORDESTE**
Diferentes abordagens

CAPA E DIAGRAMAÇÃO: Telma Custódio

**CIP-BRASIL. CATALOGAÇÃO NA PUBLICAÇÃO
SINDICATO NACIONAL DOS EDITORES DE LIVROS, RJ**

S662

A sociolinguística no nordeste [recurso eletrônico] : diferentes abordagens / organização
Cristina dos Santos Carvalho, Silvana Silva de Farias Araújo, Valter de Carvalho Dias. -
1. ed. - São Paulo : Pá de Palavra, 2023.
recurso digital ; 3 MB

Formato: epub
Requisitos do sistema: adobe digital editions
Modo de acesso: world wide web
Inclui bibliografia
ISBN 978-85-68326-46-6 (recurso eletrônico)

1. Sociolinguística - Brasil, Nordeste. 2. Livros eletrônicos. I. Carvalho, Cristina
dos Santos. II. Araújo, Silvana Silva de Farias. III. Dias, Valter de Carvalho. IV. Título.

23-83579

CDD: 306.44
CDU: 81'27



Gabriela Faray Ferreira Lopes - Bibliotecária - CRB-7/6643

Direitos reservados à
PÁ DE PALAVRA
[O selo de autopublicação da **Parábola Editorial**]
Rua Dr. Mário Vicente, 394 - Ipiranga
04270-000 São Paulo, SP
pabx: [11] 5061-9262
home page: www.padepalavra.com.br
e-mail: producao@padepalavra.com.br

Todos os direitos reservados. Nenhuma parte desta obra pode ser reproduzida ou transmitida por qualquer forma e/ou quaisquer meios (eletrônico ou mecânico, incluindo fotocópia e gravação) ou arquivada em qualquer sistema ou banco de dados sem permissão por escrito da editora.

ISBN: 978-85-68326-46-6

© da edição: Pá de Palavra, São Paulo, maio de 2023.

SUMÁRIO

PREFÁCIO.....	7
APRESENTAÇÃO.....	11

PARTE I **VISÃO GERAL SOBRE ESTUDOS SOCIOLINGÜÍSTICOS NO NORDESTE**

PANORAMA DA SOCIOLINGÜÍSTICA NO NORDESTE: CONTRIBUIÇÕES E PROSPECTIVAS.....	16
<i>Claudia Roberta Tavares Silva</i>	

PARTE II **ESTUDOS DE PRODUÇÃO E PERCEPÇÃO SOCIOLINGÜÍSTICAS**

CONCORDÂNCIA VERBAL COM AS FORMAS PRONOMINAIS <i>NÓS</i> E <i>A GENTE</i> : O EFEITO DA VARIÁVEL TEMPO VERBAL SOBRE A MANUTENÇÃO DA FORMA <i>NÓS</i> COM CONCORDÂNCIA PADRÃO.....	37
<i>Hebe Macedo de Carvalho e Hugo Leonardo Gomes dos Santos</i>	
VARIAÇÃO <i>TU</i> E <i>VOCÊ</i> NA POSIÇÃO DE SUJEITO NA FALA ALAGOANA.....	54
<i>Elyne Giselle de Santana Lima Aguiar Vitória</i>	
A PERCEPÇÃO SOCIOLINGÜÍSTICA DO SUBJUNTIVO E DO INDICATIVO.....	66
<i>Wendel Silva dos Santos</i>	

PARTE III **ESTUDOS SOCIOLINGÜÍSTICOS EM INTERFACES**

SOCIOFUNCIONALISMO: AINDA UMA CONVERSA NA DIFERENÇA?.....	83
<i>Camilo Rosa Silva</i> <i>Cleber Tourinho de Santana</i>	
ABORDAGEM SOCIOFUNCIONAL DE <i>VOCÊ</i> E <i>CÊ</i> NA FALA POPULAR DE SALVADOR - BA.....	99
<i>Rosane Bispo dos Santos e Cristina dos Santos Carvalho</i>	

PESQUISAS EM SOCIOLINGÜÍSTICA HISTÓRICA NA BAHIA: ALGUNS PERCURSOS E POSSIBILIDADES	123
<i>Huda da Silva Santiago e Rosana Carvalho Brito</i>	
ASSENTOS DE CASAMENTOS DA FREGUESIA DE SANTO ANTÔNIO DO URUBU DE BAIXO DO RIO SÃO FRANCISCO DO SÉCULO XVIII: ANÁLISE DA HABILIDADE DOS <i>SCRIPTORES</i>	136
<i>Lécio Barbosa de Assis; Jorge Augusto Alves da Silva e Vera Pacheco</i>	
SOBRE OS ORGANIZADORES.....	161
SOBRE OS AUTORES.....	162

PREFÁCIO

Como idealizadora do Encontro de Sociolinguística, sou, ao mesmo tempo, uma testemunha do seu legado.

Estávamos em 2011 e nós, professores/pesquisadores da área da sociolinguística do Programa de Pós-Graduação em Estudo de Linguagens (PPGEL) da Universidade do Estado da Bahia (UNEB) e participantes do Grupo de Pesquisa Fala e Contexto no Português Brasileiro (GConPort)¹, sentíamos a necessidade de promover algo que pudesse congregiar nossos mestres, mestrandos e pesquisadores da área da sociolinguística da UNEB e de outras instituições — locais e de outras regiões baianas —, para que houvesse o diálogo, o debate e a divulgação de trabalhos na área, realizados ou em andamento, bem como a definição de pesquisas de interesse geral e específico para cada região.

A realização de encontros locais e regionais se traduz na oportunidade de contato entre pesquisadores e no desenvolvimento de redes de pesquisa regionais, de início, em ambientes pouco atendidos pelos grandes eventos. É nesse contexto que se insere a série de Encontros de Sociolinguística que, em 2021, realizou a sua décima primeira edição. O 1º Encontro de Sociolinguística, realizado em apenas um dia, em 01 de dezembro de 2011, explorou o tema *Varição ~ História ~ Educação*. A edição inicial do evento contou apenas com duas conferências: uma de abertura, sobre a variação e formação do português brasileiro, com a Profa. Dra. Maria Marta Scherre, da Universidade Federal do Espírito Santo (UFES); e outra de encerramento, sobre o ensino de português, com a Profa. Dra. Stella Bortoni-Ricardo, da Universidade de Brasília (UnB), além de duas mesas-redondas. Esse encontro constou oficialmente, no seu início, como produção do projeto de pesquisa *Linguagem na cidade: uma fotografia sócio-discursiva de Salvador*, financiado pela FAPESB, por mim coordenado, o que possibilitou a publicação de alguns textos apresentados nessa primeira edição no livro *Linguagem na cidade: estudos sócio-discursivos* (Quarteto Editora, 2012), organizado por mim, Norma Lopes, Gilberto Sobral e Jaciara de Oliveira, professores da UNEB.

Esse primeiro Encontro de Sociolinguística causou um efeito muito positivo: marcou o início das discussões sobre a pesquisa sociolinguística entre os participan-

¹ Em 2023, houve uma mudança do nome do GConPort para Grupo de pesquisa Usos e Contextos da Língua Portuguesa.

tes, que reuniu, além de pesquisadores e estudantes dos diversos níveis da UNEB, pesquisadores de outras instituições universitárias baianas e de outros estados do Nordeste, da Secretaria de Educação do Estado da Bahia (SEC). É importante registrar o grande número de participantes da Universidade Estadual de Feira de Santana (UEFS), que se tornou, a partir desse momento, a primeira das universidades parceiras na realização do Encontro.

Os 11 encontros já realizados ocorreram ininterruptamente a cada ano, de forma presencial em 9 deles e de forma remota nos dois últimos, diante da pandemia, e exploraram diversos temas, nas diversas vertentes e interfaces dos estudos da relação entre língua e sociedade. No segundo encontro, com o tema *Linguagem na cidade: uma fotografia sócio-discursiva de Salvador*, associado diretamente ao tema do projeto financiado pela FAPESB, essa edição originou a publicação de um livro, *Sociolinguística: estudos da variação, da mudança e da sócio-história do português brasileiro; sociolinguística paramétrica; sociofuncionalismo*, organizado pelas professoras doutoras da UNEB Norma Lopes, Lígia Bulhões e Cristina Carvalho e publicado pela UEFS Editora, em 2013. A partir desse evento, o Encontro se ampliou de um para dois dias e, além das conferências e mesas-redondas, contou com apresentação, em forma de comunicação, de investigações sociolinguísticas de pesquisadores dos distintos níveis (doutores, mestres, doutorandos, mestrandos e bolsistas de iniciação científica) de universidades diversas.

O III Encontro centralizou-se no tema *A pesquisa em Salvador*. Ampliando o número de participações de diferentes universidades, o evento atraiu o interesse de pesquisadores da Universidade Federal da Bahia (UFBA), da Universidade Estadual do Sudoeste Baiano (UESB), do Instituto Federal de Educação, Ciência e Tecnologia da Bahia (IFBA), além da UEFS e da Secretaria de Educação do Estado da Bahia, instituições presentes nas edições anteriores. O livro *Salvador, sob o olhar da Sociolinguística*, organizado pelas professoras doutoras da UNEB Norma Lopes, Lígia Pellon e Lúcia Parcerro, publicou textos dos conferencistas e de apresentações em mesas-redondas (UEFS Editora, 2013). Acoplados a esse livro, apresentam-se, em formato de CD, os *Anais* do evento, com textos de comunicações aprovados pelo conselho editorial.

O IV Encontro de Sociolinguística abordou o tema *Diferentes olhares sobre o português brasileiro* e contou com apresentações de pesquisas de fora do Estado da Bahia. A partir dessa edição, em 2014, além da UEFS, a Universidade Federal de Sergipe (UFS) e a Universidade Federal de Minas Gerais (UFMG) passaram a integrar, juntamente com a UNEB, as universidades parceiras, com a presença constante das professoras doutoras Josane Oliveira (UEFS), Raquel Freitag (UFS) e Jânia Ramos, do Programa de Pós-Graduação em Linguística (POSLIN) da Universidade Federal de Minas Gerais (UFMG), esta última motivada pela aprovação do Projeto *Linguagem, Sociedade e Discurso: uma parceria entre a UNEB e a UFMG*, financiado pela FAPESB/CAPES e coordenado no PPGEL pelo Prof. Dr. Gilberto Sobral. Nesse momento, o Encontro já era um evento realizado conjuntamente por 4 instituições, a UNEB, a UEFS, a UFS e

a UFMG. O livro *Diferentes olhares sobre o português brasileiro*, organizado por Norma Lopes (UNEB), Jânia Ramos (UFMG) e Josane Oliveira (UEFS), reuniu textos de conferências, mesas-redondas e comunicações aprovadas pelo conselho editorial do evento.

Os Encontros de Sociolinguística passaram a ser itinerantes a partir de 2015, sendo realizados a cada dois anos na UNEB e, nos anos intermediários, as universidades parceiras passaram a coordenar e assumir o evento. Iniciando o caráter itinerante, o V Encontro ocorreu na UEFS, promovido pelo PPGEL/UEFS, e explorou o tema central *Diálogos entre Brasil e África*. Motivada pelo tema dessa edição, houve a parceria com a Universidade de Integração Internacional da Lusofonia Afro-Brasileira (UNILAB). O evento produziu uma publicação em número Especial da *Revista Interdisciplinar* (UFS, 2016, vol. 24 e 25).

O Nordeste foi o foco do VI Encontro, ocorrido na UNEB em 2016, com a participação de pesquisadores, mestres e pós-graduandos da Universidade Federal da Bahia, além das universidades parceiras e outras instituições. A Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB) atuou como órgão de fomento do evento, inserido, no Projeto de pesquisa *Linguagem, Sociedade e Discurso*, uma parceria entre a UNEB e a UFMG. Houve dois *e-books* publicados pela editora Blucher, resultantes desse encontro: (i) *Estudos sobre o português do Nordeste: língua, lugar e sociedade*, organizado por Norma Lopes (UNEB), Josane Oliveira (UEFS) e Lúcia Parceró (UNEB) (2017); e (ii) *A fala nordestina: entre a sociolinguística e a dialetologia*, organizado por Norma Lopes (UNEB), Silvana Araújo (UEFS) e Raquel Freitag (UFS) (2018).

O VII Encontro ocorreu na UFS, em 2017. Tomando como tema *Redes e Contatos*, a edição teve ampla participação de pesquisadores e pós-graduandos de Sergipe e de outros estados nordestinos, e passou a ter um caráter regional, dando enfoque principal ao Nordeste. A publicação decorrente desse evento deu-se na Revista *A Cor das letras* (UEFS), vol. 19 (2018), número especial organizado pela Profa. Dra. Raquel Freitag (UFS): *VII Encontro de Sociolinguística: redes e contatos*.

O VIII Encontro voltou ao PPGEL/UNEB, em 2018, quando passamos a atuar novamente na presidência da comissão organizadora. Explorou-se o tema *Variação e ensino: unidade e diversidade*, e a publicação resultante desse evento ocorreu na Revista *Tabuleiro de Letras*, vol. 12, número especial (UNEB), no mesmo ano.

O IX Encontro foi sediado pelo Instituto de Letras da Universidade Federal da Bahia (UFBA), em 2019. Teve como tema *Sociolinguística: quebrando tabus e inovando na escola*, homenageando a Profa. Dra. Jacyra Mota (UFBA). A publicação decorrente desse evento deu-se na Revista *A Cor das letras* (UEFS), vol. 22, número especial: *Sociolinguística, Dialetologia e Ensino de Língua Portuguesa*, organizado pelos professores doutores Josane Oliveira (UEFS), Clézio Gonçalves (UFOP) e Gredson dos Santos (UFBA), em 2021.

O X Encontro ocorreu de forma remota com transmissão pelo YouTube, em 2020, tendo como tema *Desafios do Isolamento para a Pesquisa Sociolinguística e Interfaces*. Desse encontro, resultou a publicação do livro *Desafios para Pesquisa em Sociolinguística*, orga-

nizado por Raquel Freitag (UFS), Silvana Araújo (UEFS), Valter Dias (IFBA), publicado pela editora Blucher, em 2022.

Em 2021, ocorreu o XI Encontro, ainda de forma remota e com transmissão pelo YouTube, com o tema *A Sociolinguística no Nordeste*. Desse evento, originou-se a publicação do presente livro, *A sociolinguística no Nordeste: diferentes abordagens*, organizado pelos professores doutores Cristina Carvalho (UNEB), Silvana Araújo (UEFS) e Valter Dias (IFBA), sob o selo Pá de Palavra, em 2023.

Os Encontros de Sociolinguística têm possibilitado rico debate interinstitucional entre pesquisadores da área, com envolvimento de profissionais e estudantes dos diversos níveis interessados em vertentes da sociolinguística. Como este texto apresenta, o Encontro de Sociolinguística buscou contribuir com a difusão do conhecimento, a produção da pesquisa e a discussão na área central de interesse. Concluiu afirmando que o empreendimento tem sido bem-sucedido. Como se pode atestar através da larga produção sociolinguística, os eventos se revelaram grande contribuição para a consolidação de pesquisas sociolinguísticas no Nordeste. Nessas onze edições, com a publicação de textos de conferências e de mesas-redondas, além de comunicações apresentadas e constantes debates, motiva-se, divulga-se e produz-se ciência, mesmo em momentos de extrema dificuldade, como na pandemia que assolou o mundo e que teima ainda em nos assustar.

Norma da Silva Lopes

APRESENTAÇÃO

O presente livro reúne textos de pesquisadora/es que participaram do XI Encontro de Sociolinguística, realizado no início de dezembro de 2021. Nessa edição, o evento teve como tema *A Sociolinguística no Nordeste* e, nessa perspectiva, buscou “ampliar a rede de pesquisas e pesquisadoras/es em Sociolinguística e dar maior destaque às produções acadêmicas na região Nordeste” (DIAS, 2021)².

Sendo assim, com esta coletânea, socializamos resultados de trabalhos sociolinguísticos desenvolvidos, sob diferentes abordagens, por docentes e discentes de programas de pós-graduação de instituições de ensino superior de estados — Alagoas, Bahia, Ceará, Maranhão, Pernambuco e Paraíba — da região Nordeste do nosso país. Tais estudos, além de mostrar o rigor metodológico das investigações empreendidas, constituem evidências científicas contra o mito da homogeneidade linguística no Nordeste, tão propalado pelas mídias televisiva e publicitária.

Nesta obra, a partir dos textos compilados, inicialmente, fornecemos uma visão geral sobre a sociolinguística no Nordeste (capítulo 1) e, posteriormente, mostramos estudos que enfocam produção e/ou percepção sociolinguísticas (capítulos 2, 3 e 4) e que representam uma interlocução entre a vertente variacionista e outros modelos teóricos, tais como o funcionalismo norte-americano (capítulos 5 e 6) e a linguística histórica (capítulos 7 e 8).

No capítulo “Panorama da sociolinguística no Nordeste: contribuições e perspectivas”, Claudia Roberta Tavares Silva discute o processo de consolidação dos estudos sociolinguísticos desenvolvidos na região, tendo em vista seus contributos e perspectivas advindos de uma agenda de trabalho comprometida desde sempre em compreender, sob a ótica da intrínseca relação língua-sociedade, a constituição dos diferentes falares nordestinos a partir da descrição e sistematização da variação linguística que lhes é inerente. A autora, depois de fazer uma retrospectiva histórica da sociolinguística no Brasil, apresenta um mapeamento sociolinguístico do Nordeste, citando inúmeras investigações já realizadas nessa região sobre os usos linguísticos produzidos em diferentes modalidades da língua e sobre as crenças e atitudes lin-

² DIAS, Valter de Carvalho. Apresentação. In: ENCONTRO DE SOCIOLINGUÍSTICA. 11, 2021. *Anais eletrônicos*. Salvador, 2021. Disponível em: <https://doity.com.br/anais/xiencontrodesociolinguistica>. Acesso em: 9 nov. 2022.

guísticas dos falantes diante desses usos. Por fim, sinaliza novas reflexões/prospecções que devem ser incluídas em uma agenda de trabalho que vise à compreensão do mosaico sociolinguístico do Nordeste.

Em “Concordância verbal com as formas pronominais *nós* e *a gente*: o efeito da variável tempo verbal sobre a manutenção da forma *nós* com concordância padrão”, Hebe Macedo de Carvalho e Hugo Leonardo Gomes dos Santos analisam a atuação da variável tempo e tipo de paradigma verbal na variação da concordância verbal com as formas pronominais de primeira pessoa do plural, na fala de Fortaleza da década de 1980. As ocorrências de *nós* e *a gente* foram coletadas de dezoito entrevistas sociolinguísticas do projeto Dialeto Sociais Cearenses. Os resultados quantitativos obtidos a partir do *software* R Studio (R CORE TEAM, 2020) mostram que: há, na amostra, um maior emprego de *a gente* do que *nós*; prevalecem as estruturas de concordância padrão, sendo a estrutura *a gente* com *-mos* rara na base de dados e *nós* sem *-mos* pouco frequente; na análise apenas das estruturas de concordância padrão, se verifica que, apesar de todos os tempos verbais favorecerem o uso de *a gente* sem *-mos*, o pretérito perfeito do indicativo se configura como um ambiente morfossintático de competição entre as variantes.

Em “Variação *tu* e *você* na posição de sujeito na fala alagoana”, Elyne Giselle de Santana Lima Aguiar Vitória descreve a representação da 2ª pessoa do singular na função de sujeito como fenômeno variável sob a perspectiva da produção e percepção sociolinguísticas. A autora mostra como *tu* com concordância, *tu* sem concordância, *você*, *ocê* e *cê* têm se comportado na variedade alagoana. Constata que, em Alagoas, há o registro dessas cinco variantes, com *você* e *tu* sem concordância sendo mais frequentes. Também evidencia que, enquanto *você* é a variante coringa, ocorrendo em qualquer situação comunicativa, *tu* sem concordância exprime simetria, proximidade e familiaridade, associado a situações mais informais, e *tu* com concordância exprime assimetria e maior distanciamento entre os falantes, sendo associado a situações comunicativas mais formais.

No texto “A percepção sociolinguística do subjuntivo e do indicativo”, Wendel Silva dos Santos, a partir da teoria da variação e mudança, propõe uma análise de percepção estimulada por orações subordinadas no modo indicativo ou subjuntivo, organizadas de acordo com a técnica *matched-guise* (LAMBERT *et al.*, 1960; CAMPBELL-KIBLER, 2006). Os estímulos (enunciados por 4 falantes — dois ludovicenses e dois paulistanos) constituem-se de orações subordinadas no modo indicativo ou subjuntivo: adverbiais com *embora* ou *talvez* e substantivas introduzidas por *querer* ou *acreditar*. No geral, as análises de regressão linear (R CORE TEAM, 2021) apontam que: os falantes foram percebidos como mais inteligentes quando ouvidos no subjuntivo; os ludovicenses foram avaliados como mais formais, mas os falantes paulistanos foram percebidos como mais antipáticos; ludovicenses foram percebidos como menos escolarizados quando ouvidos com o indicativo, enquanto não houve diferença para os falantes paulistanos.

Em “Sociofuncionalismo: ainda uma conversa na diferença?”, Camilo Rosa Silva e Cleber Tourinho propõem uma reflexão teórico-metodológica sobre a perspectiva dos estudos sociofuncionalistas, destacando pontos convergentes ou aproximativos entre a sociolinguística variacionista e o funcionalismo norte-americano. Para isso, inicialmente, fazem uma revisão de literatura na qual focam a sociolinguística e o funcionalismo, seus objetos de estudo e seus pressupostos mais relevantes. Em seguida, explicitam a base do construto teórico do sociofuncionalismo, considerando seus avanços e possibilidades. Ainda citam, como modelo de aplicação dos princípios sociofuncionais, um estudo de caso e apresentam uma breve análise da qualidade e alcance do referido trabalho para presentes e futuras investigações. Concluem que o sociofuncionalismo representa uma possibilidade viável de análise da variação e da mudança, contribuindo com resultados robustos e refinados para as pesquisas dessa área dos estudos linguísticos.

No capítulo “Abordagem sociofuncional de *você* e *cê* na fala popular de Salvador-BA”, Rosane Bispo dos Santos e Cristina dos Santos Carvalho analisam os usos de duas formas de expressão da segunda pessoa do singular — *você* e *cê* — na função sintática de sujeito, em posição pré-verbal. Do ponto de vista teórico, fundamentam-se em postulados do sociofuncionalismo, conciliando as perspectivas de variação e gramaticalização. Do ponto de vista metodológico, examinam, em viés qualiquantitativo, dados extraídos de dezesseis entrevistas sociolinguísticas integrantes do acervo do Programa de Estudos do Português Popular de Salvador. As autoras apresentam apenas resultados relativos às variáveis linguísticas consideradas relevantes pelo programa GoldVarb X para o fenômeno variável estudado: natureza semântico-funcional, paralelismo formal e tipo semântico verbal. Quanto à intersecção variação-gramaticalização, constatam que há, na fala soteropolitana, de um lado, uma estratificação entre *você* e *cê* e, do outro, uma tendência à gramaticalização de *cê* como clítico.

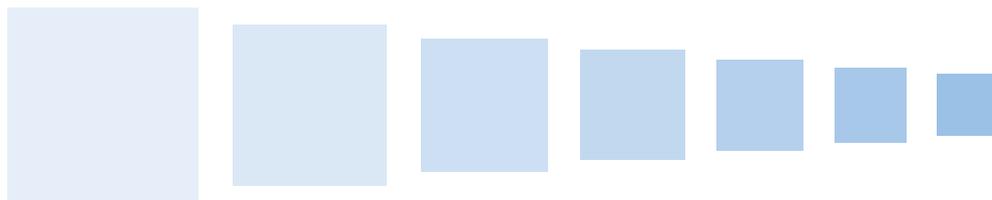
Em “Pesquisas em sociolinguística histórica na Bahia: alguns percursos e possibilidades”, Huda da Silva Santiago e Rosana Carvalho Brito discutem alguns enfrentamentos e avanços conquistados no que se refere ao tratamento metodológico dos *corpora* utilizados em pesquisas realizadas sob a perspectiva da sociolinguística histórica. Inicialmente, as autoras apresentam uma breve revisão bibliográfica dos estudos em sociolinguística histórica e os desafios enfrentados por essa disciplina, por exemplo, de localização e constituição de *corpora* escritos que sejam representativos, de identificação do perfil dos escreventes e de estabelecimento dos fatores sociais relacionados à comunidade da qual faziam parte. Posteriormente, comentam alguns estudos desenvolvidos por pesquisadores baianos, que seguem o viés da sociolinguística histórica, embora não assumam explicitamente, em seus textos, a filiação a esse modelo.

Em “Assentos de casamentos da Freguesia de Santo Antônio do Urubu de Baixo do Rio São Francisco do século XVIII: análise da habilidade dos *scriptores*”, Lécio Barbosa de Assis, Jorge Augusto Alves da Silva e Vera Pacheco, norteando-se pelo suporte

da leitura paleográfica e da sociolinguística histórica, objetivam analisar a habilidade dos *scriptores*, através dos aspectos físico-gráficos, gráficos e fonéticos, como o traçado das letras, substituições de segmentos, segmentação vocabular, acentamento de vogais, substituições consonantais, formas etimologizadas, entre outros. Para tanto, examinam um *corpus* de 231 assentos de casamentos retirados do primeiro livro da *Freguesia de Santo Antônio do Urubu de baixo do Rio São Francisco*. Os resultados revelam que, apesar de os *scriptores* fazerem parte de uma elite com maior nível de instrução, uma vez que as habilidades para ler e escrever seriam fundamentais para desempenhar os trabalhos exigidos em sua paróquia de atuação, a escrita era influenciada pelos conflitos referentes aos períodos históricos da ortografia da língua portuguesa. Os autores verificam, ainda, a existência de grafias que podem ser decorrentes da influência da oralidade, especialmente em algumas variações vocálicas, nas posições não acentuada final, inicial absoluta e pretônica.

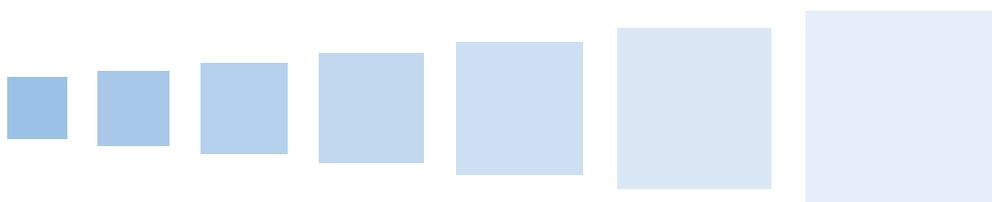
Esperamos que as reflexões aqui apresentadas, a partir de pesquisas sociolinguísticas em diferentes abordagens, possam contribuir para desmitificar a ideia de um Nordeste *uno* e, por conseguinte, para reforçar a visão de uma região plural do ponto de vista (socio)linguístico no que tange aos falares dos seus estados. Acreditamos que tais reflexões sejam de interesse para discentes e docentes de graduação e pós-graduação das áreas de letras e linguística e áreas afins, no sentido de que, entre outras ações, eles possam atuar como multiplicadores no combate a estereótipos concernentes ao português do Nordeste. Fazemos-lhes, então, um convite à leitura desta obra.

Cristina dos Santos Carvalho
Silvana Silva de Farias Araújo
Valter de Carvalho Dias



PARTE I

VISÃO GERAL SOBRE ESTUDOS SOCIOLINGÜÍSTICOS NO NORDESTE



PANORAMA DA SOCIOLINGUÍSTICA NO NORDESTE: CONTRIBUIÇÕES E PROSPECTIVAS

CLAUDIA ROBERTA TAVARES SILVA

1 PRIMEIRAS PALAVRAS

Por ser toda e qualquer língua intrinsecamente heterogênea e submetida a distintos condicionamentos internos e externos no sentido de que “[p]odemos definir uma *variável sociolinguística* [grifo do autor da citação] como correlacionada com alguma variável não linguística do contexto social: o falante, o interlocutor, o público, o ambiente etc”, (LABOV, [1972] 2008, p. 275), abordaremos, neste capítulo, sob o viés da sociolinguística, como a diversidade linguística na região Nordeste do Brasil vem sendo estudada, tendo em mente dois importantes eixos: (i) a descrição das formas linguísticas variáveis e (ii) a avaliação social dessas formas.

A produção bibliográfica e os trabalhos investigativos realizados revelam que o mito do “português nordestino” não se sustenta e que, portanto, há, na dinâmica ininterrupta língua e contexto social, um rico e complexo mosaico sociolinguístico constituído por diferentes falares quando comparamos os usos linguísticos entre diferentes estados do Nordeste e entre diferentes regiões de um mesmo estado. Para a construção desse mosaico, a compreensão das dimensões sincrônica, diacrônica (“história interna”) e sócio-histórica (“história externa”) torna-se crucial.

Entendendo que a relação língua e contexto social dá-se de forma intrínseca, Labov (1966), ao desenvolver a pesquisa intitulada *The Social Stratification of English in New York City*, introduz, nos estudos linguísticos, uma nova abordagem de compreender essa relação, tomando por base que as formas linguísticas variáveis, longe de serem assistemáticas, apresentam uma distribuição estruturada, atrelada a diferentes graus de estratificação socioeconômica e que, por serem os falantes os agentes de propagação dessas formas, sua avaliação social sobre elas é também objeto de interesse. Diante desses dois aspectos nucleadores da proposta laboviana, Freitag (2016) pontua que os estudos sociolinguísticos feitos no Brasil têm centrado a atenção, em sua grande maioria, na descrição dos usos linguísticos advindos da constituição de amostras e de bancos de dados. Essa abordagem tem suas bases no modelo hegemônico dos estudos sociolinguísticos da primeira onda (FREITAG; MARTINS; TAVARES, 2012),

não sendo, portanto, o foco de interesse, por algumas décadas, os valores sociais atribuídos pelos brasileiros às formas linguísticas variáveis.

Não obstante, mais precisamente no século XXI, observamos um movimento distinto do século XX, na medida em que pesquisas sociolinguísticas no Nordeste, por exemplo, começam a se intensificar no que se refere à temática da avaliação s. Além disso, novos bancos de dados começam a ser construídos a partir de indicadores sociodemográficos mais amplos, o que nos leva a inferir que uma nova fase está emergindo, baseada em uma visão pluridimensional para o tratamento da diversidade linguística no Nordeste.

Para a discussão aqui pretendida, são objetivos norteadores deste capítulo: a) situar e primordialmente discutir o processo de consolidação dos estudos sociolinguísticos desenvolvidos no Nordeste, tendo em mente seus contributos e perspectivas, advindos de um agenda de trabalho comprometida desde sempre em compreender a constituição dos diferentes falares nordestinos; e b) discorrer sobre a construção do mapeamento sociolinguístico dessa região, tomando por base as inegáveis contribuições dos estudos dialetológicos e sociolinguísticos. Para tanto, na seção 2, serão feitas incursões no percurso histórico desses estudos sobre a diversidade linguística no Brasil; na seção 3, algumas considerações serão apresentadas sobre a construção dos bancos de dados no Nordeste; na seção 4, abordaremos a constituição do mosaico sociolinguístico da região, tendo em mente dois eixos contemplados nos estudos: o uso das formas linguísticas e a avaliação social dessas formas e, por fim, apresentaremos, as prospecções advindas do panorama aqui delineado.

2 PERCURSO HISTÓRICO DA SOCIOLINGÜÍSTICA NO BRASIL: BREVES INCURSÕES

2.1 ESTUDOS DIALETOLÓGICOS: CONTRIBUTO PARA A COMPREENSÃO DA DIVERSIDADE LINGÜÍSTICA DO BRASIL

Considerada a “ciência-mãe da Sociolinguística” (FREITAG, 2021, p. 26), a dialetologia

[...] apresenta-se, no curso da história, como uma disciplina que assume por tarefa identificar, descrever e situar os diferentes usos em que a língua se diversifica, conforme a sua distribuição espacial, sociocultural e cronológica. Dois aspectos fundamentais estão, pois, na sua gênese: o reconhecimento das diferenças ou das igualdades que a língua reflete e o estabelecimento das relações entre as diversas manifestações lingüísticas documentadas ou entre elas e a ausência de dados registrados, circunscritos a espaços e realidades pré-fixados” (CARDOSO, 2002a, p. 1).

Três trabalhos sobre a diversidade linguística no Brasil ganham destaque no cenário dialetológico: *O dialeto caipira*, de Amadeu Amaral (1920), *O linguajar carioca*, de Antenor Nascentes (1922) e *A língua do Nordeste: Alagoas e Pernambuco*, de Mário Marroquim (1934). Para Mota e Cardoso (2006), esses estudos instauram a segunda fase

da dialetologia no Brasil (1920-1952), por não estarem circunscritos apenas ao domínio lexical, ao contrário do que se observa na primeira fase (1826-1920), cujo enfoque é a publicação de dicionários e de glossários.

Através do decreto 30.643, de 20 de março de 1952, instaura-se, no Brasil, a terceira fase, marcada pelo incentivo à elaboração do *Atlas Linguístico do Brasil* e, por sua vez, em 1996, inaugura-se a quarta fase, com o início do projeto *Atlas Linguístico do Brasil* (ALiB), coordenado pela Profa. Suzana Alice Marcelino Cardoso (UFBA), cuja meta norteadora é “contribuir para o entendimento da língua portuguesa no Brasil como instrumento social de comunicação diversificado, possuidor de várias normas e usos, mas dotado de uma unidade sistêmica”¹ Essa última fase:

Do ponto de vista metodológico, [...] coincide com a incorporação dos princípios implementados pela Sociolinguística a partir da década de 60 do século passado, abandonando-se a visão monodimensional – monoestratégica, monogeracional, monogenérica, monofásica, etc. – que predominou na geolinguística hoje rotulada de ‘tradicional’” (MOTA; CARDOSO, 2006, p. 21).

Assumindo agora uma visão pluridimensional, a metodologia do ALiB agrega, sob influência da sociolinguística variacionista, iniciada na década de 1960, variáveis sociais à variável diatópica. Sobre essa influência, destaca Brandão (1991, p. 26):

[...] hoje, torna-se imperativo, por exemplo, incluir, entre os critérios de escolha dos indivíduos que servirão de informantes para a formação do *corpus* de um atlas linguístico, variáveis como idade, sexo, nível de instrução, ou mesmo situação socioeconômica, a fim de que se revelem ao máximo as particularidades do sistema dialetal focalizado e se possam melhor conhecer os condicionamentos socioculturais que presidem a distribuição geográfica dos fenômenos linguísticos (BRANDÃO, 1991, p. 26).

Indo nessa direção, seis dimensões são contempladas no ALiB, segundo verificam Isquierdo e Romano (2012, p. 903): a) *diatópica* (250 cidades, incluindo as capitais de Estados — exceto Palmas e Brasília); b) *diageracional*: faixa etária (18 a 30 anos; 50 a 65 anos); c) *diassexual*: feminino e masculino (*total geral*: 1.100 informantes); d) *diastrática*: ensino fundamental (interior e capital) e ensino superior (capital); e) *diafásica*: questionários linguísticos, textos para leitura, discursos semidirigidos; e f) *diarreferencial*: apuração dos fatos metalinguísticos e epilinguísticos.

Em se tratando dos atlas linguísticos produzidos no Nordeste, em particular, verificamos, numa escala temporal, que, desde a década de 60, se deu início a essa produção, conforme mostrado a seguir:

- Atlas Prévio dos Falares Baianos (APFB) (ROSSI; FERREIRA; ISENSEE, 1963);
- Atlas Linguístico da Paraíba (ALPB) (ARAGÃO; BEZERRA DE MENDES, 1984);
- Atlas Linguístico de Sergipe (ALS) (FERREIRA, 1987);
- Atlas Linguístico do Maranhão (AliMA) (RAMOS, 200-);

¹ <http://www.alib.ufba.br/contente/objetivos>

- Atlas Linguístico de Sergipe II (ALS II) (CARDOSO, 2002b);
- Atlas geolinguístico do Litoral Potiguar (ALiPTG) (PEREIRA, 2007);
- Atlas Linguístico da Mata Sul de Pernambuco (ALMASPE) (ALMEIDA, 2009);
- Atlas Linguístico do Ceará (ALECE) (BESSA, 2010);
- Atlas Linguístico do Centro-Oeste Potiguar (SILVA, 2012);
- Atlas Linguístico de Pernambuco (ALiPE) (SÁ, 2013);
- Atlas Linguístico do Estado de Alagoas (ALEAL) (DOIRON, 2017);
- Atlas Linguístico do Litoral Pernambucano (ALITOP)² (coordenador: prof. Edmilson José de Sá);
- Atlas Linguístico do Piauí (AliPI) (coordenador: prof. Juscelino Francisco do Nascimento (UFPI))³.

Em geral, verificamos, a partir do século XXI, como o título de alguns atlas sugere, uma tendência ao refinamento na escolha do espaço geográfico a partir do qual estão distribuídas e descritas as particularidades do sistema dialetal no Nordeste: o espaço geográfico deixa de ser o estado como um todo e passa a abarcar uma região específica dele, como é o caso do ALiPTG, do ALMASPE, do ALITOP e do Atlas Linguístico do Centro-Oeste Potiguar.

2.2 PRIMEIROS CENTROS E GRUPOS DE PESQUISA SOCIOLINGÜÍSTICA

Não se pode negar que o conhecimento sobre a realidade linguística brasileira teve início no âmbito da Dialectologia e atingiu o ápice com a Sociolinguística, mais especificamente, com a sociolinguística variacionista laboviana e os sofisticados métodos de análise estatística, um modelo de análise seguido em centenas de estudos na área (CALLOU, 2010, p. 33).

Após as discussões sociolinguísticas iniciadas na década de 1960 nos EUA, tomando por base os três eixos: a) indissociabilidade da língua com o contexto social; b) a distribuição das variáveis linguísticas, atrelada à sua sistematização no âmbito da estratificação social; e c) a avaliação social dessas variáveis, verificamos que os estudos pioneiros iniciados no Brasil na década de 1970 centraram a atenção sobretudo nos dois primeiros eixos, voltando-se à descrição das formas linguísticas variáveis e seus possíveis condicionamentos linguísticos e extralinguísticos. Paiva e Scherre (1999) verificam que, anterior a essa década, mais precisamente, no ano de 1969, é idealizado pelo professor Nelson Rossi (UFBA) o Projeto de Estudos da Norma Urbana Culta do Brasil (NURC) cujo título já sinaliza para o fato de que essa norma tem uma identidade que lhe é própria, distinguindo-se, por exemplo, da norma urbana culta de Portugal. Tal perspectiva amplia a visão de que estamos diante de uma das variedades do português, indo de encontro à ideia de homogeneidade linguística, um dos pilares dos estudos sociolinguísticos.

² Em construção desde 2019.

³ Em construção desde 2020.

Conforme pontuam Scherre e Roncarati (2008), as discussões em torno da sociolinguística variacionista iniciam no estado do Rio de Janeiro, em 1976, durante uma disciplina ministrada pelo professor Anthony Julius Naro no curso de mestrado da PUC-RJ. Pesquisas pioneiras nessa área foram orientadas por ele em diversos temas, segundo verificam Votre e Roncarati (2008, p. 21-26 apud OLIVEIRA, 2016, p. 483): a) concordância nominal (BRAGA, 1977; SCHERRE, 1978); b) construções relativas (MOLLICA, 1977); c) derivação sufixal (MONTALVO, 1978); d) segmentos nasal e vibrante em final de vocábulos (VOTRE, 1978); e) pronome pessoal de terceira pessoa em posição acusativa (OMENA, 1978); e f) distribuição da vibrante (CALLOU, 1979). Vale mencionarmos que, no ano de 1977, em parceria com a professora Miriam Lemle, o professor Naro coordenou o projeto Competências Básicas do Português, cuja proposta reside na ideia da heterogeneidade linguística:

A língua portuguesa, código comunicativo de uma sociedade complexa e heterogênea, é ela mesma um sistema heterogêneo, uma rede de subsistemas, de variedades dialetais, decorrentes dos diferentes graus de coesão e contato dos diversos subgrupos da comunidade total. Assim sendo, ter plena capacidade de uso da língua portuguesa implica em ter a capacidade de traquejo com mais de uma variedade dialetal (LEMLE; NARO, 1977, p. 1).

A década de 1970, portanto, destaca-se como emblemática na pesquisa sociolinguística, na medida em que surgem os primeiros projetos e grupos de pesquisa (CEZARIO; VOTRE, 2008; PAIVA; SCHERRE, 1999), a saber:

- Grupo do projeto MOBREAL Central (coordenação: profa. Miriam Lemle);
- Grupo do projeto da Norma Urbana Oral Culta do Rio de Janeiro (NURC) (coordenação: prof. Celso Cunha);
- Grupo do Projeto Censo da Variação Linguística no Estado do Rio de Janeiro (Censo) (coordenação: prof. Anthony Naro);
- Programa de Estudos dos Usos da Língua (PEUL) (idealizado pelo prof. Anthony Naro).

Outros grupos e projetos de investigação surgem, por sua vez, em décadas posteriores, em outras regiões brasileiras (PAIVA; SCHERRE, 1999), com o intuito de ampliar a compreensão do mosaico sociolinguístico brasileiro:

1988 – Projeto: Dialeto Social Cearense (coordenação: profa. Claudia Roncarati – UFC);

1990 – Variação Linguística Urbana da Região Sul (VARSUL) (coordenação: prof. Carlos Alberto Faraco – UFPR);

1990 – Programa para a História da Língua Portuguesa (PROHPOR) (coordenação: profa. Rosa Virgínia Mattos e Silva – UFBA);

1991 – A Língua Usada em Alagoas (LUAL) (coordenação: profa. Maria Denilda Moura – UFAL);

1993 – Variação Linguística no Estado da Paraíba (VALPB) (coordenação: prof. Dermeval da Hora – UFPB);

1996 – Projeto: A Língua Portuguesa no Semi-Árido Baiano (coordenação: profa. Norma Lúcia Fernandes de Almeida e profa. Zenaide Oliveira Novais Carneiro – UEFS);
1998 – Projeto para a História do Português Brasileiro (PHPB) (coordenação: prof. Ataliba de Castilho – USP).

Vale referir que, em 1985, se cria, por iniciativa dos professores Jürgen Heye (PU-C-RJ), Sebastião Votre (UFRJ) e Paulino Vandresen (UFSC), o Grupo de Trabalho (GT) de Sociolinguística da ANPOLL (BRANDÃO, 1994). Mapeando os grupos de pesquisa cadastrados no CNPq nessa área, Salomão (2011) contabilizou um total de 48, sendo assim distribuídos por região: a) região Sudeste: 19 grupos; b) região Sul: 10 grupos; c) região Norte: 9 grupos; d) região Nordeste: 8 grupos; e) região Centro-Oeste: 2 grupos. Conforme observamos, a região Sudeste, pioneira nos estudos sociolinguísticos, concentra mais grupos, sendo tímida ainda a formação de novos grupos em outras regiões, com destaque para a região Nordeste, a penúltima em menor quantidade.

Comparando diferentes regiões brasileiras no que se refere à tradição de trabalhos desenvolvidos na área da sociolinguística, Freitag (2021, p. 24) destaca:

Enquanto as capitais e regiões urbanas, especialmente do Sudeste e do Sul do Brasil, apresentam uma tradição de pesquisa sociolinguística já consolidada, a região Nordeste carece de empreendimentos científicos nesse campo: as pesquisas linguísticas ainda são incipientes e ainda concentradas na região litorânea, em capitais, geograficamente limitada em relação ao todo da região. Muito pouco se sabe e muito se generaliza, alimentando estereótipos e o preconceito, e construindo um outro mito, o do ‘português nordestino’ (FREITAG, 2021, p. 24).

Ademais, a autora pontua que, embora sejam inegáveis as contribuições dos projetos regionais sociolinguísticos criados no Brasil, algumas limitações são verificadas: “[...] estratificam as amostras em função do perfil sociolinguístico da região, o que torna a comparação mais difícil” e “[...] seu escopo de atuação fica circunscrito ao escopo da sede do projeto, ou, mais especificamente, da instituição de ensino superior que o abriga” (FREITAG, 2021, p. 33). Para além desses aspectos, teceremos algumas considerações, na próxima seção, sobre a documentação linguística que vem sendo realizada no Brasil, advinda da construção de bancos de dados sociolinguísticos.

3 BANCOS DE DADOS SOCIOLINGUÍSTICOS DO BRASIL: ALGUNS APONTAMENTOS

Conforme mencionado na seção precedente, desde a segunda metade do século XX, sociolinguistas vêm investindo esforços para descrever a diversidade linguística no Brasil e, para tanto, a construção de bancos de dados torna-se um expediente necessário para que seja possível documentar as formas linguísticas variáveis. Essa documentação entendida aqui como um “protocolo de orientação sociolinguística para a geração de dados para alimentar descrições linguísticas e testagem de teorias” (FREITAG, 2016, p. 456) levanta uma questão primária: que critérios devem ser levados em conta durante a constituição de um banco de dados sociolinguístico? A resposta a essa pergunta tem variado ao longo do tempo.

Por um lado, se levada em conta a gênese dos estudos sociolinguísticos desenvolvidos no Brasil, Freitag, Martins e Tavares (2012) pontuam que a metodologia adotada para a constituição dos bancos de dados vincula-se a um modelo hegemônico que perdura ao longo das décadas, na medida em que, enquadrada nos estudos sociolinguísticos da primeira onda, seguindo a proposta de Eckert (2012), possibilita, através de técnicas estatísticas, a compreensão das “correlações entre variáveis linguísticas e categorias sociais primárias, como classe socioeconômica, sexo, idade, escolaridade etc.” (p. 920). Sob esse prisma, descrever a comunidade de fala é crucial e, para tanto, é importante padronizar as coletas dos dados, destacando-se as entrevistas sociolinguísticas como o principal instrumento para essas coletas.

Os autores verificam também que “as amostras são construídas em curto prazo”, “as categorias são definidas *a priori*” e “a distribuição é homogênea, tanto quanto ao tamanho da amostra quanto às categorias controladas” (p. 931). Além disso, algumas limitações são apresentadas por Freitag (2016) sobre os seguintes critérios, adotados para a constituição das amostras:

- (i) “controle da classe social através do controle da escolarização” (p. 455);
- (ii) “atendimento dos informantes a critérios básicos como: ter nascido e vivido na comunidade” (p. 456);

O problema que se apresenta em (i) tem a ver, segundo a autora, com o fato de que, na atual dinâmica da sociedade brasileira, é inviável controlar a classe social a partir do grau de escolarização do falante, pois tem havido mais estudantes na educação básica, e muitos têm ingressado no ensino público superior, que ganhou grande expansão para além dos principais centros urbanos. Quanto aos critérios mencionados na alínea (ii), eles apresentam-se limitantes considerando que os fluxos migratórios têm se intensificado no país por diversas razões, sendo uma delas a maior mobilidade dos estudantes de uma região para outra, em decorrência da expansão das instituições de ensino superior (IES).

Assemelhando-se aos estudos da primeira onda por manter o foco na comunidade de fala, os estudos da segunda onda, ainda escassos no Brasil, ampliam o universo da análise que, além de ser quantitativa, é etnográfica. Nesse sentido, para coletar os dados, o sociolinguista deve manter um “maior envolvimento com a comunidade” através do qual categorias sociodemográficas mais abstratas devem ser consideradas; elas “não são identificáveis em uma coleta rápida, como o julgamento de pertencimento à comunidade, valores, atitudes, etc.” (FREITAG, 2015, p. 4)

Por outro lado, contrariamente aos estudos da primeira e da segunda onda, os estudos sociolinguísticos da terceira onda são entendidos como uma

[...] proposta de retomada do significado social da variação, mudando o foco da estrutura para a prática linguística. Estudos dessa natureza tomam como unidade comunidades de prática – grupo de indivíduos engajados em função de um propósito – ao invés de comunidades linguísticas, permitindo investigar o papel do indivíduo, em termos de hierarquia, inovação e adesão a variáveis linguísticas (FREITAG, 2015, p. 4).

Assim, embora esteja prevista a realização de análise quantitativa, ela só poderá ser conduzida, segundo Freitag, Martins e Tavares (2012, p. 931), quando, por exemplo: (a) as categorias sociodemográficas mais amplas e abstratas relativas à comunidade de prática forem definidas *a posteriori*, implicando considerar que “a estratificação é baseada em valores localmente estabelecidos”; e (b) a amostra for constituída a longo prazo, em virtude de a coleta etnográfica dos dados exigir um grau de envolvimento maior do pesquisador com o grupo investigado através da “observação participante e da interação entre grupos”. Estudos dessa natureza têm ganhado fôlego nos últimos anos no Brasil e novos bancos de dados começam a contemplar indicadores sociodemográficos mais amplos não previstos no modelo hegemônico. Destacam-se, nesse cenário, os seguintes bancos de dados construídos/em construção na região Nordeste: (a) 2011 – Banco de fala culta de Itabaiana/SE (coordenação: profa.: Raquel Freitag); (b) 2011 – Banco de Falares Sergipanos/SE (coordenação: profa. Raquel Freitag); (c) 2012 – Banco de dados FALA-Natal (coordenação: prof. Marco Antônio Martins e profa. Maria Alice Tavares).

Diante do quadro esboçado aqui, sabemos que têm sido inegáveis as contribuições dos *corpora* constituídos a partir de bancos de dados provenientes dos estudos sociolinguísticos da primeira onda, resultado de um incansável trabalho de linguistas brasileiros de diferentes regiões para compreender como se estrutura a diversidade linguística em diferentes comunidades de fala:

[...] a elaboração desses *corpora* permitiu a descrição do português brasileiro em diferentes aspectos linguísticos e considerando distintas variedades. De algum modo, têm-se uma descrição da variação na(s) gramática(s) do português do Brasil envolvendo diferentes fenômenos e a correlação destes com variáveis sociais (FREITAG; MARTINS; TAVARES, 2012, p. 921).

Contudo, diante da vasta documentação sociolinguística já realizada no Nordeste, por exemplo, a partir da atuação de diferentes grupos de pesquisa, consideramos relevante ampliar o olhar para novas questões sociolinguísticas que convocam a constituição de bancos de dados adicionais que permitam uma análise pluridimensional da variação linguística, destacando-se, entre outros, o significado social das formas linguísticas em diferentes comunidades de prática.

4 MOSAICO SOCIOLINGUÍSTICO DO NORDESTE BRASILEIRO: LOCUS DE DIFERENTES FALARES

Como o próprio título desta seção sugere, não é possível assumir a ideia generalizante de um “português nordestino”, mas de um Nordeste constituído por diferentes falares. Segundo Mota (2016), existem subáreas dialetais nessa região. Além disso, esses falares inscrevem-se em uma região que vem passando por alterações em seu perfil socioeconômico, sobretudo quando verificados os grandes fluxos migratórios de pessoas para regiões mais afastadas dos grandes centros urbanos em decorrência do processo de interiorização das IES, que ocorre em três fases (FREITAG, 2021):

- 1ª fase (2003-2006): criação de 10 IES e 49 *campi* universitários (BRASIL, 2015). Foram criadas, em 2003, a Universidade Federal Rural do Semi-Árido (UFERSA) - Mossoró (RN) e, em 2005, a Universidade Federal do Recôncavo da Bahia (UFRB) - Cruz das Almas (BA);
- 2ª fase (2007-2012): criação de 4 IES e implantação e consolidação de 86 *campi*. Foi criada, em 2010, a Universidade da Integração Internacional da Lusofonia Afro-Brasileira (UNILAB) - Redenção (CE);
- 3ª fase: criação de 4 IES e 47 *campi*. Foram criadas, em 2013, a Universidade Federal do Oeste da Bahia (UFOB) - Barreiras (BA), a Universidade Federal do Sul da Bahia (UFSB) - Itabuna (BA) e a Universidade Federal do Cariri (UFCA) - Juazeiro do Norte (CE).

Com essa expansão, o fluxo migratório de estudantes se intensifica no Nordeste. Segundo pontua Freitag (2021, p. 39), os contatos linguísticos nessa região

[...] acontecem com muita proficuidade [...] tanto em perspectiva diacrônica (línguas de base africana e indígena), como em perspectiva sincrônica: os movimentos de retorno e os fluxos migratórios motivados pela mobilidade acadêmica da expansão da educação superior instauram um cenário de contatos linguísticos propício ao estudo sobre efeitos de interferência de variedades e de acomodação dialetal. Ao mesmo tempo, a expansão da educação superior permite a ampliação da pesquisa sociolinguística para além das capitais nordestinas (FREITAG, 2021, p. 39).

Exemplificando a intensificação desses contatos na atual sincronia, Rubio, Torres e Cá (2021) realizam um estudo sobre a diversidade linguística e a prática docente no contexto multicultural da UNILAB-CE cuja missão é “proporcionar um espaço de convívio harmonioso das diferenças entre os povos e culturas e do respeito à diversidade étnico-racial, religiosa, cultural, de gênero e linguística” (p. 400). O quadro discente dessa instituição é composto por brasileiros do interior do Ceará e por estudantes da Comunidade dos Países de Língua Portuguesa (CPLP).

Em 2018, a UNILAB contava com mais de 3.300 estudantes, sendo quase 1000 estrangeiros e, em 2019, com mais de 4.619 discentes, sendo 1156 estrangeiros. Nesse cenário, dá-se o contato entre diferentes variedades do português no interior do Nordeste: uma variedade adquirida como L1 pelos brasileiros e outra variedade adquirida como L2 por grande parte dos estrangeiros, que convive com suas L1 (por exemplo: línguas crioulas e bantas). Tomando por base essa vasta diversidade linguística, Rubio, Torres e Cá (2021) consideram que a prática docente deve

- “realçar [nas grades curriculares] a importância [dessa diversidade] como fator de integração, disseminação e democratização do saber”, destacando, por exemplo, “o caráter variável e heterogêneo da língua portuguesa falada no Brasil” (p. 410-411);
- “dar-se conta de que, para grande parte dos estudantes estrangeiros, [...] o português é uma língua não materna, aprendida em situação escolar [...], de modo que

a variedade que trazem não lhes garante uma proficiência suficiente exigida no ambiente acadêmico, o que pode gerar desestímulo e uma sensação de impotência”. Além disso, “fenômenos variáveis e as características do português brasileiro não são reconhecidos nem fazem sentido como manifestação linguística legítima para [esses estudantes], que ainda convivem com o mito de que aprenderam a língua portuguesa padrão” (p. 411).

Ao propor essas ações/estratégias, Rubio, Torres e Cá sinalizam para a necessidade de as pesquisas sociolinguísticas investirem mais no ativismo sociolinguístico que prevê uma interface, um diálogo multidisciplinar entre diferentes domínios (social, cultural, etnográfico, linguístico e educacional). Para tanto, faz-se necessária uma “postura ativa dos pesquisadores da Sociolinguística em mobilizar os resultados de seus estudos para revertê-los em tecnologia social: materiais didáticos e de difusão no suporte, reconhecimento e valorização das variedades” (FREITAG, 2016, p. 451).

Feitas essas considerações, abordaremos, na próxima subseção, o que chamamos de “fotografias sociolinguísticas do Nordeste”, delineadas sob diferentes focos (a produção linguística, a compreensão sócio-histórica e a avaliação social). Longe de exaurirmos todo o percurso investigativo realizado nessa região, apresentaremos alguns *flashes* que poderão suscitar a geração de novos empreendimentos investigativos.

4.1 FOTOGRAFIAS SOCIOLINGUÍSTICAS DO NORDESTE: FOCALIZANDO ALGUNS FLASHES

Tal como apresentado nas seções anteriores, os estudos sociolinguísticos no Brasil embasam-se, em grande medida, no modelo hegemônico da primeira onda, por centrarem a atenção em comunidades de fala, enfocando a distribuição e a sistematização de formas linguísticas variáveis. Centrando nossa atenção na região Nordeste, verificamos que grupos de pesquisa têm se voltado, em geral, à constituição de amostras e, conseqüentemente, à documentação sociolinguística. Esse fato se evidencia pela intensa construção de bancos de dados decorrentes de projetos de investigação sociolinguística realizados em seus diferentes estados, conforme verificado a seguir, a partir de uma linha cronológica que se inicia na segunda metade do século XX:

- 1988: Projeto Dialetos Sociais Cearenses (coordenação: profa. Claudia Roncarati – UFC);
- 1990: Programa para a História da Língua Portuguesa (PROHPOR) (coordenação: profa. Rosa Virgínia Mattos e Silva – UFBA);
- 1991: Projeto A Língua Usada em Alagoas (LUAL) (coordenação: profa. Maria Denilda Moura – UFAL);
- 1993: Projeto: Variação Linguística no Estado da Paraíba (VALPB) (coordenação: prof. Dermeval da Hora – UFPB);
- 1996: Projeto A Língua Portuguesa no Semi-árido Baiano (coordenação: profa. Norma Lúcia Fernandes de Almeida e profa. Zenaide Oliveira Novais Carneiro – UEFS);

- 1998: Núcleo de Estudos da Língua Portuguesa (NELP) (coordenação: profa. Norma Lúcia Fernandes de Almeida e profa. Zenaide Oliveira Novais Carneiro – UEFS);
- 1998: Programa de Estudos sobre o Português Popular de Salvador (PEPP) (coordenação: profa. Norma da Silva Lopes – UNEB);
- 2002: Projeto Vertentes do Português Popular do Estado da Bahia (coordenação: Prof. Dante Lucchesi – UFBA);
- 2010: Projeto Aspectos Gramaticais do Português Falado por Estudantes Teresinenses (PORFATER) (coordenação: profa. Maria Auxiliadora Ferreira Lima e profa. Maria Anecy Calland Marques Serra – UFPI);
- 2012: Projeto Variação e Processamento da Fala e do Discurso: Análises e Aplicações – PROFALA (coordenação: profa. Maria Elias Soares e profa. Maria do Socorro Silva de Aragão);
- 2015: Projeto Língua Usada em Pernambuco (LUPE) (coordenação: prof. Adeilson Pinheiro Sedrins – UFRPE);
- 2015: Projeto Laboratório de Edição e Documentação Linguística de Pernambuco (LEDOC) (coordenação: prof. Cleber Ataíde – UFRPE);
- 2016: Projeto Falares Sergipanos virtual: variedade, diversidade, contato e os direitos linguísticos (coordenadora: profa. Raquel Freitag – UFS);
- 2018: Projeto Variação Morfossintática com base no Português Maranhense (Var-Sint) (coordenação: Profa. Cibelle Corrêa Béliche Alves – UFMA).

Dentre outras metas, esses grupos de pesquisa e muitos outros têm contribuído, de forma significativa, para o fortalecimento do campo de estudos voltados à sociolinguística variacionista e para a geração de futuras investigações a partir da definição e sistematização de questões adicionais, possibilitando o acesso e a formação de recursos humanos no âmbito da graduação e da pós-graduação. Com o advento da expansão das IES no século XXI, verificamos também que novos projetos e grupos de pesquisa sociolinguística vêm sendo implementados e consolidados, para além das capitais nordestinas, por pesquisadores que atuam nas instituições recém-criadas, agregando conhecimentos aos estudos sociolinguísticos sobre a compreensão da diversidade linguística no Nordeste.

Levando em conta que há “uma polarização entre grandes centros urbanos, localizados no litoral, e todo o interior seja semiárido, seja sertão, que corresponde a mais de 50 de todo o território político-administrativo [dessa região]” (FREITAG, 2021, p. 28), torna-se, portanto, insustentável a defesa do mito do “português nordestino” que se apresenta generalizante e, ao mesmo tempo, limitante:

O português nordestino se constitui no imaginário popular brasileiro como o português das novelas da Globo, com ‘sotaque forte’(SOUZA; MARCOLINO, 2016), como se todo o Nordeste fosse um conjunto homogêneo e que parou no tempo, num período rural, das secas e dos retirantes, pré-industrialização. No entanto, assim como o português brasileiro é um mito a ser

desfeito (FREITAG *et al.*, 2015), a partir de estudos sociolinguísticos, o mesmo pode se dizer do mito do 'português nordestino' [...] (FREITAG, 2021, p. 24).

Ainda sobre os projetos de pesquisa realizados no Nordeste, Freitag (2021, p. 33) compara-os a projetos nacionais (a exemplo do NURC) e tece as seguintes considerações:

[...] enquanto os projetos nacionais apresentam uma metodologia de constituição de amostras padronizada, que permite a comparação, os projetos regionais estratificam suas amostras em função do perfil sociolinguístico da região, o que torna a comparação mais difícil. Outro ponto de limitação dos projetos regionais é que, ainda que ampliem o panorama da investigação linguística da região Nordeste, seu escopo de atuação fica circunscrito ao escopo da sede do projeto, ou mais especificamente, da instituição de ensino superior que o abriga (FREITAG, 2021, p. 33).

Vale referir que, além das inegáveis contribuições dos projetos regionais, as publicações voltadas à descrição dos usos linguísticos no Nordeste em diferentes níveis de análise linguística também têm se intensificado ao longo dos anos, sobretudo no século XXI. A lista apresentada a seguir, longe de ser exaustiva, vai desde o trabalho seminal de Marroquim (1934) até produções mais recentes:

- *A língua do Nordeste: Alagoas e Pernambuco*, de Mário Marroquim (1934);
- *Análise fonético-fonológica do falar paraibano*, de Maria do Socorro Silva de Aragão (1978);
- *Projeto Variação Linguística no Estado da Paraíba (VALPB)*, de Dermeval da Hora e Juliene Lopes Ribeiro Pedrosa (2001);
- *O português falado no Maranhão: estudos preliminares*, de Conceição de Maria de Araújo Ramos, José de Ribamar Mendes Bezerra e Maria de Fátima Sopas Rocha (2005);
- *A diversidade do português falado no Maranhão: o Atlas Linguístico do Maranhão em foco*, de Conceição de Maria de Araújo Ramos, Maria de Fátima Sopas Rocha e José de Ribamar Mendes Bezerra (2006);
- *Amostras da língua falada na zona rural de Anselino da Fonseca (Piemonte da Diamantina)*, de Norma Lúcia Fernandes de Almeida e Zenaide de Oliveira Novais Carneiro (2008);
- *Resquícios de Palmares: o que uma comunidade quilombola nos diz*, de Denilda Moura (2009);
- *O português falado no Maranhão: múltiplos olhares*, de Conceição de Maria de Araújo Ramos, José de Ribamar Mendes Bezerra e Maria de Fátima Sopas Rocha (2010);
- *Variação e mudança linguística no estado de Alagoas*, de Januacele Francisca da Costa, Renata Livia de Araújo Santos e Elyne Giselle de Santana Lima Aguiar Vitória (2011);
- *Aspectos descritivos e sócio-históricos da língua falada em Pernambuco*, de Adeilson Pinheiro Sedrins e Edmilson José de Sá (2015);

- *A fala nordestina: entre a sociolinguística e a dialetologia*, de Norma da Silva Lopes, Silvana Silva de Farias Araújo e Raquel Meister Ko. Freitag (2016);
- *Variação linguística em Feira de Santana*, de Norma Lúcia Fernandes de Almeida, Silvana Silva de Farias Araújo, Eliana Pitombo Teixeira e Zenaide de Oliveira Novais Carneiro (2016);
- *Estudos sociodialetais do estado do Maranhão*, de Conceição de Maria de Araújo Ramos *et al.* (2019);
- *Estudos sobre o português do Nordeste: língua, lugar e sociedade*, de Norma da Silva Lopes, Josane Moreira de Oliveira e Lúcia Maria de Jesus Parcero (2020);
- *A língua em foco no nordeste brasileiro: d`além das capitais*, de Cícero Barboza Nunes e Claudia Roberta Tavares Silva (2021).

Em linhas gerais, é perceptível o vasto e importante empreendimento de linguistas não só das capitais nordestinas, mas para além delas, em descrever os múltiplos usos linguísticos, revelado pela grande quantidade de amostras constituídas. No entanto, assumimos, com Freitag (2021, p. 40), que “é preciso investir no desenvolvimento de metodologia para a padronização de coletas, atendendo às especificidades sociodemográficas da região”.

Ao lado dos estudos voltados à descrição das formas variáveis do ponto de vista sincrônico, destacam-se, no cenário dos estudos sociolinguístico no Nordeste, a partir da década de 1980, os trabalhos investigativos realizados pela profa. Rosa Virgínia Mattos e Silva (UFBA) sobre a constituição da sócio-história do português brasileiro (PB). Segundo ela, para que as mudanças linguísticas “internas” a essa variedade sejam compreendidas em profundidade, é crucial que seja levado em conta o contexto sócio-histórico associado ao contato linguístico do português com outras línguas, iniciado no século XVI em território brasileiro:

[...] a partir da variação espacial e social sincrônica há novas preocupações com a ‘história interna’ do português brasileiro, que para ser interpretada em sua totalidade exige que melhor e mais detalhadamente se conheça a complexa ‘história externa’ da sociedade multilíngue do espaço que, a partir do século XVI, se denominou Brasil (MATTOS E SILVA, 2004, p. 30).

Visando à ampliação dos estudos sociolinguísticos sob essa perspectiva de análise, a referida pesquisadora, no ano de 1990, institui e coordena, na UFBA, o *Programa para a História da Língua Portuguesa (PROHPOR)*, um projeto pioneiro na região Nordeste, que têm aberto espaço à geração de estudos investigativos sobre as mudanças linguísticas que se processaram e ainda se processam no interior das estruturas do PB em decorrência do contexto multilíngue iniciado no século XVI, em que línguas africanas e indígenas passam a entrar em contato com o português: “Há muitas histórias por reconstruir sobre as faces indígena e negra do Brasil e, conseqüentemente, sobre as variantes do português brasileiro que aí se veicularam e se veiculam” (MATTOS E SILVA, 2004, p. 20). Dentre esses estudos, destacam-se as pesquisas de Baxter (1992), Baxter; Lucchesi (1997), Ferreira (1999), Lobo (2001), Lucchesi (2000), Andrade (2003),

Burgos (2003), Souza (2004), Araújo (2006), Cavalcante (2007), Lucchesi; Baxter; Ribeiro (2009) e outros.

Paralela à perspectiva de análise defendida por Mattos e Silva, surge, no final da década de 1980, uma nova abordagem para o tratamento da mudança linguística no PB, denominada sociolinguística paramétrica, que prevê a interface entre a sociolinguística e a teoria da sintaxe gerativa, mais especificamente, o modelo de princípios e parâmetros, cujo precursor é Fernando Tarallo (1987; 1993).

No livro intitulado *Português brasileiro: uma viagem diacrônica*, organizado por Roberts e Kato (1993), as análises desenvolvidas ao longo dos capítulos convergem para a seguinte conclusão: evidências quantitativas apontam que o século XIX — mais precisamente a sua segunda metade — é o *locus* do início de uma mudança paramétrica do PB em relação ao português europeu quando observados, por exemplo, o sistema pronominal, as estratégias de relativização e a ordem de palavras.

Até onde temos verificado, pesquisas no âmbito da sociolinguística paramétrica são escassas no Nordeste, destacando-se, entre elas, a tese de Cavalcante (2001) e a dissertação de Mendonça (2004), produzidas na Universidade Federal de Alagoas sob a orientação da profa. Denilda Moura. Os resultados desses estudos convergem em relação aos resultados de trabalhos realizados em outras regiões brasileiras, evidenciando que o processo de mudança paramétrica tem se implementado na gramática do PB.

Até o momento, discorreremos sobre o percurso de trabalhos investigativos no Nordeste que têm se concentrado na explicação dos usos linguísticos seja no plano sincrônico, seja no plano da história “externa” ou “interna”. Trata-se de uma perspectiva de análise que tem suas raízes no século XX e se estende pelo século XXI, quando comparada aos estudos voltados à avaliação social das formas linguísticas variáveis, que só passa a receber uma maior atenção dos pesquisadores dessa região no atual século.

Consideramos que desenvolver estudos sobre a avaliação social ampliará nossa visão de como os falantes avaliam positiva ou negativamente não só seus próprios usos, mas também os de outros. Sob esse enfoque, o tipo de atitude linguística pode ser determinante para uma rápida ou lenta implementação de uma mudança linguística:

Uma atitude favorável ou positiva pode fazer que uma mudança linguística se cumpra mais rapidamente, que em certos contextos predomine o uso de uma língua em detrimento de outra, que o ensino-aprendizagem de uma língua estrangeira seja mais eficaz, que certas variantes linguísticas se confinem aos contextos menos formais e outras predominem nos estilos cuidadosos. Uma atitude desfavorável ou negativa pode levar ao abandono e ao esquecimento de uma língua ou impedir a difusão de uma variante ou uma mudança linguística (MORENO FERNÁNDES, 1998, p. 179).

No que se refere às mudanças linguísticas em curso, Labov ([1978] 2008) verifica que o comportamento dos falantes se correlaciona às suas atitudes linguísticas, entendidas aqui como “como qualquer índice cognitivo ou comportamental de reações

avaliativas em direção às variedades diferentes de língua ou de seus falantes” (GILES; RYAM; SEBASTIAN, 1982, p. 7).

A exemplo de pesquisas voltadas para o significado social das formas linguísticas em diferentes níveis da gramática por falantes de diferentes estados do Nordeste, elencamos as seguintes (a lista não pretende ser exaustiva):

- Ramos (1997): *Avaliação de dialetos brasileiros: o sotaque*;
- Cardoso (2015): *Atitudes linguísticas e avaliações subjetivas de alguns dialetos brasileiros* (tese defendida em 1989);
- Raquel (2015): *Competências sociolinguísticas na escola pública cearense: saberes, atitudes linguísticas e práticas pedagógicas de docentes no 5º ano do Ensino Fundamental*;
- Freitag e Santos (2016): *Percepção e atitudes linguísticas em relação às africadas pós-alveolares em Sergipe*;
- Moura (2017): *Atitudes linguísticas de graduandos em Letras em relação à redução do ditongo final [já] em palavra paroxítona e às pessoas que a empregam*;
- Santos (2018): *Crenças e atitudes de variedades linguísticas no nordeste brasileiro: como universitários de uma zona de fronteira dialetal percebem a língua portuguesa?*;
- Hora, Henrique e Amorim (2018): *Produção e percepção: o processo da palatalização em jogo*;
- Miranda, Santos e Silva (2018): *Crenças e atitudes linguísticas: o que dizem os maranhenses sobre sua fala*;
- Souza e Araújo (2000): *Reações subjetivas de estudantes da cidade de Tucano-BA às variantes não padrão da concordância verbal*;
- Vitória (2021): *Avaliação social e a concordância verbal com o pronome tu*;
- Pereira (2021): *Avaliações subjetivas de universitários recifenses acerca da realização de artigo definido feminino diante de antropônimo masculino no português brasileiro*;

Ao nosso ver, comparar as atitudes linguísticas não só entre falantes do Nordeste — domínio interestadual e intraestadual (entre mesorregiões) —, mas também as suas com as de falantes de outras regiões brasileiras constitui-se uma das tarefas da agenda de trabalho sociolinguística que precisa se intensificar, visando contribuir, de forma adicional, para a valorização dos diferentes falares regionais e, conseqüentemente, para a luta contra o mito do “português nordestino”. Nesse sentido, concordamos com Freitag (2015, p. 5) que “as variações regionais são delimitadas a partir não apenas das regularidades de usos linguísticos, mas também pela maneira como as pessoas avaliam tais usos e os efeitos dessas avaliações no seu comportamento linguístico”

5 PALAVRAS FINAIS

Diante do panorama dos estudos sociolinguísticos aqui traçado neste capítulo, concordamos com Salomão (2011, p. 187) que as pesquisas empreendidas desde a década de 1970 têm contribuído não só para a “construção de um panorama de comunidades de fala existentes no Brasil”, como também para a “desconstrução da ideia de

homogeneização linguística”. Para além dessas contribuições, destacamos as seguintes prospecções que passam a integrar uma agenda de trabalho que, longe de ser limitante, se apresenta dinâmica e instigante para a compreensão do mosaico sociolinguístico do Nordeste:

- “busca pela inserção em políticas públicas (inserção social) da Sociolinguística do Brasil, o ativismo sociolinguístico” (FREITAG, 2016, p. 457);
- maior investimento: (a) “em estudos interdisciplinares (confluências de vários componentes linguísticos e extralinguísticos)”; (b) “na comparação entre as muitas variedades já descritas”; (c) “no desenvolvimento de pesquisas diacrônicas”; (d) na “investigação de comunidades de práticas, com ênfase na análise dos estilos; (e) “em novas ferramentas de suporte estatístico” (OLIVEIRA, 2016, p. 491);
- novos empreendimentos no campo da documentação sociolinguística (registro, salvaguarda e difusão dos usos linguísticos no Nordeste), visando não só à constituição de bancos de dados adicionais que contemplem indicadores sociodemográficos mais amplos, mas também à padronização das coletas para posteriores estudos contrastivos;
- ampliação de estudos dialetológicos pluridimensionais que contemplem diferentes áreas do Nordeste, o que contribuirá, de forma adicional, para uma maior compreensão do panorama sociolinguístico dessa região, tendo em mente, por exemplo, a inclusão de mais falantes topodinâmicos do que topoestáticos;
- retorno a comunidades de fala já descritas em bancos de dados pioneiros do Nordeste para a coleta de novos dados, a fim de mapear formas linguísticas variáveis que ainda permanecem ou não, tomando por base que a construção do perfil social dos novos informantes deve contemplar variáveis sociodemográficas mais abstratas;
- ampliação de estudos que contemplem a sócio-história do PB em diferentes estados do Nordeste;
- maior enfoque na descrição dos múltiplos usos e na avaliação social das variáveis não só nas capitais nordestinas, mas, sobretudo, d’além dessas capitais, visando à *ampliação das fotografias sociolinguísticas*, levando em conta uma dimensão pluri-dimensional, não prevista no modelo hegemônico;
- intensificação no desenvolvimento de projetos de pesquisa em rede (domínio intrarregional, interregional e internacional), visando à (re)construção e difusão de conhecimentos sobre a diversidade linguística.

São inúmeras as ações que essa agenda de trabalho nos sinaliza. Isso nos motiva a investir na formação e continuidade de projetos de pesquisa e grupos de investigação que, através de uma atuante cooperação, possam engendrar novas discussões em torno dos diferentes falares do Nordeste, enfocando não só a relação intrínseca entre língua e contexto social, mas também a distribuição e a avaliação social das formas linguísticas variáveis instauradas no domínio da estratificação socioeconômica.

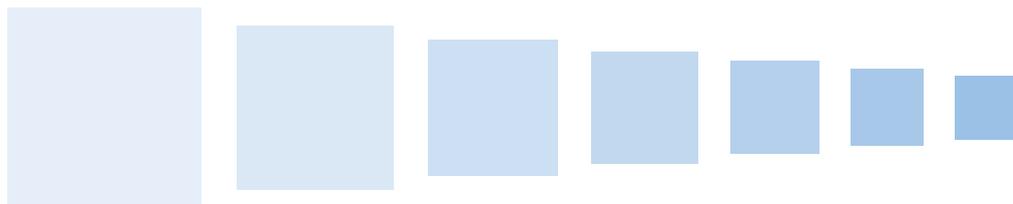
REFERÊNCIAS

- ALMEIDA, E. M. O. **Atlas Linguístico da Mata Sul de Pernambuco**. 2009. 151 f. Dissertação (Mestrado em Letras) – Universidade Federal da Paraíba, João Pessoa.
- ANDRADE, P. R. **Um fragmento da constituição sócio-histórica do português do Brasil: variação na concordância nominal de número em um dialeto afro-brasileiro**. 2003. 147f. Dissertação (Mestrado em Linguística) – Universidade Federal da Bahia, Salvador.
- ARAGÃO, M. S. S.; BEZERRA DE MENEZES, C. P. **Atlas Lingüístico da Paraíba**. Brasília: UFPB/CNPq, Coordenação Editorial, 1984; v. 1, 2.
- ARAÚJO, E. A. **As construções de tópico do português dos séculos XVIII e XIX: uma análise sintático-discursiva**. 2006. 293f. Tese (Doutorado em Linguística) – Universidade Federal da Bahia, Salvador.
- ARAÚJO, S. S. F. **A concordância verbal no português falado em Feira de Santana - BA: sociolinguística e sócio-história do português brasileiro**. 2014. 342f. Tese (Doutorado em Língua e Cultura) – Universidade Federal da Bahia, Salvador.
- BAXTER, A. A contribuição das comunidades afro-brasileiras isoladas para o debate sobre a criouliização prévia: um exemplo do Estado da Bahia. In: D'ANDRADE, E.; KIHM, A. (Orgs.). **Actas do Colóquio sobre Crioulos de Base Lexical Portuguesa**. Lisboa: Colibri, 1992. p.7-36.
- BAXTER, A.; LUCCHESI, D. A relevância dos processos de pidginização e criouliização na formação da língua portuguesa no Brasil. **Estudos Linguísticos e Literários**, Salvador, n.19, p.65-83, 1997.
- BESSA, J. R. F. (Coord.). **Atlas Linguístico do Ceará**. Vol.I – Introdução, Vol.II – Cartogramas. Universidade Federal do Ceará. Fortaleza: Edições UFC, 2010.
- BRAGA, M. L. **A concordância de número no sintagma nominal no triângulo mineiro**. 1977. 88f. Dissertação (Mestrado em Língua Portuguesa) – Pontifícia Universidade Católica, Rio de Janeiro.
- BRANDÃO, S. F. **A geografia lingüística do Brasil**. São Paulo: Ática, 1991.
- BRANDÃO, S. F. GT de Sociolinguística. **Revista da Anpoll**, 1(1), 1994.
- BURGOS, L. E. S. de. **Estratégias de uso das relativas em uma comunidade de fala afro-brasileira**. 2003. 122f. Dissertação (Mestrado em Letras e Linguística) – Universidade Federal da Bahia, Salvador.
- CALLOU, D. **Variação e distribuição da vibrante na fala culta do Rio de Janeiro**. 1979. 199f. Tese (Doutorado em Língua Portuguesa) – Universidade Federal do Rio de Janeiro, Rio de Janeiro.
- CALLOU, D. Quando dialetologia e sociolinguística se encontram. **Estudos Linguísticos e Literários**, Salvador, n.41, p.29-48, jan./jun. 2010.
- CARDOSO, D. P. **Atitudes linguísticas e avaliações subjetivas de alguns dialetos brasileiros**. São Paulo: Blücher, 2015.
- CARDOSO, S. A. M. Geolinguística no terceiro milênio: monodimensional ou pluridimensional? **Revista GELNE**, Fortaleza, v.4, n.2, p.1-16, 2002a. í
- CARDOSO, S. A. M. **Atlas Lingüístico de Sergipe II**. Rio de Janeiro, 2002b. v. 2.
- CARDOSO, S. A. M. *et al.* (Orgs.). **Atlas lingüístico do Brasil**. v. 1 (Introdução). Londrina: EDUEL, 2014a.
- CARDOSO, S. A. M. *et al.* (Orgs.). **Atlas lingüístico do Brasil**. v. 2 (Cartas lingüísticas 1). Londrina: EDUEL, 2014b.
- CAVALCANTE, A. **O sujeito pronominal em Alagoas e no Rio de Janeiro: um caso de mudança em progresso**. 2001. 128f. Tese (Doutorado em Linguística) – Universidade Federal de Alagoas, Maceió.
- CAVALCANTE, R. **A negação pós-verbal no português brasileiro: análise descritiva e teórica de dialetos rurais de afrodescendentes**. 2007. 160f. Dissertação (Mestrado em Letras) – Universidade Federal da Bahia, Salvador.
- CEZARIO, M. M.; VOTRE, S. Sociolinguística. In: MARTELOTA, M. E. (Org.). **Manual de Linguística**. São Paulo: Contexto, 2008. p. 141-155.
- DOIRON, M. P. B. A motivação semântica nas respostas dos informantes do **Atlas Linguístico do Estado de Alagoas – ALEAL**. 2017. 488f. Tese (Doutorado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina.
- ECKERT, P. Three waves of variation study: the emergence of meaning in the study of sociolinguistic variation. **Annual Review of Anthropology**, n. 41, p. 87-100, 2012.
- FERREIRA, C. *et al.* **Atlas Lingüístico de Sergipe**. Salvador: UFBA – Instituto de Letras/Fundação Estadual de Cultura de Sergipe, 1987.

- FERREIRA, G.. **Cinzeno**: memória de uma comunidade negra remanescente de quilombo. 1999. 195f. Dissertação (Mestrado em Ciências Sociais) – Pontifícia Universidade de São Paulo, São Paulo.
- FREITAG, R. M. Ko. Prefácio: Atitudes e Identidade Linguística: muito chão pela frente. In: CARDOSO, D. P. **Atitudes Linguísticas e Avaliações Subjetivas de Alguns Dialeto Brasileiros**. São Paulo: Blucher, 2015. p. 3-8.
- FREITAG, R. M. Ko. Sociolinguística no/do Brasil. **Cadernos de Estudos Linguísticos**, (58.3), Campinas, pp. 445-460 - set./dez. 2016.
- FREITAG, R. M. Ko. Português nordestino: para além das capitais. In: NUNES, C. B.; SILVA, C. R. T. (Orgs.). **A língua em foco no nordeste brasileiro**: d' além das capitais. São Paulo: Pontes, 2021. p. 23-44.
- FREITAG, R. M. Ko.; MARTINS, M. A.; TAVARES, M. A. Banco de dados sociolinguísticos do português brasileiro e os estudos da terceira onda: potencialidades e limitações. **Alfa**, São Paulo, 56 (3), p. 917-944, 2012.
- FREITAG, R. M. Ko.; SANTOS, A. D. Percepção e atitudes linguísticas em relação às africadas pós-alveolares em Sergipe. In: LOPES, N. S.; ARAÚJO, S. S. F.; FREITAG, R. M. Ko. (Orgs.) **A fala nordestina**: entre a sociolinguística e a dialetologia. São Paulo: Blucher, 2016. p. 109-122.
- GILES, H.; RYAN, E. B.; SEBASTIAN, R. J. An integrative perspective the study of attitudes toward language variation. In: GILES, H.; RYAN, E. B. (Eds.). **Attitudes towards language variation**: social and applied context. London: Edward Arnold, 1982. p. 1-19.
- HORA, D.; HENRIQUE, P. F. L.; AMORIM, A. W. D. Produção e percepção: o processo de palatalização em jogo. **Diadorim**, Rio de Janeiro, vol. 20, n. 2, p. 280-296, jul.-dez. 2018.
- ISQUERDO, A. N.; ROMANO, V. P. Discutindo a dimensão sociolinguística do projeto ALiB: uma reflexão a partir do perfil dos informantes. **Alfa**, São Paulo, 56 (3), p. 891-916, 2012.
- LABOV, W. **The social stratification of English in New York City**. Washington. D. C: Center for Applied Linguistics, 1966.
- LABOV, W. **Padrões sociolinguísticos**. Trad. Marcos Bagno, Maria Marta Pereira Scherre e Caroline R. Cardoso. São Paulo: Parábola, [1972] 2008.
- LEMLE, M.; NARO, A. J. **Competências básicas do português**. Rio de Janeiro: MOBRAL, Fundação Ford, 1977.
- LOBO, T. **Para uma sociolinguística histórica do português do Brasil: edição filológica e análise linguística de cartas particulares do Recôncavo da Bahia, século XIX**. 2001. 248f. 4 vols. Tese (Doutorado em Filologia e Língua Portuguesa) – Universidade de São Paulo, São Paulo.
- LUCCHESI, D. **A variação na concordância de gênero em uma comunidade de fala afrobrasileira**: novos elementos sobre a formação do português popular do Brasil. 2000. 364f. Tese (Doutorado em Linguística) – Universidade Federal do Rio de Janeiro, Rio de Janeiro.
- LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. (Orgs.). **O português afro-brasileiro**. Salvador: EDUFBA, 2009.
- MARROQUIM, M. **A língua do Nordeste**: Alagoas e Pernambuco. 2. ed. Companhia Editora Nacional, São Paulo, 1934.
- MATTOS E SILVA, R. V. **Ensaio para uma sócio-história do português brasileiro**. São Paulo: Parábola, 2004.
- MENDONÇA, V. A. **O objeto direto anafórico na fala matagrãndense e paulistana**: um estudo comparativo. 2004. 97 f. Dissertação (Mestrado em Linguística; Literatura Brasileira) – Universidade Federal de Alagoas, Maceió.
- MIRANDA, A. L. A.; SANTOS, V. V.; SILVA, F. C. Crenças e atitudes linguísticas: o que dizem os maranhenses sobre sua fala. **Revista de Letras Juçara**, Caxias-Maranhão, v. 02, n. 01, p. 112-128, jul. 2018.
- MOLLICA, M. C. M. **Estudo da cópia nas construções relativas em português**. 1977. 86f. Dissertação (Mestrado em Letras) – Pontifícia Universidade Católica, Rio de Janeiro.
- MONTALVÃO, S. M. **Derivação sufixal nominal e variação linguística**. 1978. 106f. Dissertação (Mestrado em Língua Portuguesa) – Pontifícia Universidade Católica, Rio de Janeiro.
- MORENO FERNÁNDEZ, F. **Principios de sociolingüística y sociología del lenguaje**. Barcelona: Ariel, 1998.
- MOTA, J. Aspectos fônicos do Nordeste a partir de dados do ALiB. In: LOPES, N. S.; ARAÚJO, S. S. F.; FREITAG, R. M. Ko. (Orgs.). **A fala nordestina**: entre a sociolinguística e a dialetologia. São Paulo: Blucher, 2016. p. 59-74.

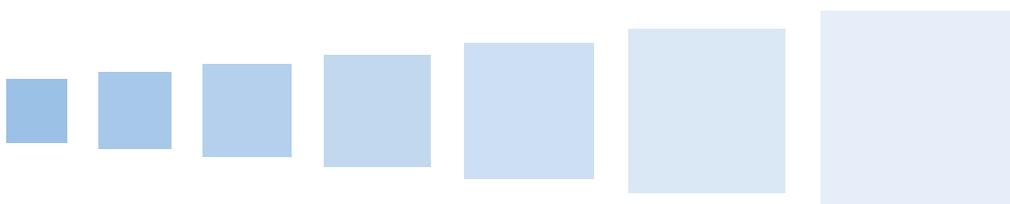
- MOTA, J. A.; CARDOSO, S. A. M. Para uma nova divisão dos estudos dialetais brasileiros. *In*: MOTA, J. A.; CARDOSO, S. A. M. (Orgs.). **Documentos 2: Projeto Atlas Linguístico do Brasil**. Salvador: Quarteto, 2006. p.15-34.
- MOURA, E. S. V. **Atitudes linguísticas de graduandos em Letras em relação à redução do ditongo final [ja] em palavra paroxítona e às pessoas que a empregam**. 2017. 96f. Tese (Doutorado em Estudos da Linguagem) – Centro de Ciências Humanas, Letras e Artes, Universidade Federal do Rio Grande do Norte, Natal.
- OLIVEIRA, J. M. A sociolinguística laboviana: festejando o cinquentenário e planejando o futuro. **Estudos Linguísticos**, 58(3), p. 481-501, set./dez. 2016.
- OMENA, N. P. **Pronome pessoal de terceira pessoa: suas formas variantes em função acusativa**. 1978. 138f. Dissertação (Mestrado em Língua Portuguesa) – Pontifícia Universidade Católica, Rio de Janeiro.
- PAIVA, M. C.; SCHERRE, M. M. P. Retrospectiva sociolinguística: contribuições do PEUL. **D.E.L.T.A.**, Vol. 15, n° especial, p. 201-232, 1999.
- PEREIRA, M. N. **Atlas geolinguístico do litoral potiguar**. 2v. 2007. Vol I: 123f. Vol. II: 189f. Tese (Doutorado em Letras Vernáculas) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro.
- PEREIRA, D. K, F. **Avaliações subjetivas de universitários recifenses acerca da realização de artigo definido feminino diante de antropônimo masculino no português brasileiro**. 2021. 170f. Tese (Doutorado em Letras) – Universidade Federal de Pernambuco, Recife, 2021.
- RAQUEL, B. M. G. **Competências sociolinguísticas na escola pública cearense: saberes, atitudes linguísticas e práticas pedagógicas de docentes no 5º ano do Ensino Fundamental**. 2015. 263f. Tese (Doutorado em Linguística Aplicada) – Universidade Federal do Ceará, Departamento de Letras Vernáculas, Programa de Pós-graduação em Linguística, Fortaleza.
- RAMOS, J. Avaliação de dialetos brasileiros: o sotaque. **Rev. Est. Ling.**, Belo Horizonte, ano 6, n.5, v.1, p.103-125, jan./jun. 1997.
- RAMOS, C. M. A. (Coord.). **Projeto Atlas Linguístico do Maranhão**. São Luís: UFMA, 200-.
- ROBERTS, I.; KATO, M. A. (Orgs.). **Português brasileiro: uma viagem diacrônica**. Campinas: Editora da Unicamp, 1993.
- ROSSI, N. **Atlas prévio dos falares baianos**. Rio de Janeiro: INL, 1963.
- ROSSI, N.; FERREIRA, C.; ISENSEE, D. **Atlas Prévio dos Falares Baianos (APFB)**. Rio de Janeiro: Ministério de Educação e Cultura; Instituto Nacional do Livro, 1963.
- RUBIO, C. F.; TORRES, F. F.; CÁ, I. N. Integração Internacional lusófona no sertão do Ceará: reflexões para o estudo da linguagem no ensino superior. *In*: NUNES, C. B.; SILVA, C. R. T. (Orgs.). **A língua em foco no nordeste brasileiro: d' além das capitais**. São Paulo: Pontes, 2021. p. 391-416.
- SÁ, E. J. **Atlas Linguístico de Pernambuco (ALiPE)**. 2013. 417f. Tese (Doutorado em Letras) – Universidade Federal da Paraíba, João Pessoa, 2013.
- SALOMÃO, A. C. B. Variação e mudança linguística: panorama e perspectivas da sociolinguística variacionista no Brasil. **Fórum Linguístico**, Florianópolis, v. 8, n. 2, p. 187-207, jul./dez. 2011.
- SANTOS, C. A. N. Crenças e atitudes de variedades linguísticas no nordeste brasileiro: como universitários de uma zona de fronteira dialetal percebem a língua portuguesa? **Leitura, [S. l.]**, v. 2, n. 59, p. 168-184, 2018.
- SCHERRE, M. M. P. **A regra de concordância de número no sintagma nominal em português**. 1978. 158f. Dissertação (Mestrado em Letras) – Pontifícia Universidade Católica, Rio de Janeiro, 1978.
- SCHERRE, M. M. P.; RONCARATI, C. Programa de Estudos sobre o Uso da Língua (PEUL): origens e trajetórias. *In*: VOTRE, S.; RONCARATI, C. (Orgs.). **Anthony Julius Naro e a linguística no Brasil: uma homenagem acadêmica**. Rio de Janeiro: 7Letras, 2008. p. 37-49.
- SOUZA, A. S. **As estruturas de negação em uma comunidade rural afrobrasileira**: Helvécia-BA, 2004. Disponível em <http://www.vertentes.ufba.br/souza.doc>. Acesso em: 18 set. 2006.
- SOUZA, N. C. C.; ARAÚJO, S. S. F. Reações subjetivas de estudantes da cidade de Tucano-BA às variantes não padrão da concordância verbal. **Revista a Cor das Letras**, Feira de Santana, v. 21, n.1, p. 249-272, jan.-abr. 2020.
- SILVA, M. B. **Atlas Linguístico do centro-oeste Potiguar**. 2012. 329f. Tese (Doutorado em Linguística) – Departamento de Letras Vernáculas, Programa de Pós-graduação em Linguística, Universidade Federal do Ceará, Fortaleza.

- TARALLO, F. Por uma sociolinguística românica ‘paramétrica’: fonologia e sintaxe. **Ensaio lingüísticos**, 13, p.51-83, 1987.
- TARALLO, F. Diagnosticando uma gramática brasileira: o português d’aquém e d’além mar ao final do século XIX. In: ROBERTS, I.; KATO, M. (Orgs.). **Português brasileiro: uma viagem diacrônica**. Campinas: Editora da UNICAMP, 1993. p. 69-105.
- VITÓRIO, E. G. S. L. H. Avaliação social e a concordância verbal com o pronome Tu. **Diadorim**, Rio de Janeiro, vol. 23, n. 1, p. 151-170, jan.-jun. 2021.
- VOTRE, S. J. **Aspectos da variação fonológica na fala do Rio de Janeiro**. 1978. 221f. Tese (Doutorado em Letras) – Pontifícia Universidade Católica, Rio de Janeiro, 1978.
- VOTRE, S.; RONCARATI, C. (Orgs.). **Anthony Julius Naro e a linguística no Brasil: uma homenagem acadêmica**. Rio de Janeiro: 7Letras, 2008.



PARTE II

**ESTUDOS DE PRODUÇÃO E PERCEPÇÃO
SOCIOLINGUÍSTICAS**



CONCORDÂNCIA VERBAL COM AS FORMAS PRONOMINAIS NÓS E A GENTE: O EFEITO DA VARIÁVEL TEMPO VERBAL SOBRE A MANUTENÇÃO DA FORMA NÓS COM CONCORDÂNCIA PADRÃO¹

HEBE MACEDO DE CARVALHO
HUGO LEONARDO GOMES DOS SANTOS

1 INTRODUÇÃO

Este estudo se situa no conjunto de outras pesquisas variacionistas desenvolvidas por diferentes grupos de diferentes regiões do país (cf. VIANNA; LOPES, 2015, p. 109), sobre o uso variável das formas pronominais *nós/a gente*, fenômeno amplamente estudado no português brasileiro (NARO; GÖRSKI; FERNANDES, 1983; OMENA, 1996, 2003; ALBÁN; FREITAS, 1991; FERNANDES; GÖRSKI, 1986; MONTEIRO, 1990, 1994; FREITAS, 1991; LOPES, 1993, 1998, 2003; FERNANDES, 1997; NARO; GÖRSKI; FERNANDES, 1999; MENDES, 2007; ZILLES, 2007; LUCCHESI, 2009; RAMOS; BEZERRA; ROCHA, 2009; RUBIO, 2012; MATTOS, 2013; FOEGER, 2014; ARAÚJO, 2016; FREITAG, 2016; VITÓRIO, 2017; FEITOSA; VITÓRIO, 2018; SILVA, 2020)². Ao longo do desenvolvimento dos estudos dedicados à alternância pronominal, à luz da teoria da variação e mudança linguística (LABOV, 2008 [1972]), destacamos como “achados” desses trabalhos o encaixamento linguístico do sintagma nominal *a gente* como pronome pessoal de primeira pessoa do plural, tendo em vista a assimilação de traço de pessoa e semântica de número [+EU, +PL], quando em competição com a forma canônica *nós* [+EU, +PL]. Destacamos ainda o rearranjo na morfossintaxe da concordância verbal em que a forma inovadora (*a gente*) concorda com verbo na 3ª pessoa do singular, parâmetro mais frequente nas bases de dados, incluindo a ampla expansão dessa forma

¹ Participaram da pesquisa as bolsistas de IC Maylle Lima Freitas (PIBIC/UFC) e Adelaide Cristina Sousa de Vasconcelos (PIBIC/CNPq), graduadas em Letras/ Português com habilitação em Espanhol e respectivas literaturas. Agradecemos imensamente o trabalho competente e responsável dedicado ao Projeto *O uso variável das formas nós/a gente no falar de Fortaleza (fase III)* – PIBIC/2020-2021.

² Para citar apenas alguns estudos a que tivemos acesso durante a pesquisa. Certamente, estudos igualmente relevantes ficaram fora dessa relação. Destacamos, ainda, que alguns desses trabalhos versam sobre o fenômeno da alternância *nós/a gente*; outros, sobre a concordância verbal com esses sujeitos pronominais.

pronominal, em termos de frequência, na fala urbana brasileira (FERNANDES, 1997; ZILLES, 2007; CALLOU; LOPES, 2004 citadas por VIANNA; LOPES, 2015; ARAÚJO, 2018; SILVA, 2020; CARVALHO; FREITAS; FAVACHO, 2020).

Nosso objetivo, neste texto, é apresentar resultados da atuação da variável *tempo e tipo de paradigma verbal* sob a variação da concordância verbal com as formas pronominais de primeira pessoa do plural (1PP), na fala de Fortaleza da década de 1980, época em que se tem notícia das primeiras investigações sobre a alternância pronominal *nós/a gente*, fenômeno que tem provocado novos arranjos morfológicos e morfossintáticos no sistema pronominal do português do Brasil. Os dados de fala de *nós/a gente* na função de sujeito explícito e implícito são da base de dados do projeto Dialectos Sociais Cearenses (DSC). Seguem excertos de fala com *supertokens* que ilustram a variantes sob controle no estudo.

a. *nós* com o verbo na primeira pessoa do plural (*nós* com *-mos*) e *nós* com verbo na terceira pessoa do plural (*nós* sem *-mos*)

- (1) **nós fomos** de ônibus, foi da igreja, INF: e aí' (...) **nós ia** subindo SERra' (+) (mulher, faixa 1, primário, nor)³.
- (2) **nós** hoje **tamos** melhor' agradecemos o Padre Batista' a ajuda de Dona Nazaré Paubica' foi quem botou essa luz pra nos alumiar' que **nós** num **tinha'** **nós morava** no escuro' (+) **nós** num **tinha** nada (+) na vida' (+) só ϕ **tinha** mesmo o filho e o escuridão... (mulher, faixa 4, analfabeta, fca).

b. *a gente* com verbo na terceira pessoa do plural (*a gente* sem *-mos*) e *a gente* com verbo na primeira pessoa do plural (*a gente* com *-mos*)

- (3) **a gente/** (+) **num tem** essas palhas' assim' **a gente corta** e fica só o::: o coisa né'' **a gente vai corta'** **corta'** aí faz um cavalinho, **corta** assim ó' aí **faz** as orelhas e aqui no coisa (...) no coisa que fica assim, INF: é **a gente corta** asSIM, aí **a gente faz** as orelhas' aí faz assim' ele num é assim' ó, INF: aí pinta, INF: é, aí pinta assim de carvão' sabe'' aí **a gente vamo** brincar de cavalo (...) (homem, faixa 1, primário, fco.jean).
- (4) eu conheci ele né, lá no clube, **a gente começou** a dançar eu e ele... (homem, faixa 3, segundo.grau, joao.bosc).

As ocorrências (1), (2), (3) e (4) foram retiradas de entrevistas sociolinguísticas dos dezoito participantes do DSC, estratificados por sexo, faixa etária e com escolaridade até o segundo grau⁴ (SANTOS; VIANA; ARAÚJO, 2021), que estão transcritas e dispo-

³ As informações em parênteses são da estratificação social do participante da pesquisa adotada com base no corpus de dados de fala do projeto DSC, sendo, nessa ordem: sexo do falante (mulher ou homem), faixa etária (faixa 1: 10-14 anos, faixa 2: 16-19 anos, faixa 3: 20-24 anos e faixa 4: 37-42 anos), nível de escolaridade (analfabeto(a): sem escolarização, primário: cursou até o 4º ano do Fundamental I, ginásio: cursou o Fundamental II e segundo grau: cursou o Ensino Médio, fazendo a devida equiparação com a nomenclatura atual paras os níveis de ensino) e, por fim, a identificação do participante na amostra.

⁴ Denominação adotada no projeto para o que hoje poderia corresponder ao chamado ensino médio.

níveis no *site* do projeto Profala⁵. Inicialmente, foram consideradas quatro variantes, ilustradas anteriormente, para efeito de seleção, coleta e quantificação de dados.

2 O CONTEXTO DA PESQUISA

Como já dito, os dados de fala deste estudo⁶ foram coletados na década de 1980 e nosso objetivo é apresentar resultados da variável *tempo e tipo de paradigma verbal no modo indicativo*, replicada do trabalho de Scherre, Yacovenco e Naro (2018), o estudo-fonte para esta pesquisa. A hipótese alternativa pressupõe o efeito estatístico do tempo-modo indicativo sob o uso de *nós* e *a gente*, e as respectivas concordâncias verbais.

Constituem fatores dessa variável os tempos verbais presente, pretérito perfeito e pretérito imperfeito do modo indicativo, com controle dos verbos que têm a mesma morfologia para o presente e o pretérito perfeito do modo indicativo na 1PP (verbos ambíguos – ocorrência 5), bem como verbos com morfologia diferente para esses tempos quando flexionados na 1ª pessoa do plural (verbos não ambíguos – ocorrência 6). A variável testada pelos autores controla também a escala da saliência fônica na relação singular/plural dos verbos para a 1ª pessoa do plural, porém, para efeito deste estudo, a saliência não será contemplada.

- (5) **nós chegamos** aqui em cinquenta e quatro... (mulher, faixa 4, segundo grau, mar. ivon).
- (6) **nós tivemos** que voltar pra cá, por causa que a gente num podia pagar aluguel de casa... (mulher, faixa 4, segundo grau, mar. ivon).

Verbos do tipo *chegar* (cf. ocorrência 5), *cantar*, *dormir*, *vender*, quando flexionados na 1ª pessoa do plural, apresentam formas homônimas para o presente e o pretérito perfeito do indicativo. Observe-se que a expressão de passado está indicada pelo ano em que ocorre o evento *chegar* — o ano de cinquenta e quatro. A ocorrência (6) ilustra formas verbais com morfologias diferentes para o presente e para o pretérito perfeito (*temos/tivemos*). A razão para categorizar essa variável controlando esses contextos, segundo Scherre, Yacovenco e Naro (2018), dá-se com base em estudos de Naro, Fernandes e Görski (1999), de Naro *et al.* (2017)⁷ e de Scherre *et al.* (2014)⁸, que têm apontado para o uso do pronome *nós* com *-mos* associado ao tempo pretérito perfeito do indicativo, bem como a tendência, sugerida por Naro, Görski e Fernandes (1999), na concordância variável *nós* ou com *-mos* (*nós comemos*) vs *nós* sem *-mos* (*nós come*), de “resolução da ambiguidade potencial entre pretérito perfeito (*nós dormimos*) e pre-

⁵ O projeto DSC foi coordenado inicialmente pela professora Cláudia Roncarati, com o apoio do professor José Rogério Fontenele Bessa, coordenador do Núcleo de Pesquisas Linguísticas – NUPEL, sediado na Universidade Federal do Ceará. Mais detalhes sobre o histórico do projeto podem ser consultados na página do projeto Profala: <https://profala.ufc.br/pt/projeto-dialetos-sociais-cearenses/>, coordenado pela profa. Maria Elias Soares.

⁶ Os resultados aqui apresentados são referentes ao projeto Pibic/UFC – 2020-2021, intitulado *O uso variável das formas nós/a gente no falar de Fortaleza* (fase III).

⁷ Citados por Scherre, Yacovenco e Naro (2018).

⁸ Citados por Scherre, Yacovenco e Naro (2018).

sente igual (*nós dormimos*), com uso preferencial de *-mos* para o pretérito perfeito (*nós dormimos*) e de sua ausência para o presente (*nós dorme*)” (SCHERRE; YACOVENCO; NARO, 2018, p. 14).

Replicar variáveis e adotar critérios de análise similares constituem prática recorrente da pesquisa empírica em sociolinguística (GUY; ZILLES, 2007, p. 23). Foram quantificadas 374 observações de *nós* e *a gente* com concordância padrão e concordância não padrão, em posição de sujeito explícito e implícito, submetidas à análise quantitativa tomando como base o *software* R Studio (R CORE TEAM, 2020). Controlamos as ocorrências de sujeitos implícitos quando a ausência do sujeito seguia um certo paralelismo formal. Seguem dois excertos de fala do DSC que ilustram esses contextos:

(7) aí no dia cinco de janeiro’ **a gente volta** (+) aí ϕ começa a fazer a/as provas finais... (homem, faixa 1, primário, fco.jean).

(8) aí eu’ a mamãe **nós trouxemos** ele pra casa, Morto de bêbado’ com as pernas Bamba que (+) não aguentava nem se segurar, ϕ **chegamos** lá ϕ **trouxemos’ ϕ trouxemos** ele... (mulher, faixa 4, segundo grau, mar.ivon).

Em (7), a implicitude do sujeito antes do verbo *começa* foi controlada como sendo o sujeito implícito *a gente*. Em (8), os sujeitos implícitos antes de *chegamos*, *trouxemos*, *trouxemos* é antecedido pelo *nós* enunciado anteriormente. Essas ausências foram controladas como *nós* sujeito implícito. Seguimos, com rigor, esse critério para a coleta dos sujeitos implícitos. Casos de *nós* sem *-mos* (concordância não padrão) com sujeito implícito (9) foram mais raros.

(9) nós hoje tamos melhor’ agradecemos o Padre Batista’ a ajuda de Dona Nazaré Paubica’ foi quem botou essa luz pra nos alumiar’ que **nós num tinha’ nós morava** no escuro’ (+) **nós num tinha** nada (+) na vida’ (+) só ϕ **tinha** mesmo o filho e o escuridão... (mulher, faixa 4, analfabeta, fca).

No excerto de fala (9), a não explicitude do sujeito antes do verbo *tinha* foi controlada como *nós* sujeito implícito, tendo em vista que a configuração das estruturas antecedentes apresenta sujeitos explícitos *nós*.

3 METODOLOGIA

A base de dados deste estudo, como já informado, é constitutiva do projeto Dialetos Sociais Cearenses (DSC), composto por dezoito falantes cujas entrevistas foram realizadas na década de 1980 (SANTOS; VIANA; ARAÚJO, 2021). A sua primeira fase tem culminância no final da década, em 1988⁹. A estratificação da amostra desta pesquisa pode ser conferida no Quadro 1.

⁹ Detalhes da história do projeto DSC podem ser conferidos na página: <https://profala.ufc.br/pt/projeto-dialetos-sociais-cearenses/historico/>. Acesso em 29 abr. 2022.

Quadro 1: Estratificação da amostra por faixa etária e sexo

Faixa etária	Sexo		Participantes
	Masculino	Feminino	
10-14 anos	inf. 2, inf.7	inf. 1, inf. 4, inf.8	5
16-19 anos	inf. 9, inf.14	inf.17	3
20-24 anos	inf. 13, inf.10	inf.11, inf. 12, inf. 15, inf.16	6
37-42 anos	inf. 6	inf. 3, inf.5, inf.18	4
Total	7	11	18

Fonte: Adaptado do Projeto DSC.

Conforme o Quadro 1, há células sociais com 01 participante, a exemplo das células da faixa etária de 16-19 anos, sexo feminino, e 37-42 anos, sexo masculino. A amostra considerou também participantes *sem escolaridade* (uma mulher), com o *primário*¹⁰ (03 participantes na faixa etária de 10-15 anos), com o *ginásio* (09 participantes distribuídos por todas as faixas etárias da amostra) e com o *segundo grau* (05 participantes com faixas etárias de 16 a 42 anos). O instrumento de coleta de dados consiste em entrevistas semiestruturadas, nos moldes labovianos.

Os dados já transcritos foram, após seleção e coleta, transportados para uma planilha .csv. Em seguida, para tratamento estatístico, análises descritiva e inferencial e elaboração dos gráficos, foi adotada a linguagem R (R CORE TEAM, 2020), por meio do ambiente RStudio. As seguintes variáveis dependentes e independentes foram consideradas para a testagem de hipóteses.

Quadro 2: Variáveis adotadas no estudo

	Variável	Níveis
Variável dependente (VD)	Ocorrências	<i>nós</i> com -mos <i>nós</i> sem -mos <i>a gente</i> com -mos <i>a gente</i> sem -mos
Variável independente	Tempo e tipo e paradigma verbal no modo indicativo	(1) Presente igual ao pretérito perfeito do indicativo (2) Presente diferente do pretérito perfeito do indicativo (3) Pretérito perfeito igual ao presente do indicativo (4) Pretérito perfeito diferente do presente do indicativo (5) Pretérito imperfeito do indicativo

Fonte: elaboração própria.

¹⁰ A base de dados do projeto DSC, constituída nos anos 1980, adota os termos “primário”, “ginásio” e “segundo grau” para a escolarização dos participantes. O primário remete aos quatro primeiros anos escolares, ou seja, da 1ª à 4ª série escolar. O ginásio compreende da 5ª à 8ª série, e o segundo grau equivale ao que hoje é denominado ensino médio.

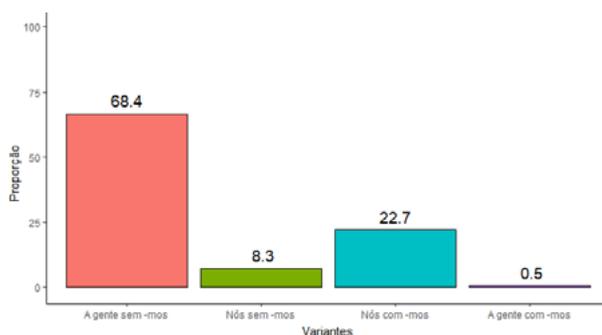
A análise descritiva contempla a VD em quatro níveis. Em decorrência do baixo número de dados, o estudo se desenvolveu com base em dois níveis (*nós* com *-mos*; *a gente* sem *-mos*) — isso pode ser conferido na seção 4: resultados e discussão. A variável linguística *tempo e tipo e paradigma verbal no modo indicativo* foi testada, nesta pesquisa, considerando a dimensão estrutural/funcional dessa variável. Não entrou em jogo, portanto, a saliência fônica, “dimensão cognitiva/perceptual de formas em competição” (cf. SCHERRE; NARO; YACOVENCO, 2018, p. 444). Nosso enfoque é o papel funcional do tempo na concordância verbal com as formas pronominais de 1PP. A escala da proeminência deverá ser tratada, nestes dados, em outra etapa do estudo.

A variável independente é composta por 05 fatores: (1) presente igual ao pretérito perfeito do indicativo (*nós cantamos*; *nós canta*; *a gente cantamos*; *a gente canta*); (2) presente diferente do pretérito perfeito do indicativo (*nós vamos*; *nós vai*; *a gente vamos*; *a gente vai*); (3) pretérito perfeito igual ao presente do indicativo (*nós cantamos*; *nós cantou*; *a gente cantamos*; *a gente cantou*); (4) pretérito perfeito diferente do presente do indicativo (*nós fomos*; *nós foi*; *a gente foi*); (5) pretérito imperfeito do indicativo (*nós cantávamos*; *nós cantava*; *a gente cantávamos*; *a gente cantava*).

4 RESULTADOS E DISCUSSÃO

Esta seção desenvolve os resultados do estudo sobre a fala de Fortaleza. Serão apresentados resultados gerais em termos de proporção e de frequência de uso das variantes *nós* e *a gente*, e as respectivas concordâncias verbais, assim como os resultados de um modelo de regressão logística binomial. Foi adotada a linguagem R (R CORE TEAM, 2020) por meio do ambiente RStudio, que fornece resultados dos efeitos estatísticos de variáveis independentes sobre a variável dependente (doravante VD) — a expressão de 1ª pessoa do plural (*nós/a gente*) e respectivas concordâncias verbais. O número de observações inicialmente foi 374. O Gráfico 1, a seguir, apresenta as proporções das variantes sob controle na pesquisa.

Gráfico 1: Proporções das variantes *nós* com *-mos* e *a gente* sem *-mos* e respectivas concordâncias padrão e não padrão na fala de Fortaleza da década de 1980 – DSC



Fonte: elaboração própria.

Em termos totais, das 374 observações, os resultados apontam 68,4% (n = 256) da forma *a gente sem -mos* (*a gente canta/a gente brinca*), 22,7% (n = 86) de *nós com -mos* (*nós cantamos/nós brincamos*), 8,3% (n = 31) de *nós sem -mos* (*nós canta/nós brinca*) e 0,53% (n = 2) de *a gente com -mos* (*a gente cantamos/a gente brincamos*).

As duas ocorrências de *a gente com -mos* podem ser conferidas na sequência.

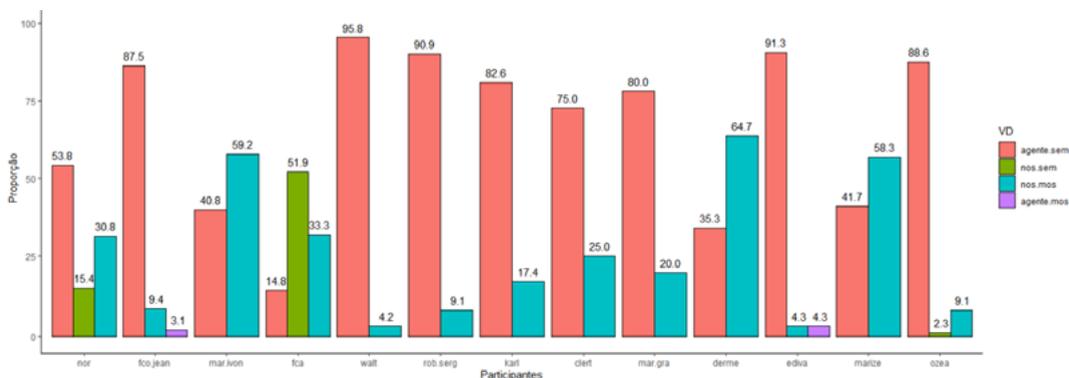
(10) *é, aí pinta assim de carvão' sabe'' aí a gente vamo brincar de cavalo.* (homem, faixa 1, primário, fco.jean).

(11) *a gente voltou ϕ fizemos* assim uma missa (mulher, faixa 3, ginásio, ediva).

Os resultados (cf. Gráfico 1) da proporção das variantes indicam que a forma *a gente sem -mos* como a mais frequente na amostra (n = 256/ 68,4%), seguindo a tendência de estudos prévios (VIANNA; LOPES, 2015; ARAÚJO, 2016) sobre o português brasileiro, o que confirma a expansão dessa variante na fala brasileira. Dados da fala de graduados fortalezenses também seguem essa tendência (CARVALHO; FREITAS; FAVACHO, 2020).

Do total de 31 ocorrências de concordância não padrão, observamos esse uso em apenas cinco informantes da amostra. O uso de *nós sem -mos* (*nós canta/cantou/cantava*) ocorre nos dados de fala de três participantes: “Nor” (15,39%, n = 2/13), “Ozea” (2,27%, n = 1/44) e “Fca” (51,85%, n = 28/54). A proporção maior de *nós sem -mos* concentra-se na única participante sem escolaridade da amostra, com idade de 42 anos, profissão operadora de chafariz. A estrutura *a gente com -mos* foi produzida por “Fco.jean” (3,12%, n = 1/32) e “Ediva” (4,34%, n = 1/23), falantes com escolaridade primária e ginásial respectivamente. Considerando o baixíssimo número de ocorrências com concordância não padrão, entendemos que a projeção desses usos para a comunidade depende de uma quantidade maior de dados e de amostras com maior número de participantes. O Gráfico 2 mostra a distribuição por participante.

Gráfico 2: Distribuição das variantes *nós* e *a gente*, e respectivas concordâncias verbais por participante da amostra DSC

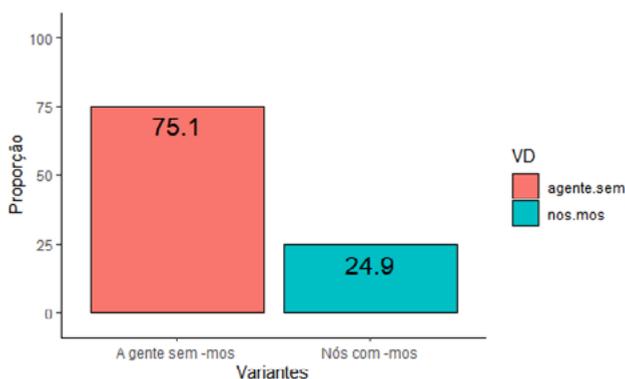


Fonte: elaboração própria.

Dado o baixo número de ocorrências da concordância verbal do *nós* com o verbo na 3PS – 30 dados (*nós cantou/brincou*) e do *a gente* com o verbo na 1PP – 02 dados (ocorrências 10 e 11), a análise desenvolvida passa a ser binária com duas variantes linguísticas: *nós* com *-mos* (padrão antigo) e *a gente* sem *-mos* (padrão emergente).

A estrutura emergente com concordância padrão - *a gente* sem *-mos* - foi a mais frequente (n = 256/341, proporção = 75,1 %), sendo o *nós* com *-mos* (n = 85/341) menos frequente, com proporção de 24,9% do total dos dados. O Gráfico 3 de barras mostra os resultados da análise binária (*nós* com *-mos* vs. *a gente* sem *-mos*).

Gráfico 3: Proporção de uso de *nós* com *-mos* e *a gente* sem *-mos* na fala de Fortaleza dos anos 1980, amostra DSC (N = 341): análise binária



Fonte: elaboração própria.

Ao retirar as variantes *nós* sem *-mos* e *a gente* com *-mos*, o número de observações passa a ser 341. Do ponto de vista da alternância das formas pronominais, os resultados expostos no Gráfico 3 apontam a expansão da forma *a gente* na função de sujeito, tendência atestada no falar de cidades do Nordeste brasileiro como João Pessoa (PB), Cacém (PE) e Fortaleza (CE) — dados do Projeto Norma Oral Popular de Fortaleza – Norpofof (cf. ARAÚJO, 2018, p. 200). A forma *nós* ocorreu na proporção de 24,9% dos dados. Diante desse cenário, a variante de referência para a regressão logística é o *nós* com *-mos* vs. *a gente* sem *-mos*. Essa escolha é motivada com o objetivo de rastrear contextos de restrição de *nós* com *-mos* (estrutura canônica) em função da variável tempo e tipo de paradigma verbal, variável previsora no modelo de regressão logística. A tabela 1 indica a proporção geral com o filtro/retirada dos dados de *nós* sem *-mos* (n = 31) e *a gente* com *-mos* (n = 2), dados de concordância não padrão. Foram retiradas ainda ocorrências de *a gente* (n = 26) e de *nós* (n = 3) associadas às formas nominais dos verbos, totalizando 312 observações. A tabela 1 mostra os percentuais da rodada sem essas estruturas. Os excertos de fala a seguir ilustram as ocorrências com formas nominais do verbo.

(12) é legal **a gente contar** nossos problemas (homem, faixa 1, primário, rob.serg).

(13) **a gente prestando** a atenção (mulher, faixa 1, ginásio, karl).

- (14) pegar um emprego pra **nós trabalhar** (mulher, faixa 4, analfabeta, fca).
 (15) agradecemos **nós viver** dentro da igreja (mulher, faixa 4, analfabeta, fca).

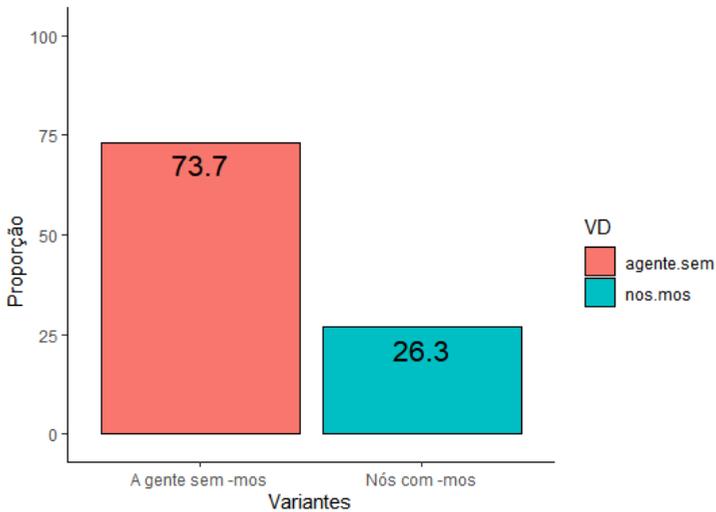
Tabela 1: Proporção do uso de *nós* com *-mos* e *a gente* sem *-mos* na fala de Fortaleza dos anos 1980, amostra DSC (N = 312): análise binária

Variante	N	Proporção
<i>A gente</i> sem <i>-mos</i>	230	73,7%
<i>Nós</i> com <i>-mos</i>	82	26,3%

Fonte: elaboração própria.

Observe-se que houve um leve aumento da proporção da estrutura canônica *- nós* com *-mos* de 24,9% (Gráfico 3) para 26,3%. Esses dados são apresentados no Gráfico 4.

Gráfico 4: Proporção do uso de *nós* com *-mos* e *a gente* sem *-mos* na fala de Fortaleza dos anos 1980, amostra DSC (N = 312): análise binária



Fonte: elaboração própria.

A análise de regressão logística tomou como parâmetro esses números. Ao mensurar a variável linguística tempo e tipo de paradigma do verbo (cf. Tabela 2), foram considerados cinco tempos verbais do modo indicativo. O tempo presente com morfologia igual para o pretérito perfeito do indicativo mostrou-se associado ao *a gente* sem *-mos* (92,3%, n = 84), desfavorecendo *nós* com *-mos* (7,7%, n = 7). Do mesmo modo, o presente diferente do pretérito perfeito do indicativo mostrou-se associado à construção *a gente* sem *-mos* (74,3%, n = 78) em detrimento do *nós* com *-mos* (25,7%, n = 27). Os resultados para o tempo verbal pretérito perfeito do indicativo com morfologia igual ao presente (formas ambíguas, por exemplo: *andar/comer/dormir*) sugerem um equilí-

brio de uso entre *nós* com *-mos* (52%, n = 26) e *a gente* sem *-mos* (48%, n = 24), sugerindo uma competição equilibrada nesse ambiente linguístico. O pretérito perfeito do indicativo com morfologia diferente do presente na primeira pessoa do plural favorece *nós* com *-mos* (57,1%, n = 20), e o imperfeito do indicativo desfavorece essa estrutura (6,5%, n = 2). Em síntese, o tempo presente mostrou-se associado ao *a gente* sem *-mos*, o pretérito perfeito parece ser um ambiente de competição entre essas construções, e o imperfeito, desfavorecedor do *nós* com *-mos*. Confirma-se a hipótese da atuação do tempo e do modo verbal do indicativo no uso da concordância verbal com *nós* e *a gente*. Os resultados, em termos de frequência (N) e proporção (%), estão na tabela 2.

Tabela 2: Proporção de *nós* com *-mos* vs. *a gente* sem *-mos* por tempo e tipo de paradigma verbal na amostra de fala de Fortaleza – DSC: análise binária

Tempos verbais do modo indicativo	Nós com -mos		A gente sem -mos	
	N	%	N	%
Presente igual ao pretérito perfeito do indicativo Nós <i>moramos/cantamos/cortamos</i> A gente <i>mora/canta/corta</i>	7/91	7,7	84/91	92,3
Presente diferente do pretérito perfeito do indicativo Nós <i>estamos/fazemos/vamos</i> A gente <i>está/faz/vai</i>	27/105	25,7	78/105	74,3
Pretérito perfeito igual ao presente do indicativo Nós <i>passamos/saimos/chegamos</i> A gente <i>passa/sai/chega</i>	26/50	52	24/50	48
Pretérito perfeito diferente do presente do indicativo Nós <i>estivemos/fomos/trouxemos</i> A gente <i>esteve/foi/trouxe</i>	20/35	57,1	15/35	42,9
Pretérito imperfeito do indicativo Nós <i>trabalhávamos/tínhamos/partíamos</i> A gente <i>trabalha/tinha/partia</i>	2/31	6,5	29/31	93,5
X ² = 56,815 (4), p < 0,001.				

Fonte: elaboração própria.

Do ponto de vista da análise descritiva, *nós* com *-mos* está associado ao tempo verbal pretérito perfeito, sendo o verbo ambíguo (52%) ou não (57,1%). Submetemos essa variável ao teste de qui-quadrado, que aponta uma significativa associação entre o tempo/modo verbal e o uso de *nós* com *-mos* vs. *a gente* sem *-mos*, apresentando o valor de 56,815, com quatro graus de liberdade (cinco fatores/tempos verbais) e p-valor menor que 0,001. Esse p-valor indica que a probabilidade de a variável tempo e tipo de paradigma verbal exercer influência sobre a VD (*nós* com *-mos* e *a gente* sem *-mos*) é bastante alta, sinalizando ser significativa para a realização desta variável linguística: a concordância verbal com formas pronominais de 1PP. Considerando o número de

observações da estrutura *nós* com *-mos* por fatores do tempo pretérito perfeito igual (n = 26) e pretérito perfeito diferente (n = 20) e as proporções e o fato de expressarem funcionalmente a mesma referência temporal, essa variável previsora foi amalgamada. Os resultados podem ser conferidos na tabela 3.

Tabela 3: Proporção de *nós* com *-mos* vs. *a gente* sem *-mos* por tempo e modo verbal na amostra de fala de Fortaleza – DSC

Tempos verbais	<i>Nós com -mos</i>		<i>A gente sem -mos</i>	
	N	%	N	%
Presente do indicativo igual ao pretérito perfeito do indicativo <i>Nós moramos/cantamos/cortamos</i> <i>A gente mora/canta/corta</i>	7	7,7	84	92,3
Presente do indicativo diferente do pretérito perfeito do indicativo <i>Nós estamos/fazemos/vamos</i> <i>A gente está/faz/vai</i>	27	25,7	78	74,3
Pretérito perfeito do indicativo <i>Nós passamos/saimos/estivemos/fomos</i> <i>A gente passa/sai/chega/esteve/foi</i>	46	54,1	39	45,9
Pretérito imperfeito do indicativo <i>Nós trabalhávamos/tínhamos/partíamos</i> <i>A gente trabalha/tinha/partia</i>	2	6,5	29	93,5
X ² = 56,534 (3), p < 0,001.				

Fonte: elaboração própria.

O teste de qui-quadrado é um indicador de que a escolha de amalgamar os dois pretéritos gera uma distribuição com melhor potencial explicativo. O valor de chi-quadrado não sofreu grande alteração, mesmo com a diminuição dos graus de liberdade da variável (de 5 fatores/tempos verbais para 4 fatores/tempos verbais). Por sua vez, o p-valor se manteve abaixo de 0,001, indicando novamente a alta probabilidade de essa variável exercer influência sobre a VD. Do ponto de vista da alternância e da concordância verbal, o pronome *a gente* sem *-mos* parece estar associado predominantemente ao presente igual ao pretérito perfeito (92,3%) e ao pretérito imperfeito (93,5%). Note-se que a frequência (N) de uso de *nós* com *-mos* associado ao fator presente igual é de apenas 7 (sete) observações, contexto de ambiguidade entre o presente e pretérito perfeito. Nesses ambientes, *nós* com *-mos* é desfavorecido, tendência atestada no estudo fonte de Scherre, Yacovenco e Naro (2018, p. 21).

Esses indicadores, mostrados a partir dos gráficos e tabelas exploratórios, fornecem uma descrição da distribuição dos dados. De posse desses resultados, realizamos a análise multivariada de regressão logística, com o objetivo de “verificar o efeito simultâneo de múltiplas variáveis predictoras, a fim de chegar a um modelo para descrever, explicar e prever o comportamento da variável resposta” (OUSHIRO,

2017) — variável dependente. A tabela seguinte apresenta os resultados da rodada de regressão logística tomando como variante de referência o *nós com -mos* para a geração das estimativas de probabilidades de sua ocorrência, associada aos fatores de cada variável controlada.

Tabela 4: Resultados do modelo de regressão logística binomial para a variante *nós com -mos* vs. *a gente sem -mos*: amostra DSC

Coefficientes	Estimativa	Erro padrão	Valor z	Pr(> z)	
Intercepto	-2,1937	0,5503	-3,986	6,71e-05	***
Tempo verbal					
Presente diferente	1,3788	0,4943	2,789	0,00528	**
Preterito perfeito	2,3578	0,5087	4,635	3,58e-06	***
Preterito imperfeito do indicativo	-1,1919	0,8807	-1,353	0,17591	
Referência					
Eu + ele.s + ela.s determinado	0,4454	0,4025	-1,107	0,26847	
Genérico	-1,6354	0,5586	-2,928	0,00341	**
Eu	-1,6369	1,0979	-1,491	0,13598	
Sexo					
Masculino	-1,0341	0,4019	-2,573	0,01008	*
Escolaridade					
2º Grau	-0,8120	0,5523	-1,470	0,14149	
Ginásio	0,1500	0,4263	0,352	0,72490	
Faixa etária					
Segunda faixa	-0,24136	0,5113	-0,472	0,63696	
Terceira faixa	1,4397	0,4596	3,133	0,001731	**
Códigos de significância: 0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '' 1					
VD ~ TEMPOVERBAL + REFERÊNCIA DO PRONOME + SEXO + ESCOLARIDADE + FAIXA.ETARIA.NOVA					

Fonte: elaboração própria.

Observe-se que a tabela apresenta os resultados de todas as variáveis controladas no estudo, quais sejam: *tempo e tipo de paradigma verbal*, *determinação do referente do pronome*, *sexo*, *escolaridade* e *faixa etária*. Contudo, apenas a variável linguística *tempo e tipo de paradigma verbal* será objeto de análise e descrição.

O intercepto apresenta a probabilidade de ocorrência da variante de referência (*nós com -mos*), tomando por base o primeiro fator das variáveis previsoras. Nesse

caso, os valores do intercepto (Tabela 4) remetem ao uso de *nós* com *-mos*, associado aos seguintes fatores:

- (I) verbos no presente igual ao pretérito perfeito do indicativo;
- (II) o pronome *nós* com referência indeterminada¹¹;
- (III) sexo feminino;
- (IV) escolarização primária; e
- (V) primeira faixa etária.

A variável previsora *tempo e tipo de paradigma verbal* é selecionada significativamente no uso da variável em estudo, mantendo a significância¹², observada no estudo que nos serviu de ponto de partida (SCHERRE; YACOVENCO; NARO, 2018), de seleção desta variável. Para mensurar a atuação dos fatores dessa variável linguística sobre a variável resposta, é necessário somar o valor de estimativa de algum de seus fatores ao valor de estimativa do intercepto. Por exemplo, a estimativa do fator presente diferente (1,3788) é somada à estimativa do intercepto (-2,1937 - cf. Tabela 4), resultando no valor de estimativa -0,8149. Como esse valor é maior do que a estimativa do intercepto, o fator presente diferente favorece o uso de *nós* com *-mos* quando comparado ao fator presente igual do intercepto.

Aplicando esse raciocínio aos demais fatores, nota-se que a estrutura padrão *nós* com *-mos* tem menor probabilidade de ocorrer com verbos no presente igual ao pretérito (cf. intercepto -2,1937) e com verbos no pretérito imperfeito (-3,3859), em concordância com a tendência atestada nas amostras de Vitória - ES e da Baixada Cuibana -CO, demonstrada por Scherre, Yacovenco e Naro (2018, p. 21), estudo de onde partimos para estruturar essa variável. Nos casos de presente igual tem-se a ambiguidade potencial, note que nesses contextos o *nós* com *-mos* (*nós cantamos*) cede espaço para a entrada de *a gente* sem *-mos* (*a gente canta*). Como dizem as autoras e o autor do estudo, a ausência de *-mos*, é marca de presente. A entrada do *a gente* sem *-mos* associada ao imperfeito do indicativo implica uma questão de saliência, o que sugere a “esquiva da proparoxítona com o morfema de plural *-mos* (*falávamos*), motivada pelo padrão fonológico preferencialmente paroxítono do português brasileiro” (SCHERRE; YACOVENCO; NARO, 2018, p. 22). Esse ambiente é favorecedor da ausência de *-mos*.

A probabilidade aumenta com verbos no presente diferente do pretérito (- 0,8149) e atinge seu maior valor com verbos no perfeito do indicativo (0,1641). Esses números

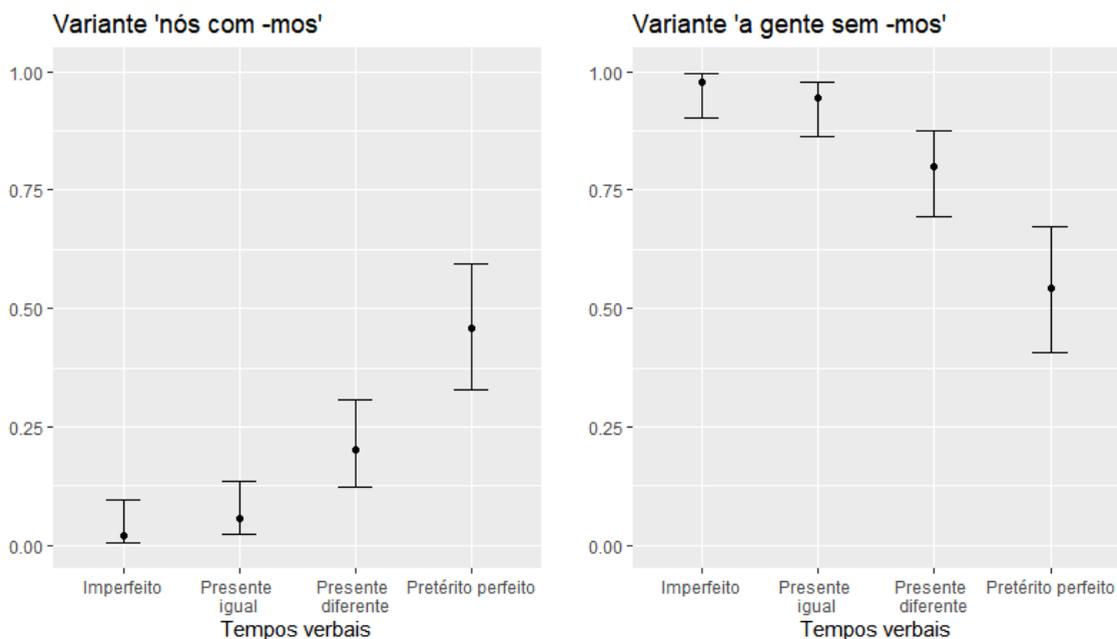
¹¹ O fator *eu + ele.s + ela.s determinado* remete aos casos de referência genérica dos pronomes — a referência ao falante (*eu*) + eles/não-pessoa indeterminada, remetendo a um grupo relativamente indeterminado de pessoas. A ocorrência seguinte ilustra esse contexto linguístico: “(...) mas aí tem a lei’ essa lei (+) que veio agora’ né” do Sarney, que ele falou que ia dar um décimo terceiro aos pensionistas, os pensionista e aposentados, (+) aí eu’ eu acho que *a gente* vai pegar, é (+) um auMENto’ porque a gente tá assim (+) em defasagem’ né” [*nós* somos pensionistas, *nós* somos/ *nós tamos* em defasagem’ então a gente tem quer recorrer, (+) pra ficar recebendo o um reajuste” (mulher, faixa 4, segundo grau, mar.ivon). Note-se que a falante faz referência a um grupo de pessoas pensionistas no qual ela se inclui.

¹² Destacamos que essa variável, aqui replicada, foi testada no GoldVarb X no estudo fonte, sendo selecionada significativamente também ao ser submetida aos moldes da linguagem R/RStudio, o que indica a forte significância dessa categoria sob o fenômeno em estudo.

sugerem que o pretérito perfeito do indicativo e o presente diferente do pretérito se configuram como ambientes linguísticos favorecedores da manutenção da estrutura padrão - *nós* com -*mos*. Note-se que, nesses casos de presente, não há ambiguidade entre presente do indicativo (*fazemos/vamos/estamos*) e pretérito perfeito (*fizemos/fomos/estivemos*), ou seja, “-*mos* pode ser usado sem possibilidade de interpretação de pretérito perfeito” (SCHERRE; YACOVENCO; NARO, 2018, p. 22).

O gráfico 5 mostra os efeitos dos fatores da variável tempo e tipo de paradigma verbal sobre a ocorrência de *nós* com -*mos* e de *a gente* sem -*mos*.

Gráfico 5: Efeito da variável tempo e tipo de paradigma verbal em função de *nós* com -*mos* e *a gente* sem -*mos*: amostra DSC – Fortaleza (CE)



Fonte: elaboração própria.

O gráfico apresenta, por meio dos pontos, a probabilidade de uso de cada variante associada aos tempos verbais controlados no estudo. As barras verticais indicam o intervalo de confiança, sugerindo uma projeção das probabilidades para a comunidade. Olhando apenas para os resultados da variante *nós* com -*mos*, observa-se que a estimativa de ela ocorrer concentra-se preferencialmente no pretérito perfeito (ver também tabela 3), resultado semelhante, em termos de tendência, aos dados de fala de Santa Leopoldina (ES) e Vitória (ES) — região Sudeste — e Baixada Cuiabana (MT) — região Centro-Oeste —, em rodada binária (*nós* com -*mos* vs. *a gente* sem -*mos*), em quatro amostras das duas primeiras décadas do século XXI (cf. SCHERRE; NARO; YACOVENCO, 2018, p. 443).

Analisando os resultados associados ao presente, é importante destacar que há pouca sobreposição entre os intervalos de confiança dos fatores ‘presente igual’ (nós recebemos/nós vendemos/colocamos) e ‘presente diferente’, indicando que os diferentes padrões de uso das variantes associados a cada fator apresentam uma significância estatística.

Ao observar os resultados referentes à variante inovadora (*a gente* sem *-mos*), nota-se que todos os tempos verbais favorecem a ocorrência de *a gente* sem *-mos*, sendo o pretérito imperfeito do indicativo um ambiente de uso quase categórico dessa forma. O pretérito perfeito do indicativo, nesse contexto, se configura como seu ambiente de menor probabilidade de ocorrência. No entanto, ao comparar os resultados desse fator para as duas variantes (gráfico 5), verifica-se a sobreposição dos intervalos de confiança indicados pelas barras verticais no fator ‘pretérito perfeito’, sugerindo ser esse um ambiente de competição das duas variantes na amostra dos dados de fala de Fortaleza da década de 1980 – DSC. Esse resultado chama a atenção quando confrontado com a tendência atestada em amostras das duas primeiras décadas do século XXI, em recorte binário (*nós* com *-mos* vs. *a gente* sem *-mos*), que aponta claramente a predominância de *nós* com *-mos* no ambiente do pretérito perfeito (cf. SCHERRE; NARO; YACOVENCO, 2018, p. 443). Essa divergência, em termos de tendência, pode ser motivada pelo recorte sincrônico, uma vez que a amostra utilizada na presente investigação controla dados da década de 1980 do século XX, período bem anterior ao estudo citado.

5 CONSIDERAÇÕES FINAIS

As análises apresentadas tiveram por objetivo apresentar a distribuição geral dos processos de alternância e de concordância verbal com as formas de 1PP (*nós* / *a gente*) com dados de fala de Fortaleza de 1980. Os resultados, em termos de proporção, indicam a ampla expansão da forma *a gente* na amostra *versus* a forma canônica *nós*. Em termos totais, predominam as estruturas de concordância padrão, sendo a estrutura *a gente* com *-mos* (*a gente cantamos/brincávamos*) rara (n = 2 dados) na base de dados e *nós* sem *-mos* (*nós canta/brincava/falou*) pouco frequente (n = 31 dados). Além de pouco frequentes, essas estruturas estão concentradas na fala de cinco participantes da amostra, o que pode ser o reflexo de uma peculiaridades deles.

A partir desse cenário, fizemos um recorte binário contemplando apenas as estruturas de concordância padrão, mais recorrentes na amostra. Com esse recorte, criamos um modelo de regressão logística e discutimos a atuação da variável tempo e tipo de paradigma verbal sobre a manutenção da estrutura canônica *nós* com *-mos*. Em linhas gerais, embora todos os tempos verbais favoreçam o uso de *a gente* sem *-mos*, destacamos o fato de o pretérito perfeito do indicativo ter se mostrado um ambiente morfossintático de competição entre as variantes, resultado divergente de outros estudos que indicam a predominância do *nós* com *-mos* nesses ambientes (SCHERRE; YACOVENCO; NARO, 2018).

Nossa hipótese é a de que as diferenças de resultados podem estar relacionadas às diferenças de constituição, social e quantitativa, das amostras, bem como do intervalo de tempo dessas bases de dados. Este estudo toma como base uma amostra de dados de fala dos anos 1980. Considerando os resultados e o tempo de constituição da amostra, o desenvolvimento de pesquisas sobre o fenômeno em análise com bases de dados atuais dessa comunidade de fala poderia lançar novos olhares em relação a tal competição.

REFERÊNCIAS

- ALBÁN, M. del R.; FREITAS, J. Nós ou a gente? *Estudos linguísticos e literários*, Salvador, UFBA, n. 11, p. 75-89, 1991.
- ARAÚJO, M. A. M. de. **Será que a gente usa mais o nós?** Uma fotografia sociolinguística do falar popular de Fortaleza. 2016. 148f. Dissertação (Mestrado em Linguística Aplicada) – Programa de Pós-Graduação em Linguística Aplicada, Universidade Estadual do Ceará. Fortaleza.
- ARAÚJO, M. A. M. de. Nós e a gente no falar dos fortalezenses. In: ARAÚJO, A. A.; VIANA, R. B. de M.; PEREIRA, M. L. de S. (Orgs.). **Fotografias sociolinguísticas do falar de Fortaleza-CE**. Fortaleza: EdUECE, 2018.
- CARVALHO, H. M. de; FREITAS, M. L.; FAVACHO, L. de L. A variação dos pronomes sujeitos nós e a gente: a fala culta de fortaleza em cena. *Revista (Con)Textos Linguísticos*, Vitória, v. 14, n. 27, p. 30-45, 2020.
- FEITOSA, J. G.; VITÓRIO, E. G. de S. L. A. Variação nós/a gente no sertão alagoano: restrição e avaliação. *Estudos Linguísticos e Filológicos*, v. 19, n. 2, 2018.
- FERNANDES, E. A. “Nós” e “a gente”: variação na cidade de João Pessoa. 1997. 117f. Dissertação (Mestrado em Língua Portuguesa) – Centro de Ciências, Letras e Artes, Universidade Federal da Paraíba João Pessoa, 1997.
- FERNANDES, E.; GÖRSKI, E. A concordância verbal com os sujeitos nós e a gente: um mecanismo do discurso em mudança. In: SIMPÓSIO SOBRE A DIVERSIDADE LINGÜÍSTICA NO BRASIL, I. *Atas...* Salvador, Instituto de Letras da UFBA, 1986. p.175-83.
- FOEGER, C. C. **A primeira pessoa do plural no português falado em Santa Leopoldina**. 2014. 165f. Dissertação (Mestrado em Estudos Linguísticos) – Programa de Pós-Graduação em Linguística, Universidade Federal do Espírito Santo, Vitória.
- FREITAG, R. M. K. Uso, crença e atitudes na variação na primeira pessoa do plural no Português Brasileiro. *DELTA: Documentação de Estudos em Linguística Teórica e Aplicada* [online], v. 32, n. 4, p. 889-917, 2016.
- FREITAS, J. “Nós” e “a gente” em elocuições formais. *Estudos Linguísticos e Literários*, Salvador: UFBA, n. 11, p. 91-102, 1991.
- GUY, Gregory Riordan; ZILLES, Ana. **Sociolinguística quantitativa: instrumental de análise**. São Paulo: Parábola Editorial, 2007.
- LABOV, William. **Padrões sociolinguísticos**. São Paulo: Parábola Editorial, 2008.
- LOPES, C. R. dos S. **A inserção de “a gente” no quadro pronominal do português**. Frankfurt/Madri: Vervuert/Iberoamericana, 2003.
- LOPES, C. R. dos S. A gramaticalização de *a gente* em português em tempo real de longa e de curta duração: retenção e mudança na especificação dos traços intrínsecos. *Fórum Linguístico*, Florianópolis, v.4, n.1, p. 47-80, jul. 2004.
- LOPES, C. R. dos S. “Nós” e “a gente” no português falado culto do Brasil. 1993. 140f. Dissertação (Mestrado em Língua Portuguesa) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro.
- LUCCHESI, D. A representação da primeira pessoa do plural. In: LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. (Orgs.). **O português afro-brasileiro**. Salvador: Ed. da UFBA, 2009. p. 457-471.

- MATTOS, S. E. R. **Goiás na primeira pessoa do plural**. 2013. 137f. Tese (Doutorado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade de Brasília, Brasília, 2013.
- MENDES, R. P. da S. **O perfil da alternância do sujeito “nós” e “a gente” em Santo Antonio de Jesus: um recorte no português popular do interior da Bahia**. 2007. 140f. Dissertação (Mestrado em Letras e Linguística) – Faculdade de Letras, Universidade Federal da Bahia, Salvador, 2007.
- MENDONÇA, A. K. **Nós e A gente em Vitória: análise sociolinguística da fala capixaba**. 2010. 98f. Dissertação (Mestrado em Estudos Linguísticos) – Programa de Pós-Graduação em Linguística, Universidade Federal do Espírito Santo, Vitória, 2010.
- MONTEIRO, J. L. **Pronomes pessoais**: subsídios para uma gramática do português. Fortaleza: Edições UFC, 1994.
- MONTEIRO, J. L. Variação no uso dos pronomes pessoais no português do Brasil. **Verba**, v. 17, p. 145-157, 1990.
- NARO, A. J.; GÖRSKI, E. M.; FERNANDES, E. Change without change. **Language Variation and Change**. n. 11, p. 197-211, 1999.
- NARO, A. J.; GÖRSKI, E.; FERNANDES, E. Uma mudança linguística em curso: a concordância com o sujeito nós/a gente. *In*: SEMINÁRIO SOBRE VARIAÇÃO EM SINTAXE. **Anais...** UFRJ, Rio de Janeiro, 1983.
- OMENA, N. P. de. A referência à primeira pessoa do plural. *In*: SILVA, G. M. de O.; SCHERRE, M. M. P.(Orgs.). **Padrões Sociolinguísticos**: análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro. Rio de Janeiro: Tempo Brasileiro, 1996. p. 183-216.
- OMENA, N. P. de; BRAGA, M. L. A gente está se gramaticalizando? *In*: MACEDO, A. T. de; RONCARATI, C.; MOLLICA, M. C. (Orgs.) **Variação e discurso**. Rio de Janeiro: Tempo Brasileiro, 1996.
- OMENA, N. P. de. A referência à primeira pessoa do plural: variação ou mudança? *In*: PAIVA, M. C.; DUARTE, M. E. L. (Orgs.). **Mudança linguística em tempo real**. Rio de Janeiro: Contra capa Livraria Ltda, 2003, p. 83-60.
- OUSHIRO, Lívia. **Introdução à Estatística para Linguistas**, v.1.0.1. dez/2017. Disponível em: <https://zenodo.org/record/4755739#YmQADPPMJQI>. Acesso em: 23 abr. 2022.
- R CORE TEAM. **R: A Language and Environment for Statistical Computing**. The R Development Core Team: 2020.
- RAMOS, C. de M. de A.; BEZERRA, J. R. M.; ROCHA, M. de F. S. Do nosso cotidiano ou do cotidiano da gente? Um estudo da alternância “nós” e “a gente” no português do Maranhão. **Signum: Estudos da linguagem**. Londrina, v. 12, n. 1, p. 279-92, 2009.
- RUBIO, C. F. **Padrões de concordância verbal e alternância pronominal no português brasileiro e português europeu**: um estudo comparativo. São Paulo: Cultura Acadêmica, 2012.
- SANTOS, H. L. G. dos; VIANA, R. B. de M.; ARAÚJO, A. A. Panorama dos estudos sociolinguísticos no Ceará. *In*: ARAÚJO, A. A.; VIANA, R. B. de M.; RODRIGUES, L. da S. (Orgs.). **O falar culto de Fortaleza em foco**. São Paulo: Pimenta Cultural, 2021. p. 20-54.
- SEARA, I. C. A variação do sujeito nós e a gente na fala florianopolitana. **Organon**, v. 14, n. 28-29, p. 179-194, 2000.
- SCHERRE, M. M. P.; YACOVENCO, L. C.; NARO, A. J. Nós e a gente no português brasileiro: concordâncias e discordâncias. **Estudos de Linguística Galega**, v. especial I, p. 13-27, 2018.
- SCHERRE, M. M. P.; NARO, A. J.; YACOVENCO, L. C. Nós e a gente em quatro amostras do português brasileiro: revisitando a escala da saliência fônica. **Diadorim**, Rio de Janeiro, vol. 10 - Especial, p. 428-457, 2018.
- SILVA, F. J. de A. **A variação pronominal nós e a gente na fala de Fortaleza**. 2020. 91f. Dissertação (Mestrado em Linguística) – Universidade Federal do Ceará, Fortaleza.
- VIANNA, J. S.; LOPES, C. R. dos S. Variação dos pronomes nós e a gente. *In*: MARTINS, M. A.; ABRAÇADO, J. (Orgs.). **Mapeamento sociolinguístico do português brasileiro**. São Paulo: Contexto, 2015. p. 109-131.
- VITÓRIO, E. G. de S. L. A realização dos pronomes nós e a gente na função de sujeito e nas funções de complemento e adjunto na cidade de Maceió/AL. **Letrônica**, v. 10, n. 1, p. 122-138, 27 dez. 2017.
- ZILLES, A. M. S. O que a fala e a escrita nos dizem sobre a avaliação social do uso de a gente? **Letras de Hoje**, Porto Alegre, v. 42, n. 2, p. 27-44, jun. 2007.

VARIAÇÃO TU E VOCÊ NA POSIÇÃO DE SUJEITO NA FALA ALAGOANA

ELYNE GISELLE DE SANTANA LIMA AGUIAR VITÓRIO

1 INTRODUÇÃO

A implementação de *você* no quadro pronominal do português brasileiro ocasiona uma série de reestruturação gramatical em diferentes níveis linguísticos. Advindo, via processo de gramaticalização, da expressão nominal *vossa mercê*, mantém a especificação original de 3ª pessoa, apesar da alteração em termos semântico-discursivos. Passa a fazer referência à 2ª pessoa [- eu], preserva o traço formal de número e perde a especificação semântica de gênero [+ fem] (LOPES, 2007). Integrado ao sistema linguístico, varia com *tu* na representação da 2ª pessoa do singular em diferentes regiões brasileiras (SCHERRE *et al.*, 2015).

Scherre, Andrade e Catão (2020) apresentam que, nas variedades brasileiras, há as seguintes construções pronominais para representar a 2ª pessoa do singular: *você*, *cê*, *ocê*, *tu* sem concordância e *tu* com concordância. Assim, os autores propõem um redesenho do mapa dos pronomes *tu/você/cê/ocê* para três construções pronominais: macro *VOCÊ*, *tu* sem concordância e *tu* com concordância. Segundo os pesquisadores, a depender do condicionamento geográfico, essas construções são vistas ora como traços interacionais, ora como traços identitários, o que nos leva a perguntar: o que as pesquisas têm mostrado sobre essas construções na fala alagoana?

Para dar conta dessa questão, focalizamos a variação dos pronomes *tu* e *você* na posição de sujeito na fala alagoana. Para tanto, recorreremos às pesquisas sociolinguísticas que têm demonstrado que, nesse caso, *você* é variante mais usada, com o pronome *tu* sem concordância apresentando poucas realizações, principalmente em entrevistas sociolinguísticas, sendo associado a situações mais solidárias, igualitárias, familiares e informais. Também objetivamos, via problema de avaliação (LABOV, 2008[1972]), analisar se a variante *tu* com concordância assume um significado social de formalidade e maior distanciamento.

Com o objetivo de apresentar o *status* da variação *tu* e *você* na posição de sujeito em Alagoas, estruturamos o presente texto da seguinte forma: além desta seção introdutória, abordamos, a seguir, um panorama do uso dessas variantes na variedade

alagoana, tomando por base pesquisas sociolinguísticas; na seção seguinte, apresentamos o comportamento variável da variante *tu* sem concordância; em seguida, analisamos se a variante *tu* com concordância apresenta um valor social contrário à variante *tu* sem concordância; e, por fim, apresentamos algumas considerações, ressaltando os pontos mais relevantes das discussões.

2 SOBRE *TU* E *VOCÊ*

A publicação do texto *Variação dos pronomes “tu” e “você”*, de Scherre *et al.* (2015), que tem como um dos objetivos mapear a variação *tu*, *você*, *cê* e *ocê* nas diferentes regiões brasileiras, instigou-nos o fato de o estado de Alagoas fazer parte das áreas não mapeadas, por não haver informação sociolinguística disponível ou suficiente sobre o uso desses pronomes. Tentando suprir essa lacuna, revisitamos alguns textos que abordavam o uso de *tu* e *você* em Alagoas e propomos, no âmbito da Universidade Federal de Alagoas, projetos que contemplavam a variação *tu* e *você* na função de sujeito em diferentes comunidades alagoanas.

Tenório (2002), com o objetivo de refletir sobre o uso de *tu* e *você* em 80 diálogos de maceioenses (D2), sob uma perspectiva funcionalista, através de uma análise qualitativa dos dados, argumenta que há um uso efetivo do pronome *tu*. *Você* e *tu* sem concordância são as variantes mais registradas na amostra, havendo apenas duas ocorrências de *tu* com concordância: “*tu* achas” e “*tás* aonde?” (TENÓRIO, 2002, p. 49-50). Segundo a autora, o uso do pronome *tu* ocorre em diálogos com o objetivo de estabelecer intimidade entre os falantes ou de atribuir à frase um caráter irônico, predominando em frases interrogativas.

Cardoso (2008), ao analisar os pronomes pessoais no português brasileiro seguindo a visão da geolinguística pluridimensional, computa 64 realizações de *tu* e *você* na cidade de Maceió. A distribuição dos dados, nos questionários do Projeto Atlas Linguísticos do Brasil (Projeto ALiB), revela um predomínio no uso de *você*, com 94% das ocorrências (n = 60), contra apenas um percentual de 6% (n = 4) do pronome *tu*. Em relação à análise das distribuições diagenérica, diageracional e diastrática, a autora mostra que *você* é mais frequente entre homens, falantes da faixa etária 1 e falantes do ensino fundamental.

Silva e Vitória (2017), ao analisarem a realização de *tu*, *você* e *cê* na posição de sujeito no sertão alagoano, através da amostra do Projeto Lusa (VITÓRIO, 2017), apresentam percentuais de 94% (n = 473) de *você* contra apenas 6% (n = 33) de *cê*, com *você* sendo a variante preferida na comunidade. As poucas realizações de *cê* foram favorecidas entre os falantes do ensino médio e em contextos de paralelismo formal. O pronome *tu* sem concordância apresentou apenas três ocorrências na amostra, conforme (1), (2) e (3), que, segundo as pesquisadoras, pode estar relacionado ao tipo de coleta, que não favoreceu o seu uso.

- (1) Depois *tu vai* editá é? L9
- (2) Armária *tu é* doido é? L19
- (3) Eu lembro que naquela ladeira ali embaixo que *tu conhece* né? L96
(SILVA; VITÓRIO, 2017, p. 130).

Vitório (2018), ao analisar a variação *tu* e *você* em Maceió, argumenta que o pronome *tu* não é a primeira forma da comunidade para representar a 2PS, ocorrendo em situações que apresentam [+ intimidade] entre os interlocutores, com *você* sendo o pronome coringa no tratamento ao interlocutor. Ao analisar uma amostra constituída por 72 entrevistas sociolinguísticas e estratificada segundo as variáveis sexo/gênero, faixa etária e escolaridade, a autora computa 404 realizações – 98% (n = 398) de *você* contra apenas 2% (n = 6) de *tu* sem concordância, como (4) e (5), o que pode estar relacionado à limitação da amostra.

- (4) L48 – a violência tá demais né? /mais, mas/ isso aí eu acho que não é culpa dele *tu acha?* acho que num é culpa dele porque tem tanto policial eles mata até os puliciais
- (5) L69 – aí Hermes *tu quer* prestá queixa? claro que eu quero – lá pro Salvador Lira né? aí quando cheguei lá fiquei aguardando lá
(VITÓRIO, 2018, p. 91).

Ainda com o objetivo de explorar a variação *tu* e *você*, a pesquisadora analisa uma amostra constituída por um questionário composto por nove situações hipotéticas (três assimétricas descendentes, três assimétricas ascendentes e três simétricas), ambientadas na cidade de Maceió e aplicadas a 46 estudantes universitários maceioenses. Os dados gerais apontam 65% (n = 285) de *você*, 16% (n = 70) de *tu* sem concordância, 14% (n = 62) de outras formas (*senhor*, *senhora* e *pronome zero*) e 5% (n = 21) de *cê*, com a variante *você* sendo selecionada nas diferentes relações sociais e a variante *tu* sem concordância e preferida nas relações simétricas.

Silva (2019), ao analisar a variação *tu* e *você* na comunidade de Coité do Noia, município do agreste alagoano, a partir de uma amostra sociolinguística composta por 36 diálogos e estratificada segundo as variáveis sexo/gênero e faixa etária, apresenta percentuais de 89% (n = 463) de *você* e 11% (n = 57) de *tu* sem concordância. A autora argumenta que *tu* sem concordância pertence à comunidade, sendo associado a relações de familiaridade. É favorecido nos seguintes contextos: *tu* antecedido por *tu* na mesma sequência discursiva; nos diálogos entre homem/homem e entre falantes mais jovens.

A pesquisadora também analisa a variação entre *você* e sua variante *cê*, registrando percentuais de 80% de *você* (n = 372) contra apenas 20% (n = 91) da variante *cê*. Os dados apresentados mostram que o uso de *cê* é favorecido nos seguintes contextos: nas relações em que há mais intimidade entre os falantes, como a relação marido/mulher e amigo/amigo; na relação entre falantes do mesmo sexo, como homem/homem e mulher/mulher, e nas relações mais simétricas, ou, seja, mais igualitárias. A variante *cê* se apresenta como uma forma linguística associada a relações em que há mais intimidade e solidariedade entre os falantes.

Divino (2020), ao analisar a variação *tu* e *você* nos inquiridos do Projeto ALiB, mostra que, em Maceió, houve a realização de *você* – 94% (n = 196) e *tu* sem concordância – 6% (n = 12). Em Arapiraca, localizada no agreste alagoano e situada a cerca de 130 km a oeste de Maceió, a autora registra apenas o uso de *você* – 100% (n = 38). O uso categórico da variante *você* também é registrado em União de Palmares – 100% (n = 83), localizada a cerca de 73km a norte de Maceió. Em Santana do Ipanema, localizada no sertão alagoano a 207 km da capital, a autora registra 89% (n = 44) de *você* e 11% (n = 5) de *tu* sem concordância.

Silva (2020), ao analisar os pronomes de 2ª pessoa do singular em falares alagoanos, tomando por base dados do projeto Portal, registra o uso de *você*, *ocê*, *cê* e *tu* —não há especificação quanto à concordância verbal que acompanha a variante *tu* na posição de sujeito. A autora examina as realizações de sujeito e não sujeito conjuntamente, o que nos impossibilita de ver o comportamento dessas variantes na função de sujeito. No entanto, na análise dos dados, é possível observar exemplos não só do uso de *tu* com concordância (6), como também do uso de *ocê* (7), realização ainda não captada nas análises de fala alagoana.

(6) *Tu moras em Maceió mesmo é?* (SILVA, 2020, p. 37)

(7) [...] e o que é que *ocê* achou da eleição, foi boa? (SILVA, 2020, p. 44)

Em relação ao uso de *ocê*, Scherre, Andrade e Catão (2021) argumentam que

A quinta e última forma pronominal, *ocê*, é rara nos registros das falas da região Nordeste. Vamos encontrar relato de um caso na fala de Fortaleza [...], quatro casos na fala de Alto Parnaíba e três em Tuntum, áreas rurais do Maranhão [...] e registros em áreas rurais do semiárido baiano [...] (SCHERRE; ANDRADE, CATÃO, 2021, p. 12).

Vitório e Silva (2021), ao retornarem à amostra Lusa, ressaltam que, na fala dos documentadores do sertão alagoano, há o registro do pronome *tu* sem concordância, como em (8), (9) e (10), o que, segundo as pesquisadoras, pode ser um indicativo de que há o uso de *tu* sem concordância, mas ele não parece ser a primeira opção da comunidade; dessa forma, a entrevista sociolinguística não conseguiu captar a sua realização. Paredes Silva (2003) argumenta que esse tipo de coleta não é um contexto ideal para captar o uso da 2ª pessoa, havendo a necessidade de constituir amostras controladas que apresentem conversas naturais.

(8) e *tu* só *vende* sorvete ô faz mais alguma coisa?

(9) e:: – me fale um poco agora assim de uma viage ô do – algum passeio que *tu* já fez que acho interessante o lugar

(10) é – e:: *tu* já *passô* por alguma situação que pôs tua vida em risco?

(VITÓRIO; SILVA, 2021, p. 152-153)

Ainda com o intuito de explorar a variação *tu* e *você* na posição de sujeito no sertão alagoano, as autoras analisam uma amostra composta por nove situações hipotéticas ambientadas na cidade de Delmiro Gouveia, conforme metodologia adotada em Vitório (2018). O questionário foi aplicado a 44 estudantes universitários pertencentes

ao sertão alagoano e aponta 53% (n = 231) de *você*, 22% (n = 96) de *tu*, 16% (n = 69) de outras formas (*senhor*, *senhora* e *prônimo zero*) e 9% (n = 38) de *cê*, com a variante *tu* sendo associada a situações que apresentam [+ intimidade], [+ familiaridade] e [- formalidade]¹.

As pesquisas realizadas mostram que, apesar das limitações das amostras, *você*, *cê* e *tu* sem concordância são as variantes selecionadas para representar a 2ª pessoa do singular. Tenório (2002) e Silva (2020) também registram o uso de *tu* com concordância e Silva (2020) registra o uso de *ocê*. Os resultados mostram que são necessárias mais coletas de dados de amostras de conversas livres e naturais em diferentes comunidades alagoanas para que tenhamos mais segurança em relação ao uso do macro *VOCÊ*, principalmente quanto ao uso das variantes *cê* e *ocê*, bem como do *status* do pronome *tu* na variedade alagoana.

3 SOBRE *TU* SEM CONCORDÂNCIA

Vitório (2019), ao mensurar as percepções sociolinguísticas de estudantes universitários do agreste alagoano sobre o uso do pronome *tu*, pontua que *você* é a variante mais selecionada, mas *tu* faz parte da comunidade estudantil, não havendo uma avaliação negativa quanto ao seu uso. Segundo a autora, há uma associação entre o uso de *tu* e a origem geográfica do falante, bem como entre o uso de *tu* e as relações sociais, revelando que, na comunidade, há uma percepção de que estamos diante de um traço interacional. Nesse caso, o pronome *tu* está associado a situações comunicativas mais informais que apresentam mais familiaridade entre os falantes.

Em relação à concordância verbal com o pronome *tu*, Vitório (2021), ao mensurar a percepção de estudantes universitários do agreste alagoano, pontua que *tu* sem concordância é a variante preferida da comunidade, mostrando que a forma *tu* com concordância não apresenta uma correlação positiva nem com a comunidade, nem com a região dos estudantes. A pesquisadora argumenta que *tu* sem concordância não é avaliado negativamente, sendo mais associado a um uso prático, normal, familiar, cultural, íntimo e informal, fortemente relacionado à origem geográfica, à situação comunicativa e à pessoa com quem se interage.

Vitório (no prelo), ao explorar o *status* social do pronome *tu* sem concordância, retorna aos dados de Vitório (2018) e Vitório e Silva (2021), focalizando apenas os contextos de uso dos pronomes *tu* e *você*. Após a análise dos dados, a autora computa 684 realizações de *tu* e *você* na posição de sujeito, que apresentam a seguinte distribuição: 75% (n = 516) de *você* contra 15% (n = 168) de *tu*, revelando que, apesar de *você* ser a variante mais selecionada, há um percentual expressivo de uso de *tu*. A pesquisadora

¹ Nessa amostra, há 94 realizações de *tu* sem concordância e duas realizações de *tu* com concordância, associadas à fala de pessoas mais escolarizadas, como professor e médico, conforme Vitório (no prelo).

também mostra uma associação entre o uso de *tu* e *você* e as variáveis independentes relações sociais, tipos de relação e região.

Em relação ao uso de *tu* sem concordância, a autora observa que, apesar de ocorrer nas diferentes relações sociais, há um percentual maior de realização na relação simétrica — 50% (n = 124) contra 12% (n = 21) na relação assimétrica ascendente e 9% (n = 23) na relação assimétrica descendente. *Tu* sem concordância é uma variante linguística associada a situações socialmente vistas como solidárias, igualitárias e familiares, como é o caso das interações entre amigo-amigo, irmão-irmã, namorada-namorado, filho-mãe e pai-filho. Nesse contexto, exprime simetria e proximidade entre os falantes.

Se a variante *tu* sem concordância apresenta um *status* de intimidade e informalidade, sendo mais favorecida em contextos que apresentam [+ solidariedade] entre os falantes, o mesmo valor social parece não ser associado à variante *tu* com concordância. Vitório (2021) pontua que, enquanto *tu* sem concordância não se relaciona com a variável escolaridade, *tu* com concordância sinaliza uma possível relação, estando associada, de acordo como Vitório (no prelo), à fala de pessoas mais escolarizadas. Nesse contexto, uma pergunta emerge: *tu* com concordância assume um *status* de formalidade e maior distanciamento social na variedade alagoana?

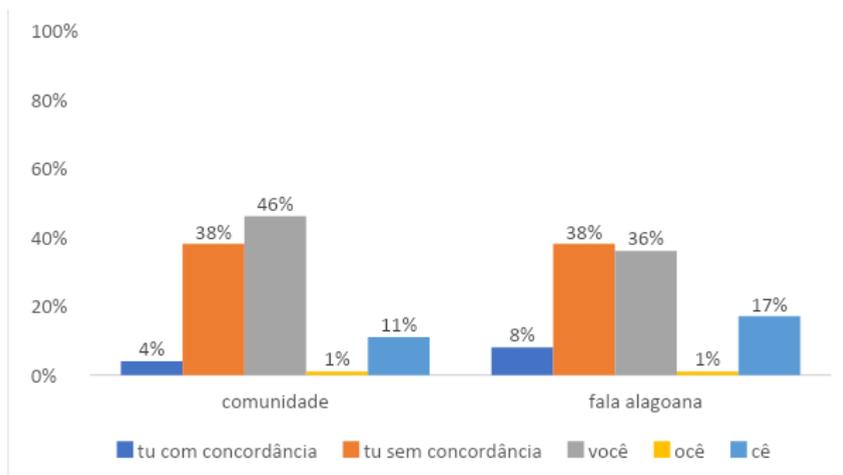
4 SOBRE *TU* COM CONCORDÂNCIA

Com o intuito de analisarmos o valor social associado a variante *tu* com concordância, recorreremos ao problema empírico da avaliação (WEINREICH; LABOV; HERZOG, 2006 [1968]) e adotamos uma abordagem direta (FASOLD, 1996), através de um questionário de avaliação linguística, respondido por 44 estudantes universitários do agreste alagoano². Para tanto, levamos em consideração os seguintes parâmetros sociais: percepção em relação ao uso linguístico; percepção quanto às variáveis sociais idade, sexo/gênero e escolaridade; percepção quanto ao contexto situacional; e percepção quanto à avaliação social.

Em relação à percepção quanto ao uso linguístico, perguntamos inicialmente aos estudantes quais variantes linguísticas — *tu* com concordância, *tu* sem concordância, *você*, *cê* e *ocê* — são usadas na sua comunidade e na fala alagoana. De acordo com o gráfico 1, observamos que, tanto na comunidade quanto na fala alagoana, há a percepção de uso das cinco variantes linguísticas, no entanto verificamos o predomínio das variantes *tu* sem concordância e *você*. A variante *cê* apresenta 11% na comunidade e 17% na fala alagoana, e a variante *tu* com concordância apresenta 4% na comunidade e 8% na fala alagoana. A variante *ocê* apresenta apenas 1%.

² Esses estudantes são alunos da UFAL – Campus Arapiraca e têm entre 20 e 35 anos. Participaram da pesquisa alunos de diferentes cursos de graduação, que, após a assinatura do TCLE, responderam ao questionário. A coleta foi realizada no próprio *campus*, nos meses de janeiro e fevereiro de 2020. Tendo em vista as discussões linguísticas, não participaram da pesquisa estudantes do curso de Letras.

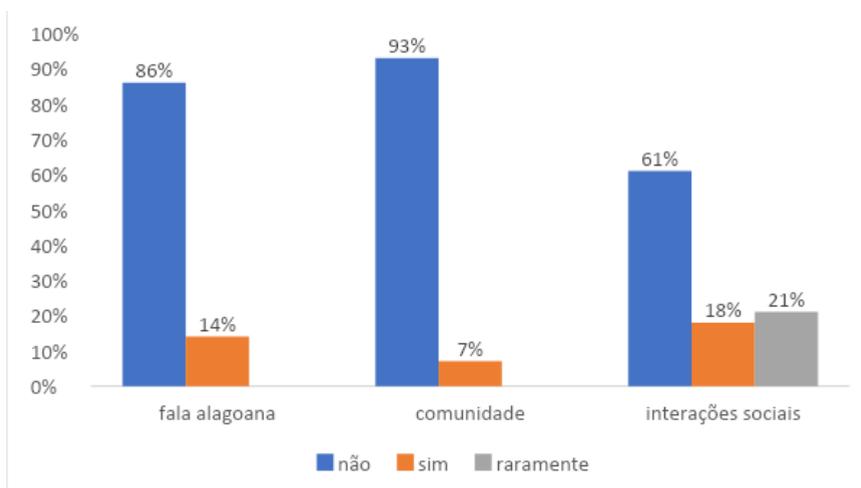
Gráfico 1: Percepção de uso dos pronomes de 2ª pessoa do singular



Fonte: elaboração própria.

Em seguida, perguntamos aos estudantes se o uso de *tu* com concordância é comum na fala alagoana, na sua comunidade e nas suas interações sociais. De acordo com gráfico 2, verificamos que predomina a percepção de que essa variante não é comum. No entanto, também notamos que 14% acreditam que *tu* com concordância é comum na fala alagoana, 7% acham que é comum na sua comunidade e 18% acreditam usá-la nas suas interações sociais. Essa percepção toma por base o argumento de que essa variante está presente em situações formais de comunicação, como observamos em (11), (12), (13) e (14). Nas interações sociais, também obtivemos um percentual de 21% para uso raro, como vemos em (15) e (16).

Gráfico 2: Percepção de uso da variante *tu* com concordância

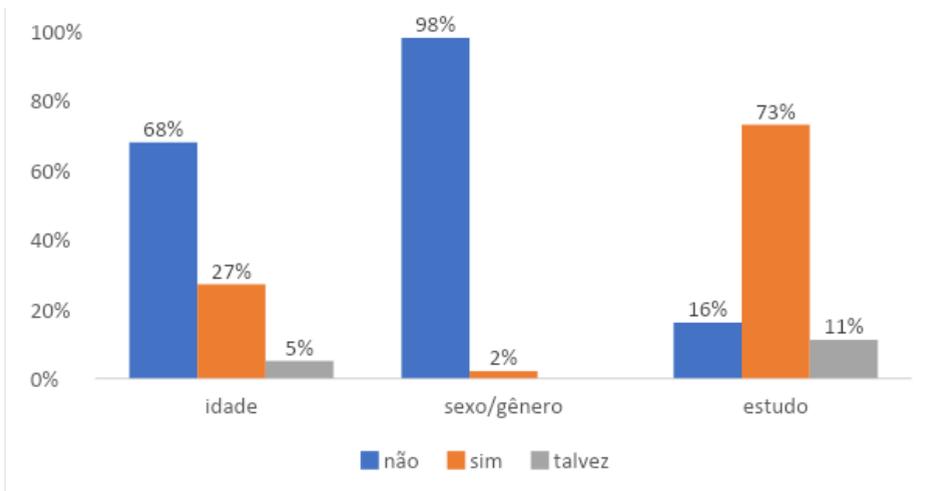


Fonte: elaboração própria.

- (11) sim, em situações mais formais – informante 3
- (12) em situações mais formais, como universidade, âmbito religioso – informante 18
- (13) em situações mais formais, como uma entrevista de emprego – informante 19
- (14) em alguns momentos sim, como, por exemplo, no ambiente de trabalho com a equipe, como *tu queres um café?* – informante 1
- (15) raramente uso, só em conversas sérias – informante 22
- (16) raramente, só em algumas situações mais formais – informante 31

Ao analisarmos a relação entre o uso de *tu* com concordância e as variáveis idade, sexo/gênero e estudo, verificamos, conforme o gráfico 3, que predomina a percepção de que idade e sexo/gênero não interferem no uso dessa variante, como observamos em (17). Quanto à variável idade, há um percentual de 27% que acreditam haver relação com tal uso, principalmente se o falante passa pelo processo de escolarização, como vemos em (18), (19) e (20), revelando que o que parece favorecer o uso de *tu* com concordância é o estudo dos falantes, que apresenta 73% para *sim*, como verificamos em (21), (22) e (23).

Gráfico 3: Percepção de uso segundo as variáveis idade, sexo/gênero e estudo



Fonte: elaboração própria.

- (17) não, acredito que esteja relacionado aos níveis de escolaridade e as situações que exigem a fala mais monitorada – informante 2
- (18) sim, porque é aprendido na escola no decorrer dos anos – informante 22
- (19) sim, creio que com o passar da idade se estuda e se aplica o que se aprende na escola quando necessário – informante 34
- (20) sim, porque as pessoas de maior idade tendem a seguir um modelo dicionarizado, dada a formação – informante 38
- (21) sim, pois através de um estudo mais aprofundado da gramática normativa o estudante conhecerá mais efetivamente tal uso – informante 10

(22) sim, geralmente, falantes menos escolarizados, grosso modo, não costumam realizar a concordância verbal – informante 18

(23) o uso desse tipo de linguagem é formal, está dentro de normas gramaticais que só são adquiridas através de um determinado nível de estudo – informante 23

Na análise dos dados, também notamos a percepção de que *tu* sem concordância seria adquirido pela criança durante a aquisição da linguagem, mas *tu* com concordância seria aprendido durante o processo de escolarização, como observamos em (24). A variante *tu* com concordância seria uma forma aprendida na escola, como vemos em (25) e (26). Ainda em relação à variável estudo, observamos um percentual de 16% para não e 11% para talvez, com o argumento de que o estudo não se relaciona com o uso dessa variante, pois está relacionado à situação comunicativa, como verificamos em (27).

(24) sim, pois é comum a criança aprender a usar *tu* sem concordância e, com anos de estudos, ela mude sua forma de usar o pronome *tu* – informante 4

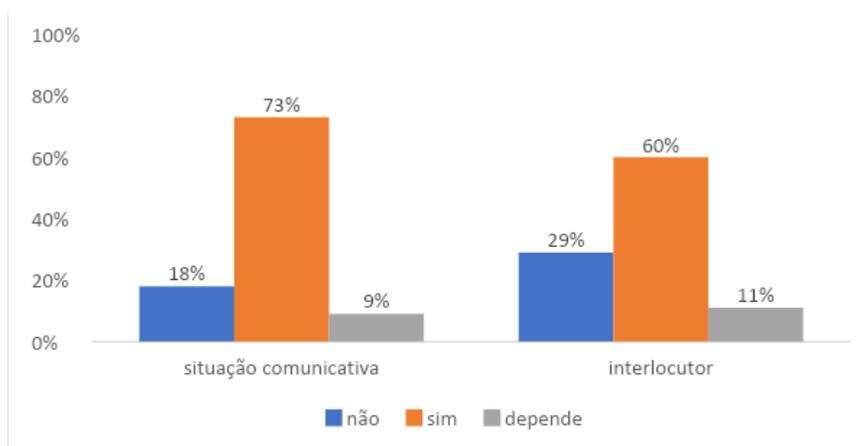
(25) com o conhecimento sobre a norma, o sujeito pode escolher usá-la – informante 44

(26) sim, pois muitos nem conhecem essa forma de uso do pronome *tu*, só começa a se familiarizar quando aprende na escola – informante 4

(27) esse uso está ligado à comunicação que o indivíduo participa – informante 26

Em relação à percepção quanto ao contexto situacional, notamos, de acordo com o gráfico 4, que tanto a situação comunicativa quanto o interlocutor são percebidos como fatores relacionados ao uso de *tu* com concordância, apresentando, respectivamente, percentuais de 73% e 60%. Essa percepção está ligada ao fato de *tu* com concordância ser associado a situações formais de uso da língua, como observamos em (28) e (29), sendo visto como uma forma de prestígio social, como verificamos em (30). Os estudantes que acreditam que essa relação não existe argumentam que tal uso é mais regional, como observamos em (31).

Gráfico 4: Percepção de uso segundo o contexto situacional



Fonte: elaboração própria.

- (28) sim, em situações mais formais, tem-se um cuidado com a fala – informante 21
 (29) sim, porque quando se fala com amigos usamos o *tu* sem concordância, mas quando se fala com estranho ou formalmente usamos com concordância – informante 1
 (30) sim, pois é de uso formal, visto o prestígio que ele tem – informante 17
 (31) não, é usado em algumas regiões e comunidades – informante 42

Com o intuito de aferirmos se há uma avaliação social positiva, negativa ou neutra em relação ao uso de *tu* com concordância, perguntamos aos estudantes quais características poderiam ser associadas a essa forma linguística. Conforme a figura 1, verificamos que predomina a percepção de estar diante de uma variante que apresenta prestígio social, sendo mais relacionada a características como formal, bonita, conservadora, prestigiosa, correta e padrão. Essa avaliação corrobora a percepção de que, quanto mais escolarizado o falante e mais formal a situação comunicativa, maior é uso de *tu* com concordância.

Figura 1: Avaliação social da variante *tu* com concordância



Fonte: elaboração própria.

Os dados indicam que *tu* com concordância, apesar de ser vista como uma variante linguística não comum na fala alagoana, assume um *status* de formalidade e distanciamento social, sendo associada à fala de pessoas mais escolarizadas e a situações comunicativas mais formais, diferentemente da variante *tu* sem concordância, que, segundo Vitório (2021), é associada a situações informais e não tem relação com a variável estudo. Nesse contexto, enquanto *tu* sem concordância exprime simetria e proximidade, *tu* com concordância exprime assimetria e distanciamento entre os falantes e é avaliada positivamente na comunidade.

5 ALGUMAS CONSIDERAÇÕES

Neste trabalho, descrevemos o estado da variação *tu* e *você* na função de sujeito na fala alagoana. Para tanto, consideramos as pesquisas de Tenório (2002), Cardoso (2008), Silva e Vitória (2017), Vitória (2018, 2019, 2021), Silva (2019), Silva (2020), Divino (2020) e Vitória e Silva (2021), que mostram, nos termos de Scherre, Andrade e Catão (2020), a realização de *você*, *ocê*, *cê*, *tu* com concordância e *tu* sem concordância, com a predominância de *você* e *tu* sem concordância. Nas comunidades alagoanas, *você* é a variante selecionada, com *tu* sem concordância apresentando poucas realizações, principalmente em entrevistas sociolinguísticas.

Também observamos que, enquanto *você* é a variante coringa para representar a 2ª pessoa do singular, ocorrendo em qualquer situação comunicativa — o que parece revelar um significado social neutro —, *tu* sem concordância e *tu* com concordância apresentam valores sociais diferentes. Em relação à primeira variante, notamos que há um *status* de intimidade e informalidade, favorecida em contextos que apresentam [+ solidariedade] entre os falantes, ao passo que a segunda variante é associada ao processo de escolarização, sendo vista como de maior formalidade e maior distanciamento social.

Essas considerações são questões que vêm sendo formuladas e analisadas em relação ao uso dessas variantes na fala alagoana. Acreditamos que a abordagem de como elas são produzidas e percebidas pelos falantes contribui para mostrar o nível de consciência social de um dado fenômeno linguístico variável, revelando que a análise de dados de produção, bem como de dados de percepção possibilita uma maior compreensão do processo de variação. Também acreditamos que são necessárias mais coletas de dados de amostras de conversas livres e naturais para que tenhamos mais informações acerca do uso dessas formas linguísticas.

REFERÊNCIAS

- CARDOSO, S. **Caminhos dos pronomes pessoais no português brasileiro: considerações a partir de dados do projeto ALiB.** 2008. Disponível em: http://dlcv.fflch.usp.br/sites/dlcv.fflch.usp.br/files/09_3.pdf. Acesso em: 10 out. 2017.
- COUTO, L.; LOPES, C. (Orgs.). **As formas de tratamento em português e em espanhol: variação, mudança e funções conversacionais.** Niterói: Editora da UFF, 2011.
- DIVINO, L. **Tu e você em cinco estados do Nordeste a partir dos dados do Projeto Atlas Linguístico do Brasil: um estudo variacionista.** 2020. 254f. Tese (Doutorado em Língua e Cultura) – Universidade Federal da Bahia, Salvador.
- FASOLD, R. **La sociolingüística de la sociedad: introducción a la sociolingüística.** Madrid: Visor, 1996.
- LABOV, W. **Padrões sociolinguísticos.** São Paulo: Parábola Editorial, 2008 [1972].
- LOPES, C. **Pronomes pessoais.** In: VIEIRA, S.; BRANDÃO, S. (Orgs.). **Ensino de gramática: descrição e uso.** São Paulo: Contexto, 2007. p. 103-119.
- LOPES, C. *et al.* **A reorganização do sistema pronominal de 2ª pessoa na história do português brasileiro: a posição de sujeito.** In: LOPES, C. (Org.). **Mudança sintática das classes de palavra: perspectiva funcionalista.** São Paulo: Contexto, 2018.

- PAREDES SILVA, V. L. O retorno do pronome tu à fala carioca. *In: RONCARATI, C.; ABRAÇADO, J. (Orgs.). Português brasileiro: contato linguístico, heterogeneidade e história.* Rio de Janeiro: 7Letras, 2003. p. 160-169.
- SCHERRE, M. *et al.* Variação dos pronomes “tu” e “você”. *In: MARTINS, M. A.; ABRAÇADO, J. (Orgs.). Mapeamento sociolinguístico do português brasileiro.* São Paulo: Contexto, 2015. p. 133-172.
- SCHERRE, M.; ANDRADE, C.; CATÃO, R. Redesenhando o mapa dos pronomes tu/você/cê/ocê no português brasileiro falado. *In: WITCHES, P. et al. (Orgs.). Conquistas e desafios dos estudos linguísticos na contemporaneidade.* Porto Alegre: Editora Fi, 2020.
- SILVA, S. **A variação pronominal de segunda pessoa do singular em Coité do Nóia /AL.** 2019. 122f. Dissertação (Mestrado em Linguística e Literatura) – Universidade Federal de Alagoas, Maceió.
- SILVA, W. **Pronomes de 2ª pessoa do singular em falares alagoanos: uma análise variacionista.** 2020. 51f. Trabalho de Conclusão de Curso (Graduação em Letras) – Universidade Federal de Alagoas, Maceió.
- SILVA, S.; VITÓRIO, E. A variação você e cê no sertão alagoano. **Revista Leitura**, v. 2, n. 59, p. 122-142, 2017.
- TENÓRIO, A. **O uso das formas tu e você em diálogos de maceioenses.** 2002. 68f. Dissertação (Mestrado em Letras e Linguística) – Universidade Federal de Alagoas, Maceió.
- VITÓRIO, E. **A língua usada no sertão alagoano: constituição da amostra.** Trabalho apresentado no III Estudos em Linguagem do Sertão. Delmiro Gouveia, jun. 2017.
- VITÓRIO, E. A variação *tu* e *você* em Maceió, Alagoas. **Todas as Letras**, v. 20, n. 2, p. 85-99, 2018.
- VITÓRIO, E. Percepções sociolinguísticas de estudantes universitários em relação ao uso do pronome *tu*. **Revista Investigações**, v.32, n. 2, p. 432-455, 2019.
- VITÓRIO, E. Avaliação social e a concordância verbal com o pronome *tu*. **Diadorim**, Rio de Janeiro, vol. 23, n. 1, p. 151-170, jan./jun. 2021.
- VITÓRIO, E. **Panorama da variação tu e você na função de sujeito em Alagoas.** (no prelo)
- VITÓRIO, E.; SILVA, S. A variação tu/você/cê na posição de sujeito e o problema da avaliação linguística no sertão alagoano. *In: NUNES, C.; SILVA, C. (Orgs.) A língua em foco no nordeste brasileiro: d’além das capitais.* São Paulo: Pontes Editora, 2021. p. 149-169.
- WEINREICH, U.; LABOV, W.; HERZOG, M. **Fundamentos empíricos para uma teoria da mudança linguística.** Trad.: Marcos Bagno. São Paulo: Parábola Editorial, 2006 [1968].

A PERCEÇÃO SOCIOLINGUÍSTICA DO SUBJUNTIVO E DO INDICATIVO

WENDEL SILVA DOS SANTOS

1 INTRODUÇÃO

O texto que se desenvolve aqui, embora seja fruto da apresentação de uma mesa-redonda que tratou de pesquisas realizadas no nordeste, traz dados de percepção de ouvintes ludovicenses e paulistanos, uma vez que é parte da pesquisa de doutorado desenvolvida por Santos (2020), que analisou, a partir da teoria da variação e mudança (WEIREICH, LABOV; HERZOG, 2006[1968]; LABOV, 2008[1972]; ECKERT, 2008; 2012), percepções sociolinguísticas associadas ao subjuntivo e ao indicativo, em São Luís, capital do estado do Maranhão, e São Paulo, capital do estado de mesmo nome.

São Paulo e São Luís distam cerca de 2.900 km uma da outra e não apenas diferem por sua geografia, mas também por avaliações sociais que, de alguma maneira, subsidiam o interesse deste estudo. De um lado, é comum ouvir que paulistanos não sabem ou não usam o subjuntivo. De outro, também se ouve que ludovicenses falam a melhor variedade do português brasileiro.

Figura 1: Distância entre São Luís e São Paulo.



Fonte: adaptado do Google Maps.

No trecho da entrevista sociolinguística apresentada a seguir, Leandro (SLM2S)³, que é ludovicense, por exemplo, ao visitar a irmã em Brasília, explica que o indicativo é característico da fala de paulistanos, ou seja, pode-se inferir que, de acordo com o informante, paulistano ou não sabe ou não usa subjuntivo.

(1)

Leandro: eu estava no carro em Brasília e um uma das pessoas que estava no carro era paulista então nós tínhamos estávamos fazendo alguma coisa eh pra ajeitar o carro ah o paulista saia ia sair do carro e aí [o...]

Doc.: [era carro de quem?]

Leandro: era um carro da minha irmã e aí o rapaz (es)tava com dificuldade em colocar o banco pra frente pra poder passar e aí ele perguntou assim: “você *QUER* que eu *ajudo*?” aí eu ach/ notei logo que ele não era de Brasília.

Doc.: uhn run.

Leandro: justamente por causa disso aí eu perguntei aí depois que ele saiu eu perguntei pros pro meu pra minha irmã “esse esse rapaz não é daqui né ele é de de São Paulo” minha irmã perguntou “como é que tu sabe?” eu falei “olha ele acabou de falar de de um modo que só em São Paulo é mais característico.

No próximo trecho (2), o informante Neto (SLM3B) afirma que ludovicenses falam o melhor português do Brasil — o que, em princípio, pressupõe que eles, mais do que paulistanos, empregam subjuntivo pelo menos nos contextos canônicos, como no caso de uma oração subordinativa objetiva direta que complementa o verbo *querer*, como no exemplo anterior, dado por Leandro.

(2)

Doc: quando o senhor chegava nessas cidades seu seu seu N. as pessoas percebiam logo de cara que o senhor não era... que o senhor era de São Luís?

Neto: percebe porque o cara de São Luis ele é um cara... eu não sei eu não digo mais inteligente mas se (inint.)

Doc: uhn rum.

Neto: nego diz que **a gente fala o melhor português do Brasil**

Doc: uhn rum.

Neto: agora eu não sei **mas nós falamos melhor do que o português mesmo a gente fala**

Doc: do que o português de Portugal?

Neto: é é.

Doc: é?

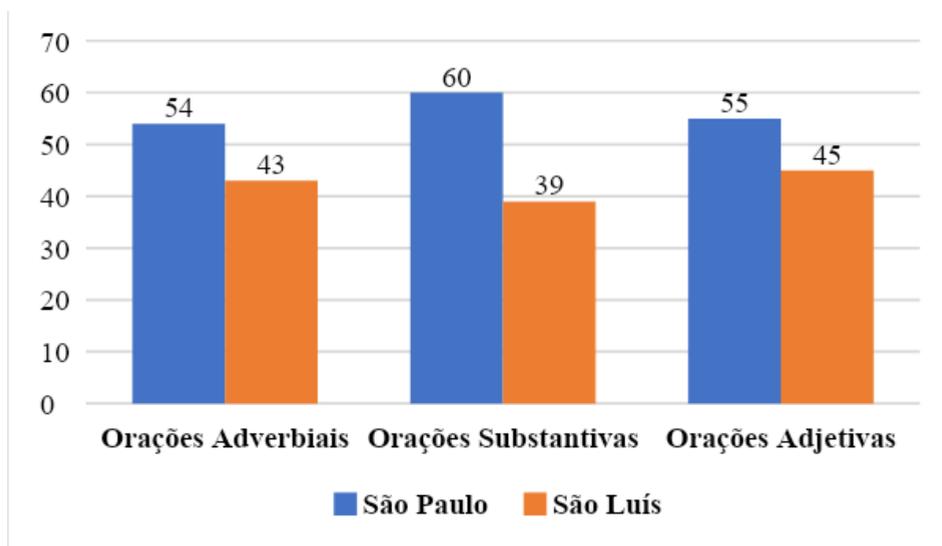
³ Santos (2015) e Santos (2020) atribuíram pseudônimos aos participantes de suas respectivas pesquisas. A abreviação **Doc** refere-se ao documentador. Nos exemplos, os códigos entre parênteses, logo após os trechos das entrevistas, indicam a cidade do informante (São Paulo ou São Luís), o seu sexo/gênero (Masculino ou Feminino), a faixa etária (1, para informantes que estejam entre 18 e 35 anos; 2, para informantes que estejam entre 36 e 59 anos; e, 3, para informantes que estejam com 60 anos ou mais) e seu nível de escolaridade (**B** para aqueles que têm até a educação básica, e **S** para os que têm ensino superior). Essa identificação dos participantes segue a proposta do Projeto SP2010, coordenado pelo Prof. Dr. Ronald Beline Mendes (USP).

Discursos metapragmáticos acerca da variedade ludovicense do português, bem como da variedade paulistana, podem ser acessados em Santos (2020), que apresenta fortes indícios de que, de um lado, ainda é bastante recorrente a máxima de que o Maranhão é o lugar onde se fala a melhor variedade do português brasileiro — ainda que tal mito já tenha sido debatido amplamente (Bagno, 2017) —, e, de outro lado, paulistanos não sabem ou não gostam do subjuntivo.

Considerando que o discurso popular associa a ausência do subjuntivo à fala de paulistanos e que o subjuntivo é uma forma fortemente associada a noções de correção gramatical, Santos (2015) inclui, em sua pesquisa, vários grupos de fatores nas análises quantitativas que estabeleceram os padrões de uso de formas indicativas e subjuntivas em São Paulo e São Luís.

Embora não seja possível discutir cada uma dessas variáveis independentes, dado ao espaço da escrita deste capítulo, reproduzem-se, a seguir, as taxas de usos gerais de indicativo nas duas cidades analisadas.

Figura 2: Tendências de uso, em pesos relativos, do uso do indicativo em São Paulo e São Luís



Fonte: Adaptado de Santos (2015).

Os resultados evidenciam que, apesar de as taxas de uso do indicativo serem maiores entre os paulistanos — isto é, esses falantes se correlacionam mais a essa morfologia verbal —, comparativamente aos falantes ludovicenses, Santos (2015, p. 111) explica que, na análise geral, essas diferenças são pequenas, e as maiores delas, observadas entre as orações subordinadas substantivas. Assim, “São Paulo e São Luís estão mais próximas entre si no caso das adverbiais e das relativas”. O autor resume que:

- (i) os padrões de alternância de formas indicativas e subjuntivas são diferentes nas duas cidades;
- (ii) essas diferenças são estatisticamente significativas e
- (iii) são maiores quando se trata

de orações substantivas e, finalmente (iv) ainda que o indicativo tenda a ser relativamente mais empregado em São Paulo, há indícios de mudança na direção do indicativo em São Luís (SANTOS, 2015, p. 111).

Essas distribuições de dados em São Paulo e em São Luís dizem respeito a uma amostra de fala com 36 falantes de cada cidade, que revela que, na verdade, não há muita diferença entre eles. Além disso, muito pontuais, essas diferenças não conseguem responder à pergunta acerca dos usos desses modos verbais por paulistanos, tampouco confirmam a “crença” de que os ludovicenses são mais conservadores no uso da língua.

A partir daí, surge o interesse de verificar se a máxima de que paulistanos não sabem o subjuntivo poderia ser confirmada por meio de um trabalho de percepção. Assim, a pergunta de pesquisa lançada é a seguinte: *Falantes — paulistanos ou ludovicenses — soam mais competentes (no sentido de escolarização e inteligência, por exemplo) e como pessoas de classe social mais alta quando são ouvidos no subjuntivo ou no indicativo?*

Especificamente, buscou-se responder às seguintes questões: i) *o indicativo, no lugar em que o subjuntivo seria uma forma possível, senão prescrita, é semelhantemente percebido como “erro” por paulistanos e ludovicenses?*; ii) *o indicativo no lugar do subjuntivo é percebido como uma característica de paulistanidade por esses grupos de ouvintes?*; iii) *os ouvintes percebem formas superficiais da língua, mas não características mais profundas* (LABOV, 1993; LABOV *et al.*, 2011)?

Para responder a essas questões, propôs-se a realização de um experimento de percepção sociolinguística, com base na técnica dos estímulos pareados (LAMBERT *et al.*, 1960; CAMPBELL-KIBLER, 2006, 2009; OUSHIRO, 2015; MENDES, 2016, 2018), cujo interesse era o de acessar os significados sociais associados a formas subjuntivas e indicativas em orações adverbiais introduzidas por *embora* e *talvez*, e em orações substantivas introduzidas pelos verbos *querer* e *acreditar*.

O desenho desse experimento é detalhado mais adiante, na metodologia da pesquisa. Antes, porém, na seção a seguir, apresenta-se o interesse teórico nos estudos de percepção pela sociolinguística variacionista. Depois, apresentam-se estudos que tenham se ocupado em acessar significados sociais de variáveis linguísticas, especialmente as gramaticais. O capítulo se encerra com a apresentação dos resultados alcançados a partir da aplicação do experimento elaborado especialmente para esta pesquisa e com a relação dos trabalhos citados ao longo deste texto.

2 O PAPEL DO OUVINTE NA VARIAÇÃO SOCIOLINGUÍSTICA

Para Campbell-Kibler (2006), o papel do ouvinte no projeto da estruturação linguística por meio da percepção evidencia que os falantes que diferem em suas preferências linguísticas, com base em seu senso social de significado, também devem diferir enquanto ouvem, no sentido de que atribuem valores distintos a variantes

linguísticas distintas⁴. Na verdade, o modo como os falantes ouvem — e, portanto, avaliam certas formas linguísticas — é só uma parte de como funciona a heterogeneidade inerente às línguas naturais, que, conjuntamente ao comportamento linguístico, devem dar a dimensão da uniformidade da variação linguística, algo já amplamente aceito nos estudos de produção. Weinreich, Labov e Herzog (2006[1968], p.102) explicam que “os correlatos subjetivos da alternância de linguagem revelam ser mais uniformes que o próprio comportamento”. Assim, a busca por padrões nas avaliações e percepções sociolinguísticas pode subsidiar explicações importantes para complexos processos de mudança linguística.

Em seu estudo sobre modularidade linguística, Levon e Buchstaller (2014) analisaram o efeito desse mecanismo nas avaliações sociais dos ouvintes. Os autores partiram da proposta de Labov (1993) e Labov *et al.* (2011), que sugerem, com base no chamado princípio da interface, que variáveis morfossintáticas não causam os mesmos efeitos perceptuais que variáveis fonológicas, o que poderia ser explicado por meio do fato de que “os membros da comunidade de fala avaliam a forma superficial da linguagem, mas não características estruturais mais abstratas” (LABOV, 1993, p. 4).

Levon e Buchstaller (2014) partem do modelo denominado monitorador sociolinguístico, proposto por Labov *et al.* (2011), para quem os falantes “monitoram” o signo linguístico a partir de formas socialmente significativas para se dizer a mesma coisa, passando pela avaliação do ouvinte com base no significado social atribuído a esses signos linguísticos. Em outras palavras, Mendes (2018: 32) esclarece que o monitorador sociolinguístico se refere a uma capacidade que o falante tem de “varrer o sinal linguístico”, em busca de diferentes modos de dizer a mesma coisa, avaliando essas formas de acordo com o que aprende em comunidade, ao longo de sua experiência sociolinguística.

Ao fim, Levon e Buchstaller observam que as variantes linguísticas em níveis mais altos (BERLINCK *et al.*, 2020) estão disponíveis para o monitorador sociolinguístico e, além dos fatores comumente analisados nas pesquisas sociolinguísticas, como categorias macrosociais, outras variáveis de ordem cognitiva devem ser inseridas em estudos dessa natureza, já que as reações dos ouvintes em relação às variantes linguísticas se correlacionam a efeitos que independem do lugar de origem, do estilo ou da faixa etária, por exemplo.

O trabalho de Levon e Buschstaller (2014) é apenas um dos que têm se dedicado, nos últimos anos, a acessar significados sociais de variantes linguísticas por ouvintes de diversas localidades. No Brasil, tal empreendimento é recente, mas também tem recebido destaque. Citem-se, por exemplo, a pesquisa de Oushiro (2015), a respeito da percepção da realização de /r/ em coda silábica, como tepe e retroflexa, na cidade de São Paulo, e a de Mendes (2018), que analisou os efeitos da pronúncia de (EN), em

⁴ Em sua tese de doutorado, Campbell-Kibler (2006: 11) explica: “When contemplating the role of listener perceptions in the structuring of linguistic variation, it is apparent that speakers who differ in their linguistic preferences and social sense of meaning while speaking must also differ while listening”.

casos como *fazenda*, que pode ser realizada como ditongo [ejn] ou monotongo [en], e da realização de concordância nominal, por meio de suas variantes não padrão CNØ ou padrão CNp, esta última podendo estar associada a noções de gênero, como soar gay. Além desses trabalhos, merecem destaque o de Melo (2020), que apresenta resultados para dados de produção e percepção da realização de /s/ em coda silábica, em palavras como me[z]mo ~ me[3] mo ~ me[v]mo~ me[h] mo ~ me[Ø]mo⁵, e o de Santos (2020), que analisa a percepção de formas indicativas e subjuntivas em São Luís, capital do Maranhão, e na cidade de São Paulo.

Este último autor verificou que, na esteira do que propõem Niedzielski (1999) e Campbell-Kibler (2006), bem como Levon e Buchstaller (2014), ouvintes ludovicenses e paulistanos não foram capazes de determinar o lugar de origem dos quatro falantes que cederam suas vozes para serem submetidas a avaliações sociolinguísticas pelos participantes. Isso pode ser um indício de que categorizações sociais dos falantes (CAMPBELL-KIBLER, 2006), como a cidade de origem, no caso do trabalho de Santos (2020), operam em níveis muito baixos para afetar consideravelmente a percepção dos ouvintes quando são colocados de frente com uma forma linguística específica.

Para acessar significados sociais associados a formas linguísticas, a maior parte dos estudos sociolinguísticos tem se valido da técnica *matched-guise*, ou estímulos pareados, ou, ainda, pares falsos, desenvolvida por Lambert *et al.* (1960). Esses estímulos — ou pares falsos —, extraídos das falas de certo número de falantes, são manipulados no sentido de que sejam iguais em todos os aspectos, diferenciando-se apenas com respeito à variável em foco. Assim, ouvintes diferentes são estimulados com dados semelhantes da fala de um mesmo ouvinte, mas em disfarces diferentes — como formas com morfologia do indicativo e do subjuntivo, tal como se descreve mais adiante, neste texto. Campbell-Kibler (2006) explica que, ao perceberem que, como o preconceito linguístico é mais frequentemente sancionado socialmente do que outras formas, Lambert *et al.* (1960) decidiram que os participantes do seu estudo expressariam suas opiniões mais abertamente se as respostas fossem motivadas por desempenhos linguísticos. Assim, pediram que um mesmo participante gravasse a si próprio falando de maneiras diferentes. O objetivo era investigar as crenças e atitudes dos ouvintes em relação às categorias de pessoas que geralmente se acredita que falam da maneira como são ouvidas.

Ainda que Weinreich, Labov e Herzog (2006[1968], p. 120) tenham afirmado que os estudos de percepção sociolinguística oferecem “um rico volume de dados sobre estágios recentes da mudança linguística”, esse tipo de estudo acabou ficando em segundo plano nas pesquisas de mudança linguística, e só mais recentemente tal método tem sido replicado de maneira mais recorrente na sociolinguística. O trabalho que aqui se propõe é uma delas, e esse desenho é detalhado na seção 3, a seguir.

⁵ Exemplos extraídos de Melo (2020: 807).

3 PROCEDIMENTOS METODOLÓGICOS

O presente estudo tem interesse em investigar os efeitos da morfologia do subjuntivo e do indicativo na percepção de quão *competentes, sérios, formais e antipáticos* soavam ludovicenses e paulistanos. Para estes últimos, interessou ao autor também verificar se o modo verbal tinha efeito em quão *paulistanos* eles soavam, de acordo com ouvintes tanto de São Luís quanto de São Paulo. A discussão levantada pelo pesquisador foi a de que, comparativamente a variáveis fonéticas, variáveis gramaticais têm recebido menos atenção nos estudos de percepção e estariam menos disponíveis para avaliação social (LABOV, 1993; LABOV *et al.*, 2011). Para tanto, lançou mão dos pressupostos teóricos da sociolinguística variacionista (WEINREICH; LABOV; HERZOG, 2006[1968]; LABOV, 2008[1972]; ECKERT, 2008; 2012).

Para responder às questões levantadas, Santos (2020) propôs a realização de um experimento de percepção sociolinguística no qual os estímulos, elaborados de acordo com a técnica *matched-guise* (LAMBERT *et al.*, 1960; CAMPBELL-KIBLER, 2006; CANEVER, 2017; MENDES, 2018), são períodos com orações adverbiais introduzidas por *embora* e *talvez*, bem como orações substantivas que complementam os verbos *querer* e *acreditar*.

Embora e *talvez* foram incluídos por terem sido os subordinadores mais recorrentes na amostra que Santos (2015) examinou em seu conjunto de análises sobre as orações subordinadas adverbiais. Além disso, como esse autor mostra, trata-se de dois bons exemplares da forte noção normativa (ROCHA LIMA, 2002; CUNHA; CINTRA, 2007) que o subjuntivo carrega quando aparece em sentenças introduzidas por essas conjunções.

Quanto a exemplos com o verbo *querer*, eles são os que aparecem na fala das pessoas quando dizem que paulistanos não sabem o subjuntivo. Foi inserido também o verbo *acreditar* no experimento porque, a partir de um levantamento feito em estudos de produção, bem como em gramáticas normativas, se verificou que o seu padrão de uso não é homogêneo com o modo subjuntivo ou o indicativo. Decidiu-se, então, usá-lo também no experimento, no sentido de verificar se essas diferenças aparecem nas percepções das pessoas, a depender do verbo da oração principal.

Foram gravados um homem e uma mulher paulistanos e um homem e uma mulher ludovicenses, todos com ensino universitário completo e com faixa etária entre 20 e 30 anos, enunciando 4 pares de sentenças — uma no subjuntivo e uma no indicativo, sendo dois pares de adverbiais introduzidas por *embora* e *talvez*, e dois pares de subordinadas substantivas, introduzidas pelo *volitivo* *querer* ou pelo *cognitivo* *acreditar*.

Para diferenciar os dois falantes ludovicenses dos dois paulistanos, foi manipulada, além da forma verbal, a pronúncia do *-r/* em coda. Nas sentenças formuladas, o tepe foi usado para identificar os paulistanos, já que, como Oushiro (2015) mostra, é um estereótipo da sua fala, e o aspirado, para caracterizar os ludovicenses. Assim, a manipulação das gravações foi feita no Praat (BOERSMA; WEENINK 2019), no sentido

de criar pares de estímulos definidos pela oposição indicativo/subjuntivo, com /r/ tepe para paulistanos e /r/ aspirado para ludovicenses. No total, são 32 os estímulos, 16 no subjuntivo e 16 no indicativo.

Figura 3: Conjuntos de Estímulos

A1	Rodrigo SP		Fabiano SL		Juliana SP		Deborah SL	
Voltivo	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind
Cognitivo	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind
Embora	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind
Talvez	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind

A2	Rodrigo SP		Fabiano SL		Juliana SP		Deborah SL	
Voltivo	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind
Cognitivo	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind
Embora	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind
Talvez	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind

B1	Rodrigo SP		Fabiano SL		Juliana SP		Deborah SL	
Voltivo	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind
Cognitivo	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind
Embora	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind
Talvez	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind

B2	Rodrigo SP		Fabiano SL		Juliana SP		Deborah SL	
Voltivo	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind
Cognitivo	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind
Embora	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind
Talvez	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind

C1	Rodrigo SP		Fabiano SL		Juliana SP		Deborah SL	
Voltivo	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind
Cognitivo	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind
Embora	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind
Talvez	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind

C2	Rodrigo SP		Fabiano SL		Juliana SP		Deborah SL	
Voltivo	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind
Cognitivo	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind
Embora	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind
Talvez	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind

D1	Rodrigo SP		Fabiano SL		Juliana SP		Deborah SL	
Voltivo	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind
Cognitivo	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind
Embora	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind
Talvez	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind

D2	Rodrigo SP		Fabiano SL		Juliana SP		Deborah SL	
Voltivo	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind
Cognitivo	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind
Embora	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind
Talvez	Sub	Ind	Sub	Ind	Sub	Ind	Sub	Ind

Fonte: extraído de Santos (2020, p. 127).

Os estímulos foram combinados e apresentados aos ouvintes de acordo com a figura 3. Foram criados 8 conjuntos de dados, e cada participante ouviu cada um dos quatro falantes apenas uma vez. Por exemplo, quem ouviu o conjunto de estímulos A1, escutou os falantes de cada cidade apenas no disfarce em azul. Observe-se que o conjunto A2 forma um par com A1, na medida em que se tem as mesmas orações, mas com outra forma verbal. Rodrigo, por exemplo, que é de São Paulo, no conjunto A1, foi ouvido falando uma oração com verbo *querer* complementado com verbo no subjuntivo, mas, no conjunto A2, foi ouvido falando a mesma oração com verbo *querer*, com complemento com verbo no indicativo. Todos os conjuntos de estímulos foram organizados de acordo com essa lógica, no sentido de que todos os falantes seriam ouvidos em algum disfarce por algum ouvinte, porém cada falante seria ouvido apenas uma vez por cada participante.

Antes de aplicar o experimento propriamente, os estímulos elaborados foram apresentados a um conjunto de ludovicenses, além de alunos da Universidade de São Paulo, a fim de que dessem suas impressões gerais sobre os áudios. Elas foram, então, sintetizadas em nuvem de palavras. Em seguida, as palavras com maior destaque na nuvem foram utilizadas na elaboração do questionário de percepção.

Tabela 1: Número de ouvintes de São Luís e de São Paulo que participaram do experimento

	São Luís	São Paulo
Feminino	125	169
Masculino	92	115
Ensino Fundamental	13	16
Ensino Médio	67	98
Ensino Superior	137	170
Total	217	284

Fonte: adaptada de Santos (2020, p. 136).

O formulário do teste de percepção sociolinguística foi preenchido por 501 ouvintes, e, como cada um deles respondeu a 4 questionários (um para cada falante), tem-se um total de 2004 respostas, analisadas estatisticamente por meio do programa R (R CORE TEAM, 2022).

4 APRESENTAÇÃO DOS RESULTADOS

Antes da análise estatística propriamente dita, fez-se uma análise de componentes principais com os dados de São Paulo e São Luís juntos, para verificar se as respostas nas diferentes escalas eram independentes umas das outras, ou se havia algum padrão de interação entre elas.

Tabela 2: Correlação entre as escalas de diferenciais semânticos.
Método: análise de componentes principais com rotação Promax

	SÃO LUÍS				SÃO PAULO			
	COMPETÊNCIA	ANTIPATIA	PAULISTANIDADE	FORMALIDADE	ANTIPATIA	COMPETÊNCIA	PAULISTANIDADE	SERIEDADE
Escolaridade	1,07	0,06	-0,10	-0,11	0,06	0,89	0,06	-0,04
Inteligência	1,00	0,03	-0,06	-0,02	0,12	0,84	0,23	0,04
Formalidade	0,26	0,05	-0,06	0,74	-0,13	0,53	-0,55	0,17
Seriedade	-0,15	0,03	-0,07	1,06	-0,17	0,01	-0,26	0,84
Educação	0,41	-0,22	0,42	0,22	0,78	0,04	0,18	0,45
Amigabilidade	-0,01	-0,94	-0,02	0,05	0,94	0,08	-0,19	-0,21
Antipatia	0,07	0,93	0,10	0,11	-0,92	-0,04	0,17	0,24
Paulistanidade	-0,12	0,11	1,00	-0,11	-0,25	0,31	0,85	-0,20
Eigenvalue	2,39	1,79	1,75	1,15	2,40	1,84	1,20	1,10
% Variância	0,30	0,22	0,22	0,14	0,30	0,23	0,14	0,14
%Acumulativa	0,34	0,55	0,74	0,74	0,30	0,53	0,68	0,82

Fonte: adaptado de Santos (2020)

No geral, os números apresentados na tabela 2 mostram que as respostas dadas pelos ouvintes, inicialmente em 8 escalas, podem ser reduzidas a 4 componentes,

tanto em São Paulo quanto em São Luís. Entretanto, tal tabela é apresentada para que se possa visualizar a diferença de como se correlacionam as respostas dadas por ouvintes paulistanos e ouvintes ludovicenses.

Em São Paulo, o primeiro desses componentes engloba as respostas nas escalas de educação, amigabilidade e antipatia, que se correlacionam negativamente entre si e que foram denominadas *antipatia*. Aqui, quando um falante foi ouvido como alguém que soa mais antipático, também foi ouvido como alguém que soa menos amigável. É importante não confundir educação com escolaridade, em que soar educado é o mesmo que soar polido ou cordial.

O segundo componente, denominado *competência*, é integrado pelas respostas nas escalas de escolaridade e inteligência. Essas duas se correlacionam positivamente, ou seja, quando um falante é percebido como alguém que soa mais escolarizado, também é percebido como mais inteligente.

O terceiro componente é integrado apenas pelas respostas à escala de *paulistanidade*, que não se correlacionam às de nenhuma outra e, por isso, o componente foi denominado pelo mesmo nome. A mesma decisão foi tomada para o componente *seriedade*.

Para os dados de São Luís, interessa verificar que os componentes identificados nas respostas dadas por ludovicenses são diferentes daquelas que vimos por paulistanos, ao menos na ordem em que esses componentes são identificados. O primeiro componente é o de competência e engloba as respostas nas escalas de escolaridade e inteligência, positivamente correlacionadas entre si. Observe-se que, em São Paulo, esse componente aparece em segundo. Entretanto, se nas respostas dos ludovicenses esse componente é o primeiro componente que explica a variação, nas respostas dos ludovicenses é o último componente. Em outras palavras, competência, na sua associação ao par subjuntivo/indicativo, é algo mais forte para ouvintes paulistanos do que para ouvintes ludovicenses. O segundo componente, denominado *antipatia* e que nos dados paulistanos é o primeiro componente identificado, é composto pelas respostas nas escalas de amigabilidade e antipatia. Esse mesmo componente, para os dados de São Paulo, inclui ainda a escala de educação. Para São Luís, a correlação também é negativa, no sentido que, quando um falante é percebido como mais antipático, ele também é percebido como menos amigável.

O terceiro componente, *paulistanidade*, equipara-se ao modelo de análise para São Paulo, no sentido de que essa escala não se correlaciona a nenhuma outra, e, por fim, o componente *formalidade* engloba as escalas de escolaridade e inteligência, também em correlação positiva.

A partir daí, tomaram-se esses componentes como variáveis dependentes, e passou-se às análises estatísticas. Apresentam-se os resultados das análises de regressão feitas no R para as escalas de *competência* e *paulistanidade*. O subjuntivo e o indicativo passam a variáveis preditoras, bem como os falantes e as interações entre eles. O ouvinte foi incluído como efeito aleatório. A falante Débora, de São Luís, e o modo indicativo são tomados como valor de *intercept* para as análises

Tabela 3: Resumo dos resultados do modelo de regressão para o CP Competência – São Luís (efeitos mistos)

Efeitos Fixos	Estimativa	Erro padrão	Valor-t	<i>p</i>
<i>(Intercept)</i>	-0,545	0,070	-7,778	<0,001 ***
Subjuntivo	1,018	0,099	10,236	<0,001 ***
Fabiano	-0,218	0,099	-2,203	<0,05 *
Juliana	-0,143	0,099	-1,440	0,150
Rodrigo	-0,142	0,099	-1,436	0,151
Fabiano:Subjuntivo	0,299	0,140	2,130	<0,05 *
Juliana:Subjuntivo	0,443	0,140	3,152	<0,01 **
Rodrigo:Subjuntivo	0,556	0,140	3,958	<0,001 ***

Fórmula: Competência ~ Modo * Falante + (1|Ouvinte), Data: PERCEPCAO.PCASL⁶
Intercept = Debora; Indicativo.
 Fonte: adaptada de Santos (2020)

Tabela 4: Resumo dos resultados do modelo de regressão para o CP Competência – São Paulo (efeitos mistos)

Efeitos Fixos	Estimativa	Erro padrão	Valor-t	<i>p</i>
<i>(Intercept)</i>	-0,626	0,053	-11,861	<0,001 ***
Subjuntivo	0,740	0,074	9,920	<0,001 ***
São Paulo	0,486	0,074	6,516	<0,001 ***
São Paulo:Subjuntivo	0,050	0,105	0,477	0,633

Fórmula: Competência ~ Modo * Cidade.Falante + (1|Ouvinte), Data: PERCEPCAO.PCASP
Intercept = Indicativo; São Luís.
 Fonte: adaptada de Santos (2020)

Para o componente competência, é significativa a diferença entre as respostas baseadas no subjuntivo e no indicativo: no geral, os falantes Juliana e Rodrigo, de São Paulo, soam mais competentes quando estão no seu disfarce com subjuntivo. Não há, contudo, diferenças significativas em São Paulo, nas respostas, para os falantes Juliano e Débora, ambos de São Luís, o que permite entender que o padrão que se acaba de estabelecer, pode ser verificado para cada um deles. Há diferença significativa entre os modos verbais, no sentido de que o subjuntivo soa como mais competente. A interação entre falante e modo não é significativa. Em São Luís, ao contrário, observa-se que há diferenças significativas nas respostas a depender do falante, de modo mais explícito: Fabiano tende a ser percebido como menos competente quando ouvido no subjuntivo,

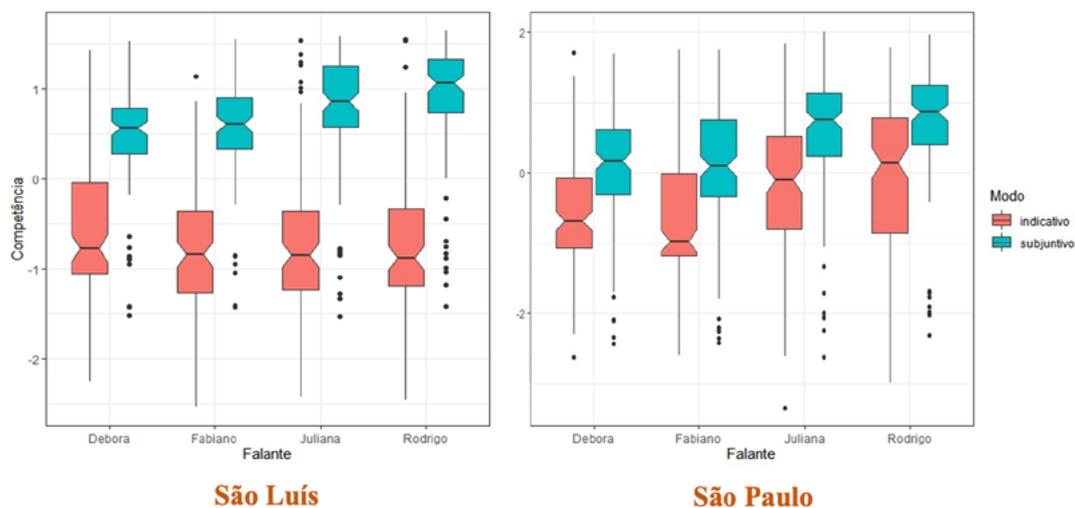
⁶ A fórmula, ao pé da tabela, indica que o CP competência foi analisado como dependente do falante e do modo verbal, bem como em função da interação entre essas duas variáveis. Além disso, a variável 'ouvinte' foi inserida como efeito aleatório, uma vez que as respostas entre os participantes podem variar muito, o que interferiria nos padrões encontrados nas análises.

relativamente a Débora, no indicativo. Também é significativa a diferença entre as respostas baseadas no subjuntivo e no indicativo. Assim, nos dados ludovicenses, a percepção de competência depende de qual falante ele ouve, e em qual modo verbal. O mesmo ocorre para a interação entre os falantes e o modo subjuntivo, no sentido de que também é estatisticamente significativa para todos os falantes, em São Luís.

Os padrões para cada falante são bastante evidentes nas respostas de ludovicenses. As diferenças entre os modos são significativamente diferentes, bem como a interação dos falantes com o modo verbal. Esses resultados mostram que, além do modo verbal, o papel do falante importa para a compreensão da percepção de competência.

Esse resultado pode ser melhor visualizado na figura a seguir, que destaca que, embora haja diferenças na percepção de competência pelos ouvintes, a depender do modo como os falantes são ouvidos, essas diferenças são mais evidentes nas respostas dadas pelos ludovicenses, comparativamente às respostas dadas por paulistanos. Santos (2015) já tinha mostrado em sua pesquisa que, na produção, paulistanos estavam à frente no que diz respeito ao uso de indicativo nas situações em que o esperado é o subjuntivo. Logo, os resultados de percepção aqui apresentados podem subsidiar discussões acerca do comportamento da produção e da percepção sociolinguísticas, no sentido de se complementarem na explicação para o avanço de formas indicativas em contextos de subjuntivo.

Figura 6: Dispersões das respostas dos ouvintes ludovicenses e paulistanos no CP competência, de acordo com o falante e com o modo verbal

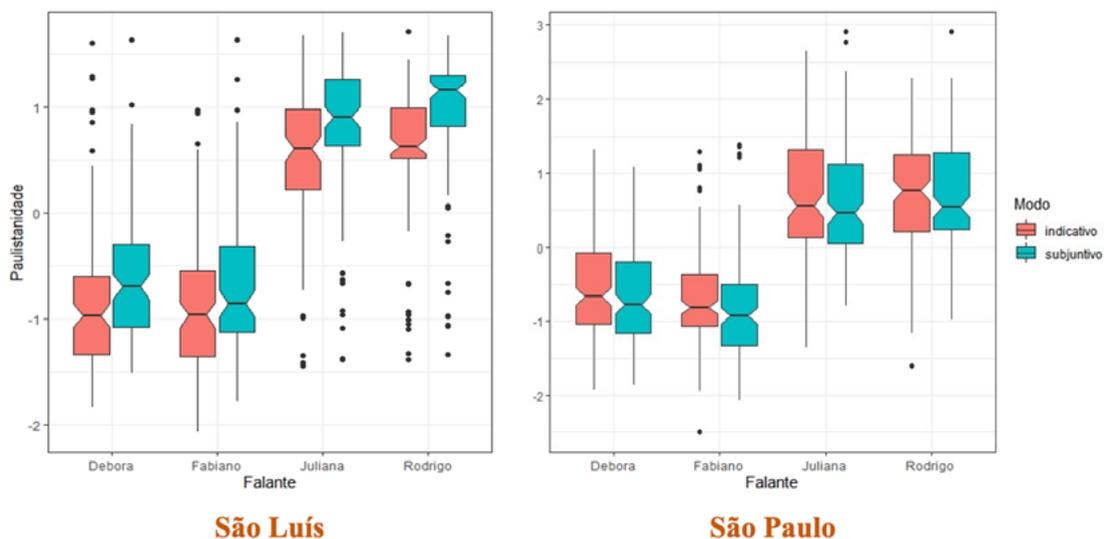


Fonte: Santos (2020: 148)

Além dos resultados para o CP competência, são apresentadas as diferenças nas respostas dos ouvintes para a percepção de paulistanidade. Os dados para as respostas de paulistanos revelam que há diferenças significativas para esses falantes, mas não para os ludovicenses. Para a percepção desse componente o modo não importa,

ou seja, o papel do falante é mais importante do que o modo na definição do que é soar paulistano.

Figura 7: Dispersões das respostas dos ouvintes ludovicenses e paulistanos para o CP paulistanidade, com base no falante e no modo (efeitos mistos)



Fonte: Santos (2020, p. 183).

De certo modo, as distribuições das respostas dadas pelos ouvintes paulistanos se aproximam das avaliações feitas por ludovicenses e paulistanos, quando foram convidados a dar suas impressões gerais sobre os trechos de fala utilizados no experimento. Aqui, do mesmo modo que nas entrevistas abertas (CAMPBELL-KIBLER, 2006)⁷, os falantes paulistanos Juliana e Rodrigo se diferenciam dos falantes ludovicenses Débora e Fabiano, mas observa-se que os padrões de respostas para o modo verbal não são muito diferentes em cada falante. Da mesma forma, entre os ludovicenses, o papel do falante faz diferença para a percepção de paulistanidade: é significativa a diferença para os falantes Juliana e Rodrigo, percebidos como mais paulistanos, comparativamente à falante Débora.

Os resultados obtidos para paulistanidade mostram que, no geral, os ouvintes percebem diferenças dialetais entre os falantes, e isso se deu mais por causa de outras características linguísticas, como, talvez, a manipulação da realização do /-r/ em coda, bem como por outros aspectos não controlados aqui, por exemplo, a prosódia ou o próprio sotaque, tomado no sentido proposto por Campbell-Kibler (2007: 32), para quem o “sotaque é uma construção carregada, que conecta padrões linguísticos a divisões sociais e econômicas entre indivíduos e grupos”⁸. Acrescente-se, para esse

⁷ Tradução proposta para o termo *open interviews*, assim chamadas por não seguirem um roteiro pré-definido.

⁸ Tradução proposta para “accent is a loaded construct, connecting linguistic patterns with social and economic divisions between individuals and groups”.

resultado, o fato de que Oushiro (2015) já tinha observado a presença desse estereótipo na fala de paulistanos que ouviram vozes femininas com /en/ ditongado, como em *fazenda*.

É importante destacar que tais distribuições mostram que *paulistanidade* não deve ser compreendida da mesma maneira para os dados ludovicenses e para os dados paulistanos, já que, *a priori*, não faz sentido responder a *quão paulistano soa um ludovicense*. De todo modo, ressalte-se o fato de que os ouvintes ludovicenses perceberam os falantes paulistanos — Juliana e Rodrigo — como mais paulistanos em seus disfarces no subjuntivo, o que não confirma a hipótese de que “deixar de usar formas subjuntivas é coisa de paulistano”. Esse resultado, portanto, pode ser lido por meio do imaginário de que os paulistanos têm acesso a mais bens de cultura do que pessoas de São Luís, por exemplo⁹.

4 CONSIDERAÇÕES FINAIS

O que os resultados alcançados aqui parecem evidenciar é que a percepção de *competência* depende da atuação do falante e do modo, em São Paulo, e da atuação do falante, do modo e da interação falante/modo, em São Luís. Por sua vez, a percepção de *paulistanidade* depende da atuação do falante, para os ouvintes paulistanos, e da atuação dos falantes e da interação falante/modo, em São Luís.

De certa forma, a hipótese de que “deixar de usar subjuntivo é coisa de paulistano” não se confirma, já que os falantes paulistanos foram percebidos como mais paulistanos em seus disfarces com subjuntivo, entre os ouvintes ludovicenses.

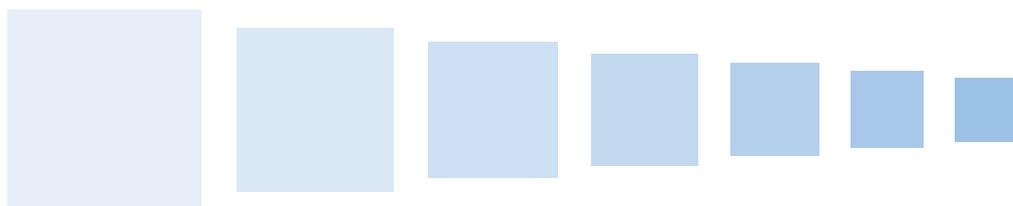
Do ponto de vista teórico, esta pesquisa se aproxima da proposta de Levon e Buchstaller (2014), no sentido de que os ouvintes percebem e avaliam variáveis morfossintáticas, afastando-se, portanto, da proposta de Labov (1993) e Labov *et al.* (2011), para quem essas mesmas variáveis estariam menos disponíveis para avaliação por parte dos ouvintes.

REFERÊNCIAS

- BAGNO, M. **Preconceito linguístico**: o que é, como se faz. 49. ed. São Paulo: Parábola Editorial, 2015.
- BERLINCK, R. de A.; BRANDÃO, S. M.; SENE, M. G. de. Desafios e caminhos na compreensão da variação sintática: design de um teste de percepção. In: CARVALHO, C. dos S.; LOPES, N. da S.; RODRIGUES, A. (Orgs.). **Sociolinguística e Funcionalismo**: vertentes e interfaces. Salvador: EDUNEB, 2020. pp. 26-58.
- BOERSMA, P.; WEENINK, D. **Praat**: doing phonetics by computer. 2015. Disponível em: <http://www.fon.hum.uva.nl/praat/>. Acesso em: 10 de jul. 2022.
- CANEVER, F. **Infinitivo flexionado em português brasileiro**: frequência e percepções sociolinguísticas. 2017. 138f. Tese (Doutorado em Linguística) – FFLCH – Universidade de São Paulo, São Paulo.

⁹ Sousa (2014) apresentou dados de uma pesquisa Datafolha mostrando que 46% dos entrevistados — que não incluía paulistanos — acreditam que São Paulo é a “locomotiva do país” e recebe mais bens de cultura que outras localidades do país.

- CAMPBELL-KIBLER, K. **Listener perceptions of sociolinguistic variables: the case of (ing)**. 2006. 282f. Tese (Doutorado em Linguística) – Stanford University, Palo Alto/California.
- CAMPBELL-KIBLER, K. Accent, (ING), and the social logic of listener perceptions. **American Speech**, vol. 82, p. 32-61, 2007.
- CAMPBELL-KIBLER, K. The nature of sociolinguistic perception. **Language Variation and Change**, vol. 21, p. 135-156, 2009.
- CUNHA, C.; CINTRA, L. **Nova Gramática do Português Contemporâneo**. Rio de Janeiro: Lexikon, 2007.
- ECKERT, P. Variation and the indexical field. **Journal of Sociolinguistics**, n. 12, vol. 4, p. 453-476, 2008.
- ECKERT, P. Three Waves of Variation Study: the emergence of Meaning in the Study of Sociolinguistic Variation. **Annual Review of Anthropology**, vol. 41. p. 87-100, 2012.
- LABOV, W. **The unobservability of structure and its linguistic consequences**. Paper presented at the 22nd New Ways in Analyzing Variation conference, University of Ottawa, October 22-25, 1993.
- LABOV, W. **Padrões Sociolinguísticos**. São Paulo: Parábola, 2008[1972].
- LABOV, W.; ASH, S.; RAVINDRANATH, M.; WELDON, T.; BARANOWSKI, M.; NAGY, N. Properties of the sociolinguistic monitor. **Journal of Sociolinguistics**, vol. 15 (4): Blackwell Publishing, pp. 431-463, 2011.
- LAMBERT, W. E., HODSON, R. C., GARDNER, R. C.; FILLENBAUM, S. Evaluational reactions to spoken languages. **Journal of Abnormal and Social Psychology**, vol. 60(1), 44-51, 1960.
- LEVON, E.; BUCHSTALLER, I. Perception, cognition, and linguistic structure: the effect of linguistic modularity and cognitive style on sociolinguistic processing. **Language Variation and Change**, vol. 27, p. 319-348, 2014.
- MELO, M. A. S. L. de. Furando as ondas: a contribuição de falantes à margem da estrutura social para a discussão sobre o significado social da variação. **Revista Linguística**. Rio de Janeiro, v. 16, Número Especial Comemorativo, p. 798-816, nov. 2020.
- MENDES, R. B. Percepções Fonéticas do /-r/ em São Paulo: principais correlações. **Todas as Letras – Revista de Língua e Literatura – Dossiê “Percepções Sociolinguísticas e atitudes: os significados sociais da variação”**, São Paulo, v. 18, n. 2, p. 133-146, maio/ago. 2016.
- MENDES, R. B. **Percepção e Performance de Masculinidades: efeitos da concordância nominal de número e da pronúncia de /e/ nasal**. 2018. 225f. Tese (Livre-docência) – FFLCH, Universidade de São Paulo, São Paulo.
- NIEDZIELSKI, N. A. The effect of social information on the perception of sociolinguistic variables. **Journal of Language and Social Psychology**, v. 18, p. 62-85, 1999.
- OUSHIRO, L. **Identidade na pluralidade: avaliação, produção e percepção linguística na cidade de São de Paulo**. 2015. 390f. Tese (Doutorado em Linguística) – FFLCH, Universidade de São Paulo, São Paulo.
- ROCHA LIMA, C. H. da. **Gramática Normativa da Língua Portuguesa**. 42. ed. Rio de Janeiro: José Olympio Editora, 2002.
- R TEAM, D. C. R. A Language and Environment for Statistical Computing. **R Foundation for Statistical Computing**, Vienna, Austria, 2022. Disponível em: <http://www.R-project.org/>. Acesso em: 10 abr.2023.
- SANTOS, W. S. dos. **A morfologia do indicativo na expressão do modo subjuntivo em São Paulo e São Luís**. 2015. 140f. Dissertação (Mestrado em Semiótica e Linguística Geral) – FFLCH, Universidade de São Paulo, São Paulo.
- SOUSA, B. Egoístas e metidos – é assim que o Brasil vê os paulistanos. **Revista Exame**, 22 de dez. 2014. Disponível em: <https://exame.abril.com.br/brasil/egoistas-e-metidos-e-assim-que-o-brasil-ve-os-paulistanos/>. Acesso em: 10 jul. 2022.
- SOUSA, B. **Percepções sociolinguísticas em São Luís e São Paulo**. 2020. 241 f. Tese (Doutorado em Semiótica e Linguística Geral) – FFLCH, Universidade de São Paulo, São Paulo.
- WEINREICH, U.; LABOV, W.; HERZOG, M. **Fundamentos empíricos para uma teoria da mudança linguística**. São Paulo: Parábola, 2006[1968].



PARTE III

ESTUDOS SOCIOLINGÜÍSTICOS EM INTERFACES



SOCIOFUNCIONALISMO: AINDA UMA CONVERSA NA DIFERENÇA?

CAMILO ROSA SILVA
CLEBER TOURINHO DE SANTANA

1 À GUIA DE INTRODUÇÃO

Habitualmente, os estudos sobre a linguagem verbal são divididos em duas grandes correntes. De um lado, está o polo denominado de funcionalismo, no qual a função codificada nos usos linguísticos assume papel preponderante; do outro, está o formalismo, abrigo de abordagens em que predomina, como propósito central de estudo, a análise da forma linguística, deixando os interesses funcionais em segundo plano (DILLINGER, 1991).

Sabemos que o dinamismo, a variação e a mudança são constantes e inerentes aos usos da língua. Ela varia, muda e se transforma, o que talvez nos faça pensar que fatores sociais ou, em sentido amplo, extralinguísticos sejam preponderantes nos estágios que a língua vai assumindo historicamente. Além disso, não se pode omitir ou negligenciar o fato de que a língua é um produto coletivo, intuitivo e tácito de qualquer comunidade. Sendo assim, a linguagem se permite desvelar numa multiplicidade de abordagens que faz as opções restritivas a um enquadramento teórico rígido ter alcance também restrito. Voltamos, assim, à célebre relação entre ponto de vista e objeto¹.

Nesses termos, apresenta-se o desafio dos estudos linguísticos contemporâneos: identificar o que é genético, inato, e o que é produto do convívio social, o que se perfaz no interior da face estrutural da língua e se substancia na riqueza dos elementos contextuais, nas condições de produção, nas intenções dos usuários. O ideal seria, talvez, descrever e analisar a língua numa mirada que coadunasse ambos os polos. No entanto, é preciso lembrar que uma pretensa confluência de abordagens distintas exige a conciliação de princípios, teses e arbítrios quase sempre de complexo equacionamento.

No âmbito da discussão possível acerca do que é acomodável em um polo e no seu antagonista, neste capítulo, pretendemos conduzir uma reflexão introdutória so-

¹ Saussure (2012).

bre a aproximação entre o funcionalismo norte-americano, cuja redundância dispensa apontar seu acento no polo funcionalista, e a sociolinguística variacionista, uma perspectiva que especialmente a partir da incorporação da noção de regra variável, poderia ser acomodada no interior da fronteira formalista².

Ao focarmos as especificidades de cada abordagem, percebemos tratar-se de duas formas distintas de encarar a mudança linguística. No entanto, essa visão não obstaculiza um desejo por movimentos mais aglutinadores, uma vez que, conforme defende Tavares (2013, p. 33), “[...] existem vários postulados teórico-metodológicos da sociolinguística e da linguística baseada no uso que guardam semelhança e, assim, podem ser relacionados para o estudo de fenômenos de variação e de mudança linguística”.

Em vista disso, nosso propósito é alçar os pressupostos centrais das duas referidas teorias, tentando identificar pontos convergentes ou aproximativos entre elas. Especificando um pouco mais esse ambiente teórico, empreendemos uma discussão introdutória a respeito das duas perspectivas de estudo que originaram o conceito, recuperando termos e noções basilares, abarcando suas origens e peculiaridades, de maneira a delinear a abordagem chamada de Sociofuncionalismo. Por coerência, defendemos essa “terceira via”, por meio da qual vem sendo possível enriquecer as pesquisas e compreender diversos fenômenos inerentes aos processos de variação e mudança linguísticas.

Para viabilizar a discussão, organizamos o texto em cinco seções que se juntam a esta introdução e à conclusão: enfocamos, inicialmente, a contextualização das principais correntes teóricas responsáveis pelo surgimento e desenvolvimento dos estudos em linguística; a seguir, situamos o surgimento da sociolinguística, seu objeto de estudo dentro do sistema vivo da língua e as percepções a respeito da noção de variação linguística; na sequência, apresentamos o campo do chamado funcionalismo norte-americano para, ato contínuo, discutir o que viria a ser, de fato, um estudo sociofuncionalista. Por último, apresentamos um “estudo de caso”, versando sobre uma pesquisa exemplar, que mostra como é possível estudar fenômenos linguísticos sob a égide da concepção sociofuncionalista, concentrando-nos nos aspectos metodológicos e seus resultados, a fim de sumarizar a contribuição desse campo emergente para a descrição e análise de fenômenos relacionados à língua em uso.

2 CONTEXTOS TEÓRICOS MAIS AMPLOS

Para a exposição das bases teóricas que delimitam este capítulo, no caso, a sociolinguística, o funcionalismo e o sociofuncionalismo, é necessário localizar, na cartografia dos estudos linguísticos, mesmo que com uma breve menção, duas das grandes correntes do pensamento linguístico, a saber: o estruturalismo e o gerativismo, que

² Segundo Camacho (2010), nem mesmo o abandono do conceito de regra variável conseguiu desviar a sociolinguística do formalismo adotado inicialmente, de tal modo que a defesa das motivações formais como prioritárias em relação às funcionais são determinantes para reivindicar e reafirmar, no ponto de vista laboviano, a autonomia do formal.

trouxeram — e trazem — importantes contribuições para a consolidação dos estudos sobre a linguagem verbal.

O estruturalismo, corrente linguística inserida no chamado polo formalista, de acordo com Dillinger (1991), diz respeito aos estudos que priorizam a *forma* linguística. Nessa corrente, a língua é entendida como um sistema autônomo que privilegia suas próprias características internas, seus constituintes e os vínculos entre eles. Desse modo, a língua é examinada nas relações entre suas partes e princípios que regulam sua organização, ocasionando explicações advindas da própria estrutura. Segundo Harris (1989), a particularidade primordial do estruturalismo de Saussure é que a própria estrutura da língua dá origem às unidades e às relações entre essas unidades; logo, a estrutura existe como um todo e não se constitui em partes. Portanto, nessa concepção, a língua é isolada de fatores externos e estudada intrinsecamente. Já se percebe o quanto essa perspectiva se distancia das abordagens funcionalistas, que concebem o estudo do significado e das funções a que se destina o uso das formas linguísticas em atos comunicativos.

Por seu turno, o gerativismo, que surgiu com Chomsky no final da década de 1950, busca criar um método para descrever de forma universal as línguas humanas através de gramáticas gerativas, procurando explicar o surgimento da linguagem no cérebro, sua organização e funcionamento. Nasce com o intuito de preencher possíveis lacunas deixadas pelo estruturalismo, que era a corrente linguística predominante à época (BORGES NETO, 2011).

O modelo gerativista chomskyano pressupõe a língua como homogênea. De acordo com essa teoria, se houvesse apenas um falante da língua — tido como ideal —, isso já seria o suficiente para que houvesse um bom representante dela. Assim, o objeto de análise dessa teoria é baseado na competência linguística do falante, e um de seus objetivos é buscar respostas para o conhecimento das regras que regem as línguas. Chomsky estabelece uma distinção entre competência e desempenho linguístico, e afirma que o desempenho diz respeito ao uso concreto da língua, logo, é irrelevante para o estudo de sua teoria (DILLINGER, 1991).

Lato sensu, o que difere os modelos teóricos estruturalistas e gerativistas das abordagens funcionalistas é que estas últimas julgam inviável um estudo da língua que desconsidere o envolvimento e a atuação de fatores sócio-históricos. Isso significa que não se pode considerar a língua como algo estático; antes, é preciso percebê-la de acordo com a dinâmica de uso, considerando-se sua natureza heterogênea, enquanto forma de interação social, levando em conta que tanto a linguagem escrita quanto a oral exercem propósitos sociais e culturais específicos (TAVARES, 2013).

Feita essa rápida tentativa de ambientação teórica que visa a organizar didaticamente as abordagens em suas respectivas vitrines, é preciso visitar as bases de cada uma das vertentes convidadas à discussão — sociolinguística e funcionalismo —, recuperando a visão de como cada uma delas está fundamentada para apontar os fatores suscetíveis de potenciais aproximações, que vão culminar no rótulo sociofuncionalismo.

3 SOCIOLINGÜÍSTICA: ALGUMAS CONSIDERAÇÕES GERAIS

De acordo com García (2004), o termo “sociolinguística” aparece pela primeira vez em 1952. A definição dessa teoria pode ser esboçada se nos concentrarmos em seu objeto de estudo, a saber, os padrões de comportamento linguístico observáveis dentro de uma comunidade de fala. Sua metodologia, afeita a métodos quantitativos, busca identificar e formalizar analiticamente tais comportamentos. Desse modo, investe-se na construção de um modelo de sistema heterogêneo, “constituído por unidades e regras variáveis” (LUCCHESI; ARAÚJO, 2006, p. 1).

Numa mirada historiográfica, pode-se defender que a sociolinguística nasce da necessidade de estudar aquilo que Saussure e Chomsky pouco consideraram em seus estudos, a saber, *a fala*. Este ramo da linguística surge num contexto em que se procurava estabelecer relações entre a linguagem e elementos da natureza social e cultural, com o intuito de desenvolver uma abordagem diferente daquela instaurada pelos estruturalistas, a chamada abordagem imanente.

Para a sociolinguística, não há sociedade sem linguagem, uma vez que indivíduos e sociedade se estabelecem e se constituem na e pela língua(gem), mediante processos de interação. Logo, língua, linguagem e sociedade estão entrelaçadas e são indissociáveis, uma influenciando e determinando o comportamento da outra (CEZARIO; VOTRE, 2008).

Labov é a referência mais destacada na área da sociolinguística variacionista, por ter estabelecido, de forma mais consistente, as propostas e os métodos assumidos hoje por essa área de estudos, também denominada de teoria da variação e mudança linguística. Seus estudos revelaram não só a existência de padrões sociais verificáveis no comportamento dos falantes, mas também o fato de que a linguagem e a sociedade estão intimamente relacionadas. Portanto, se uma dada língua serve a uma comunidade multiforme, a heterogeneidade é uma consequência natural, de modo que, para compreender as mudanças de uma língua, é preciso levar em consideração seu caráter heterogêneo, assim como verificar os aspectos que a condicionam (GARCÍA, 2004).

Foi a partir de modelos matemáticos que Labov e seus seguidores desenvolveram a natureza da pesquisa utilizada na sociolinguística variacionista. Tais modelos são capazes de quantificar a influência simultânea de vários fatores sobre um vernáculo específico e, por isso, se constituem numa “ferramenta poderosa e segura que pode ser usada para o estudo de qualquer fenômeno variável nos diversos níveis e manifestações linguísticas” (NARO, 1992, p. 25).

A noção de heterogeneidade é basilar na sociolinguística, e não está, de maneira alguma, ligada a uma espécie de caos linguístico; pelo contrário, uma prova de que o caráter heterogêneo da língua a mantém organizada ou sistematizada é o fato de que os falantes de determinada comunidade estabelecem comunicação e mútuo entendimento, mesmo em meio a variações e diversidades linguísticas (BAGNO, 2007).

Nas palavras de Tarallo (2001, p. 8), as variáveis linguísticas “são diversas maneiras de se dizer a mesma coisa em um mesmo contexto, e com o mesmo valor de ver-

dade". O conceito de regra variável implica que não existe variação livre. Uma regra variável relaciona duas ou mais formas linguísticas de maneira que, quando a regra é aplicada, ocorre uma das formas e, quando a regra não é aplicada, são outras as formas requisitadas. Ou seja, a aplicação ou não das regras variáveis é condicionada por aspectos do contexto linguístico e/ou social.

Na próxima seção, ilustramos situações reveladoras da dinamicidade da língua em uso, considerando a complexidade de fatores sociais que a envolvem, o que torna a heterogeneidade linguística objeto de análise sociolinguística e passível de sistematização.

3.1 SOCIOLINGUÍSTICA: ALGUMAS CONSIDERAÇÕES ESPECÍFICAS

Um dos fenômenos linguísticos que exemplifica o que se entende por *variação linguística* nos estudos sociolinguísticos é a alternância das formas *tu* e *você* na referência à segunda pessoa do discurso. Assim, em situações concretas de uso da língua, a opção do falante em preencher o lugar de sujeito com uma dessas formas: "**Você** vai almoçar?" ou "**Tu** vai almoçar?" tem ligação com uma série de fatores que podem depender da origem do indivíduo falante, do grau de formalidade do evento interativo, do grau de monitoramento, enfim, de especificidades contextuais. Não há dúvidas de que, embora diferentes, essas formas estão sendo utilizadas com a mesma finalidade.

Convém mencionar que, assim como há terminologias específicas em outros ramos da ciência da linguagem, nos estudos da variação, não é diferente. Logo, uma importante distinção diz respeito aos termos *variável* e *variantes*. No exemplo dado anteriormente, relativo ao uso dos pronomes *tu* e *você*, o termo *variável* se aplica à expressão pronominal da segunda pessoa do singular, correspondente a um fator ou categoria da língua que se encontra em variação. E *variantes* se aplica às formas individuais, no caso os pronomes *tu* e *você*, duas formas que concorrem em uma variável.

Fato é que a variação linguística não se dá aleatoriamente, pois existem condicionadores que regulam a escolha entre uma ou outra variante. Eles ajudam a delimitar quais são os contextos mais propícios para a ocorrência das variantes de um estudo e se dividem em dois grandes grupos, de acordo com a condição de estarem mais ligados às características internas ou externas do sistema linguístico. De tal modo, conforme lembra Coelho (2010, p.28), "os condicionadores linguísticos e sociais são também tratados por variáveis independentes (ou grupos de fatores), enquanto a variável propriamente dita também pode ser tratada por variável dependente".

Outra noção importante na sociolinguística é a ideia de *vernáculo*, que é a língua falada em situação natural de comunicação, estando especialmente presente nas narrativas de experiência pessoal. Para Tarallo (2001), trata-se da língua falada no dia a dia, nas rodas de amigos, na família, em parques e clubes, nos pátios das escolas, bem longe da monitoração dos professores e outros agentes "reguladores".

Tarallo (2001) explica, ainda, que o objetivo da coleta de dados sociolinguísticos não é apenas descrever o vernáculo, pois os dados coletados que se apresentarem como “não naturais” podem ser utilizados como centelha para novos objetivos, por exemplo, para o estudo da variação estilística do desempenho do informante. Além disso, o escopo da pesquisa sociolinguística é, primariamente, verificar “como as pessoas falam quando não estão sendo sistematicamente observadas” (COELHO *et al.*, 2012, p. 116), embora, para fins técnicos, existam estratégias que ajudam a contornar as dificuldades de se conseguir esse material oralizado.

O modelo laboviano, também conhecido como *sociolinguística quantitativa*, consiste na utilização de um tratamento estatístico aplicável à linguagem, sendo possível apurar a relevância de determinados fatores, linguísticos e sociais, que influem na aplicação de uma regra variável na língua. Ademais, vislumbra-se a possibilidade de se pensar em mudança linguística através de um maior ou menor uso de alguma das variantes identificadas como proeminentes (KOELLING, 2005; GUY; ZILLES, 2007). De fato, a variação linguística “regularmente mostra taxas mais altas ou mais baixas de ocorrência em determinados ambientes” ou “predomina em determinados grupos sociais ou estilos de fala” (GUY; ZILLES, 2007, p. 48).

Tentando rastrear genericamente as investigações sobre regras variáveis desenvolvidas nos centros de pesquisa existentes no Brasil, vem-nos automaticamente à memória alguns projetos, tais como o do IEL/Unicamp; o Projeto Vertentes/UFBA; o ValPB/UFPB; o Varsul/UFSC, entre outros. Todos eles fazem uso de programas computacionais específicos para o tratamento de dados coletados em entrevistas sociolinguísticas. Um dos mais utilizados é o pacote de programas Varbrul (acrônimo em inglês para *Variable Rules*) e suas atualizações mais recentes, o GoldVarb e o Rbrul. Para Guy e Zilles (2007, p. 106), “o Varbrul tem certas vantagens que fazem dele uma boa opção para o sociolinguista. Em primeiro lugar, ele é dedicado à estruturação dos dados que encontramos na linguagem natural”.

É importante, neste ponto, citar, no âmbito da avaliação linguística, alguns itens da taxonomia proposta por Labov (2008) o trinômio: *indicadores*, *marcadores* e *estereótipos*, a fim de encaminharmos uma análise desse fator nas variedades linguísticas.

São considerados **indicadores**, as variáveis linguísticas nas quais é possível observar variação social, mas não variação estilística, estando a avaliação “abaixo do nível da consciência” (LUCCHESI, 2015, p. 80), ou seja, as variedades que têm pouca força avaliativa na comunidade de fala (SEVERO, 2011). **Marcadores** são aqueles traços linguísticos que recebem avaliação social e estilística estratificada, em geral apontados como o ideal de fala/escrita, cujo uso causa nos interlocutores uma avaliação positiva ou de determinação de origem social/educacional. Para Labov (2008, p. 360), “exibem estratificação estilística, tanto quanto estratificação social”. Quanto aos **estereótipos**, também conforme Labov (2008, p. 360), eles representam as “formas socialmente marcadas, rotuladas enfaticamente pela sociedade”.

No caso do Brasil, avaliações categorizadas como **estereótipos**, quase sempre negativas, são mencionadas por Lucchesi (2015, p. 80) e por Freitag *et al.* (2016, p. 141),

que se referem ao *-r* retroflexo como marca da fala caipira, em São Paulo, à oclusão de /t/ e /d/ antes de /i/ como fala nordestina (em oposição ao fenômeno da palatalização das oclusivas alveolares) e ao rotacismo do /l/ em sílaba travada, como nos pares assembleia/assembleira; bicicleta/bicicreta; Clóvis/Cróvis, como fala popular/não escolarizada.

Concluído esse esboço com noções básicas acerca da sociolinguística variacionista, cuidamos, na seção seguinte, de apresentar os pilares da vertente funcionalista, remetendo, mais especificamente, ao chamado funcionalismo linguístico norte-americano e a seus diálogos com outros desdobramentos também funcionalistas.

4 FUNCIONALISMO: ALGUMAS CONSIDERAÇÕES GERAIS

O ponto de partida para o que consideramos hoje como estudos funcionalistas remete à década de 1970, nos Estados Unidos, quando os linguistas Sandra Thompson, Paul Hopper e Talmy Givón começam a empreender estudos tomando como parâmetro a língua em sua feição maleável e socialmente influenciada. Nesse ambiente, o funcionalismo linguístico, segundo Silva e Oliveira (2017, p. 249), nutre-se das contribuições dos mencionados estudiosos, as quais se fundamentavam na tendência de “[...] observar a língua no âmbito contextual, levando em consideração a situação extralinguística”. Já nos anos 90, o movimento se redimensiona, dando ênfase à ação interativa dos usuários. Desse modo, o traço identitário de uma abordagem funcionalista passa a ser a observação da língua na emergência das situações de uso, que promoveriam mudanças na própria estrutura, em todos os seus níveis, neles incluídos os contextos linguístico-gramaticais, internos ao sistema, e os extralinguísticos, socialmente motivados (SILVA; OLIVEIRA, 2017).

Contextualizado no quadro dos estudos linguísticos delineado à época, o funcionalismo emerge como uma reação à perspectiva formalista, especialmente se contrapondo aos princípios gerativistas. Oliveira (2003, p. 96-7) deixa claro que os funcionalistas

[...] criticam o formalismo pelo fato de os formalistas estudarem a língua como um objeto descontextualizado, sem levar em consideração os falantes-ouvintes ou as circunstâncias nas quais a língua é usada. Para os funcionalistas, a língua não pode ser desvinculada de suas relações com as diversas maneiras de interação social.

Dillinger (1991) compara essas duas escolas e sustenta que, apesar de tanto o funcionalismo quanto o formalismo tratarem do mesmo fenômeno (a língua), percebem-na de maneiras diferentes e, portanto, lançam mão de metodologias diferentes para estudar os diversos fenômenos. Na visão de Oliveira (2003, p. 97), “[...] funcionalismo e formalismo não podem ser comparados nem vistos como alternativos”.

No Brasil, mesmo com alguns trabalhos funcionalistas realizados entre o fim da década de 1970 e ao longo da de 1980, predominavam na academia o gerativismo e a

sociolinguística. Por isso, toma-se como marcante a publicação do artigo *Mecanismos funcionais do uso da língua função e forma* (NARO; VOTRE, 1992), considerado um “texto fundador” no âmbito nacional, em relação às investigações funcionalistas, tendo gerado muitas reações tanto pelo aceite da publicação na revista DELTA como pelas discussões relacionadas à abordagem funcional (CASTANHEIRA; CASEIRA, 2020).

Furtado da Cunha e Tavares (2016, p. 14) enfatizam que a versão chamada linguística baseada no uso assume a linguagem como estritamente enraizada no uso comunicativo, na atividade social diária, sem a qual sequer poderia ser moldada para acomodar o fenômeno da mudança. Numa perspectiva funcionalista, por exemplo, deve-se estudar a sintaxe sempre atrelada ao discurso. Nesse sentido, é vital a presença de conceitos como marcação, planos discursivos, iconicidade, informatividade, prototipicidade e projeção metafórica. Alguns deles são ligeiramente abordados na subseção que segue³.

4.1 FUNCIONALISMO: ALGUMAS CONSIDERAÇÕES ESPECÍFICAS

A noção de gramática funcionalista é a de um sistema aberto e maleável, logo, passível de se ajustar a pressões do discurso, do próprio uso no dia a dia dos falantes, enfim, nas interações comunicativas que os seres humanos praticam durante a vida. Por isso, Ford, Fox e Thompson (2003, p. 122) a definem como “[...] um conjunto de formas, padrões e práticas que surgem em serviço às funções que os falantes precisam desempenhar mais frequentemente”. Nesse sentido, devem-se analisar as formas linguísticas de acordo com as funções que desempenham nas mais variadas situações sociodiscursivas e interacionais, conforme condicionamentos evidenciados pelos dados e contextos, tendo em vista o momento da interação (SILVA; OLIVEIRA, 2017).

A fim de oferecer uma explicação eficiente da organização e funcionamento da gramática, o funcionalismo se baseia em princípios de natureza cognitiva, sendo os mais recorrentes a marcação, a iconicidade, a gramaticalização, a prototipicidade, o dinamismo comunicativo e a coerência temática (GIVÓN, 2001). Tavares (2013, p. 38) sinaliza também os princípios da “economia, da expressividade retórica, da marcação estilística, da marcação identitária, além de princípios relacionados ao próprio processo de gramaticalização, como o princípio da persistência e o princípio da especialização”. Desses todos, pela presença mais recorrente nas análises, destacamos os seguintes:

Iconicidade: refere-se à possibilidade de isomorfismo motivado funcionalmente entre estruturas morfossintáticas e suas funções semânticas ou pragmáticas correspondentes (FURTADO DA CUNHA; TAVARES, 2016). Para Givón (1995), esse prin-

³ Devido a restrições de espaço neste capítulo e aos próprios objetivos que orientam a presente exposição, não nos alongamos na descrição desses tópicos. Ao leitor interessado em ampliar conhecimentos sobre eles, recomendamos a leitura de FURTADO DA CUNHA, M. A.; OLIVEIRA, M. R. de.; MARTELOTTA, M. E. (Orgs.). *Linguística funcional: teoria e prática*. São Paulo: Parábola, 2015.

cípio se refere à correlação motivada entre forma e função, tanto no nível sintático quanto no morfológico e no morfossintático.

Marcação: refere-se à presença *versus* ausência de determinadas propriedades nos membros em competição no interior das categorias linguísticas. Givón (2001) estabelece três critérios para a análise da marcação: complexidade estrutural, complexidade cognitiva e distribuição de frequência. Atualmente, as análises funcionalistas que aplicam esses critérios geralmente o fazem na perspectiva da escalaridade, considerando o aparato contextual que envolve o uso de duas ou mais formas linguísticas concorrentes.

Gramaticalização: uma das mais referenciadas definições de gramaticalização deve ser creditada a Hopper e Traugott (1993, p. XV). Para eles, ela consiste num “processo pelo qual itens lexicais e construções passam, em determinados contextos linguísticos, a servir a funções gramaticais, e, uma vez gramaticalizados, continuam a desenvolver novas funções gramaticais”.

De um modo geral, esse fenômeno tem sido compreendido como um processo de mudança que opera de forma predominante numa só direção, resultando na alteração da relação forma/função, considerando que as formas estão sempre sujeitas às pressões gerais da cognição e do uso. A recorrência do uso faz com que as formas se rotinizem, tornem-se regulares, e conseqüentemente sejam absorvidas pela comunidade linguística. (SILVA; OLIVEIRA, 2017)

Martelotta (2011) apresenta exemplos bem ilustrativos, contemplando a abrangência que a mudança por gramaticalização pode alcançar: (i) passagem de vocábulo livre a afixo; (ii) passagem de verbo pleno a auxiliar; (iii) passagem de advérbio a conjunção; (iv) passagem de advérbios de modo a modalizador; (v) passagem de advérbios de modo a marcador discursivo. Além disso, ela também pode ser entendida como responsável por outros fenômenos de mudança, tais como a mudança no preenchimento da posição de sujeito no português brasileiro e na ordenação de advérbios, do latim ao português.

Após essa breve contextualização de ambas as vertentes teórico-metodológicas, a sociolinguística e o funcionalismo, passamos, na próxima seção, a discutir a abordagem denominada sociofuncionalismo. Manteremos o foco nos fatores que dão consistência à junção teórica e tornam a “terceira via” uma perspectiva produtiva, especialmente para análise de fenômenos de mudança linguística.

5 SOCIOFUNCIONALISMO: CAMINHOS QUE SE ATRAVESSAM?

Martelotta (2011) traça um paralelo entre sociolinguística e funcionalismo, explicando que a primeira está totalmente centrada no uso e toma a variação e a mudança como seus principais objetos de investigação. Entretanto, para o autor, o modelo laboviano não alcança todos os tipos de mudança linguística, a exemplo da questão semântico-pragmática, que conduz à disfunção entre forma e significado, vista es-

pecialmente nos itens polissêmicos ou multifuncionais de determinado elemento da língua; ou seja, a uma mesma forma associa-se mais de uma função. Logo, tais espécies de mudança estariam mais afeitas a um estudo referente à gramaticalização, que oferece outro “modelo” explicativo quanto à variação não da forma, mas de sentido/função, motivada pelo uso.

Na esteira dessas observações em relação à sociolinguística e em função dos anseios de novas perspectivas criadas pelo funcionalismo dos anos 1980, surgiram estudos que buscavam investigar fenômenos de variação e de mudança linguística. Nesse sentido, ao invés de promover a discórdia acadêmica, pretendia-se coadunar os pontos de contato entre as duas teorias e permitir que determinados pressupostos teóricos da sociolinguística variacionista e do funcionalismo linguístico norte-americano fossem articulados e aplicados às pesquisas linguísticas.

No Brasil, os primeiros estudos inseridos na seara do sociofuncionalismo vêm do Programa de Estudos sobre o Uso da Língua (PEUL/RJ). Merece destaque Claudia Roncarati que, em 1996, estudou a negação na língua falada e, depois, a variação e a aquisição da negação em L2 (português de contato) e em L1, primando pela utilização dos postulados de ambas as teorias, com resultados que influenciaram, com o tempo, diversos estudos pelas regiões do país (NEVES, 1999).

As análises continuaram e ainda são presentes nos estudos a respeito da língua portuguesa do Brasil. Tavares, no artigo “Sociofuncionalismo: um duplo olhar sobre a variação e a mudança linguística”, publicado em 2013, elenca mais de 30 produções muito relevantes produzidas até aquele momento. Certamente, não seria irresponsável afirmar que o passar do tempo consolidou o reconhecimento à produtividade da abordagem sociofuncionalista em diversos centros de pesquisa do país.

Nesse contexto, considerando a aglutinação de princípios convergentes que aproximam as duas teorias originais, Tavares (2003, p. 101), que insinua haver similitudes dessa aludida junção com uma união conjugal, afirma:

[...] o casamento teórico não incorpora os conceitos de cada teoria ‘mãe’ exatamente como foram propostos originalmente, mas sim se fundamenta na interpretação dada a eles pelos pesquisadores casamenteiros que levam avante a junção. Não se trata da soma ou da combinação de pressupostos teórico-metodológicos de um modelo e de outro, e sim do estabelecimento de pressupostos que resultam da conversa entre os modelos. A cada conversa ocorrem novas convergências e os conceitos são alterados, definindo-se como seres voláteis, transitórios, filiados ao momento e, dessa guisa, sujeitos a reinterpretções e a revisões constantes. Em decorrência, na trajetória de avanço das discussões, o sociofuncionalismo constitui-se e reconstitui-se (TARALLO, 2003, p. 101).

Isso posto, é preciso localizar onde sociolinguística e funcionalismo se tocam, e de que maneira esse encontro pode produzir novos ou complementares modos de se estudar a linguagem.

Resumidamente, uma pretensa fortuna crítica que liste estudos sociofuncionalistas produzidos no Brasil vai fatalmente fazer saber que as duas teorias: 1) encaram

o fenômeno da linguagem e a própria língua sob uma perspectiva heterogênea; 2) aceitam a mudança linguística como fenômeno ligado à variação, de caráter gradual, contínuo e socialmente influenciado; 3) interpretam os dados sempre relacionando-os a fatores sociais, cuja “pressão” permite a existência da variação e mudança; 4) percebem que os dados sincrônicos são evidência pontual de uma mudança diacrônica; 5) têm como base a análise da língua conforme seu uso real, especialmente de maneira espontânea, vernacular; 6) advogam o princípio do uniformitarismo, isto é, que os mesmos fatores linguísticos e extralinguísticos que atuam hoje sobre a língua também atuaram no passado, desenvolvendo, assim, o processo de variação e mudança; 7) trabalham com diferentes níveis da língua (fonológico, morfológico, sintático, semântico); 8) tomam como essencial a frequência de uso nas pesquisas, ou seja, usa-se a estatística como forma de compreensão de certos fenômenos⁴ (GÖRSKY; TAVARES, 2013; TAVARES, 2013; TAVARES; GÖRSKY, 2015).

A despeito de as convergências serem robustas, não existe consenso entre os estudiosos sobre a possibilidade de junção das duas teorias; por vezes, sequer se admite uma aproximação mais íntima entre elas. Não raramente, vem à tona uma virtual impossibilidade de alinhamento entre variação e discurso que parece agredir o princípio da iconicidade, ou seja, na relação motivada entre forma e função, especialmente quando defendida numa improvável biunivocidade.

Görsky e Tavares (2013) chamam a atenção para outra razão, a nosso ver mais grave, quanto às críticas negativas: admitir a validade de estudos sociofuncionalistas seria um acinte ao orgulho acadêmico de alguns, uma vez que se estaria admitindo que o campo de estudo ao qual o pesquisador está vinculado seria “incompleto”, e necessitaria de outra teoria para preencher lacunas teóricas ou metodológicas, a fim de descrever e analisar com precisão os fenômenos linguísticos em estudo.

Nesse contexto, concluímos esta seção dando eco às palavras de May (2009, p. 71): “o sociofuncionalismo é uma perspectiva viável de pesquisa, que em seu hibridismo possibilita uma aplicação não estanque, maleável de acordo com o objeto e com os objetivos de pesquisa que se apresentam a esse novo tipo de ‘linguista camaleão’”.

5.1 A PESQUISA SOCIOFUNCIONALISTA: CONSIDERAÇÕES PONTUAIS E AMPLA DEFESA

Neste ponto, interessa-nos entender de que forma uma análise baseada numa perspectiva sociofuncionalista pode colaborar para a descrição de fenômenos linguísticos em processo de variação e mudança, ou para esclarecer como a gramática vem sendo influenciada por fatores relacionados ao uso, de forma que tal estudo seja pertinente cientificamente. Para isso, resolvemos proceder a um “estudo de caso” que se apropria de uma das produções acadêmicas que tratam desse tema e servem de refe-

⁴ Silva e Oliveira (2017, p. 256) destacam, entretanto, que “para o Funcionalismo, a frequência prevê regularidade da gramática e, para a Sociolinguística Variacionista, frequência prevê o índice de difusão sociolinguística [...]”.

rência constante sobre o assunto, escrutinando os objetivos, metodologia e resultados, a fim de compreender o processo de análise.

Com esse propósito, visitamos a tese de doutorado da pesquisadora Maria Alice Tavares, defendida em 2003, na Universidade Federal de Santa Catarina⁵. Este é, a nosso ver, um dos documentos mais importantes no tabuleiro do sociofuncionalismo brasileiro, não apenas pelo tratamento dos dados, mas especialmente pela habilidade e profundidade teórica e metodológica com que trata o assunto.

No referido trabalho, Tavares estuda um possível processo de gramaticalização dos conectores sequenciadores *e*, *aí*, *daí* e *então*, analisando fenômenos de estratificação/variação e de mudança no domínio funcional, a que chama “sequenciação retroativo-propulsora de informações”. Para a autora, na relação coesiva existente entre um enunciado precedente e um posterior, alguns conectores podem atuar gerando a expectativa de que algo novo será introduzido no discurso, em continuidade e consonância com o que já fora referido. A fonte de dados para análise é a língua falada na cidade de Florianópolis, a partir de entrevistas do banco de dados do Projeto VARSUL (Variação Linguística Urbana na Região Sul do Brasil) e dados extraídos de textos escritos em português do século XIII ao século XX.

Percebemos, ao longo de suas páginas, a preocupação, conforme a própria pesquisadora assume, de estabelecer uma “conversa na diferença”⁶, ou seja, de conciliar aspectos funcionalistas e sociolinguísticos, estratégia que permite a sua produção alcançar um patamar de excelência teórica. Primeiramente, define um *corpus* de análise dividido em duas partes: 1) textos escritos antigos, sendo os principais: *o Foro Real*, “por ser um dos mais antigos escritos em língua portuguesa, o que permite mostrar funções desempenhadas por *e*, *aí*, *daí* e *então* nos primórdios da língua” (TAVARES, 2003, p. 132, grifos nossos); e *As Vinhas da Ira*, a fim de perceber um quadro mais completo de funções marcadas por eles no início do século XX; 2) dados da fala atual de Florianópolis, disponíveis no banco de dados do Projeto VARSUL.

Em seu estudo, realiza inicialmente uma ampla análise funcionalista, delineando os percursos de gramaticalização de seu objeto de estudo (*e*, *aí*, *daí* e *então*), permitindo ao leitor compreender a aplicação da sequenciação textual, sequenciação temporal, introdução de efeito, retomada e finalização do “*e*”; a dêixis locativa, anáfora locativa, anáfora temporal, sequenciação temporal e introdução de efeito do “*aí*”; a dêixis locativa, anáfora locativa, anáfora discursiva/introdução de efeito, introdução de efeito, sequenciação textual, finalização e sequenciação temporal do “*daí*”; e a anáfora temporal, sequenciação temporal e introdução de efeito — além da sequenciação textual e finalização — do “*então*”.

⁵ O trabalho foi publicado em versão livro, em 2014, pela Editora da Universidade Federal do Rio Grande do Norte (EDUFRN).

⁶ A expressão remete ao célebre artigo de Roberta Pires de Oliveira, publicado na Revista DELTA, intitulado “Uma história de delimitações teóricas: trinta anos de semântica no Brasil” (PIRES DE OLIVEIRA, 1999).

Em seguida, passa a testar sua hipótese de maneira sociolinguisticamente orientada, voltando-se para os recortes de fala coletados nos bancos de dados. Assim, ela parte para uma análise do conjunto das influências linguísticas e sociais visando traçar a distribuição estatística do *e*, *aí*, *daí* e *então* na fala florianopolitana, usando, para isso, cinco grupos de fatores condicionadores linguísticos e três extralinguísticos, quais sejam: subfunções sequenciadoras, tipos de discurso, níveis de articulação discursiva; graus de conexão, traços semântico-pragmáticos do verbo da oração introduzida pelo conector, idade, escolaridade e sexo. Em seguida, codifica os seus achados e os submete ao programa Varbrul, amplamente usado na sociolinguística quantitativa, a fim de fornecer a frequência e o peso relativo dos fatores em relação à variável dependente, de maneira que pudesse perceber a influência deles sobre o uso de cada uma das camadas/variantes.

Como resultado de seus estudos, Tavares (2003) observa tendências gerais quanto à distribuição dos sequenciadores em grupos sociais obedecendo a duas motivações: 1) valorização social atribuída aos conectores — aparecimento com mais frequência na fala de indivíduos mais jovens, de menor escolaridade e de sexo masculino; 2) Marca de identidade, em que falantes mais jovens são os que mais utilizam os conectores de menor *status* e/ou inovadores, tomando-os como marcas típicas do grupo de pares. Além disso, o estudo sociofuncional percebe tendências gerais quanto à distribuição dos sequenciadores relativos a três princípios: i) O princípio da marcação: no qual se espera que “fatores definidos como mais marcados atraiam com mais frequência os conectores mais marcados, e os fatores definidos como menos marcados favoreçam os conectores menos marcados”; ii) o princípio de persistência: isto é, que “traços dos usos passados do conector tendem a se aderir a ele, mantendo-se ao longo do processo de gramaticalização, e interferem em sua distribuição linguística”; e iii) o princípio da marcação e o princípio da persistência como forças coatuantes em competição ou como “forças complementares ou como ambas, isto é, atuando coadunadas para alguns dos conectores e como opostas para outros” (TAVARES, 2003, p. 266).

Na referida tese, aqui observada como exemplo, percebemos que os passos para um estudo sociofuncional demandam i) levar em consideração o percurso de gramaticalização de uma forma, identificando-se os domínios funcionais possíveis ao percurso dessa e de outras formas que desempenham a mesma função em um mesmo período de tempo, de modo que, no campo sociolinguístico, possamos tomá-las como “camadas” ou variantes; ii) testar os grupos de fatores diversos a fim de restringir os fatores a serem percebidos como relevantes (linguísticos, discursivos, estilísticos e socioculturais), visando descobrir, posteriormente, os padrões de variação das formas envolvidas; e iii) interpretar os padrões de variação como indícios de gramaticalização, lendo os dados tanto de maneira funcional quanto sociolinguística (TAVARES, 2013).

O exame aqui reportado desse estudo tão extraordinário mostra um evidente amálgama entre as perspectivas funcional e sociolinguística. A pesquisa de Tavares (2003), embora já quase bidecenal, permite esboçar um percurso modelar de ataque

aos problemas que possivelmente poderiam decorrer desse diálogo entre teorias. A análise funcionalista foi inicialmente considerada, tendo sido levantados fatores discursivos ligados ao tema de estudo, que envolviam as funções possíveis aos elementos *e, aí, daí e então*. Assim foi procedido, não por atribuir mais importância aos aspectos funcionais, mas porque, dessa maneira, a autora poderia se valer de salientes exemplares de análise para a posterior comparação sociolinguística, cujos resultados retornariam para convalidar ou rechaçar as hipóteses anteriormente propostas.

A partir desses movimentos metodológicos, imprimiu-se uma retroalimentação teórica: a análise funcionalista subsidiando a questão sociolinguística, que confirmaria a posição primeira, ou a tornaria discutível. Com efeito, a abordagem realizada não parece recorrer aleatoriamente a um ou outro paradigma, como se buscasse socorro para impasses descritivos. O cenário que se afirma é o da solidez de uma perspectiva sociofuncionalista delineada em suas próprias demarcações.

6 CONSIDERAÇÕES CONCLUSIVAS

Uma das tarefas da sociolinguística é descrever as línguas em sua diversidade funcional e social, lançando mão de estratégias estatísticas e interpretativas. Predominantemente, os seguidores de Labov e sua equipe se preocupam em descrever uma variedade linguística, elegendo como problema maior as fronteiras de uso e a avaliação social, com possibilidade ou não de mudança.

O funcionalismo, por seu turno, busca basicamente descrever a linguagem no cotejo da relação função/forma, revelando a instrumentalidade discursiva que nela se revela, conforme as pressões das situações sociais. Ainda, cuida de explicar as diversas motivações que materializam as funções discursivas, assumindo, entre outras, a responsabilidade de descrever os estágios de mudança funcional verificáveis nos processos de gramaticalização.

Percebe-se que se trata de perspectivas teóricas distintas entre si. A pesquisa analisada como estudo de caso, no entanto, mostrou que uma abordagem sociofuncionalista para fenômenos em perceptível processo de variação e mudança é produtiva. O recurso ao rigor metodológico que articule princípios das duas abordagens pode conduzir a resultados satisfatórios, que provavelmente não seriam atingidos se a opção fosse por uma ou outra teoria isoladamente. O sociofuncionalismo, como se comprova no exemplo exposto, promove uma convergência interpretativa coerente e cientificamente validada.

De fato, já faz algum tempo que os “casamentos” entre teorias vêm sendo articulados, e seus agenciadores apontam sempre boas razões para celebrá-los. No caso específico da união entre sociolinguística e funcionalismo, nos parece que as afinidades são tamanhas, que não seria problemático abstrair as incompatibilidades, investindo-se numa convergência harmoniosa e produtiva. Entre os fatores que garantiriam a convergência de interesses, é possível destacar, mas não sem se limitar a isso, o quanto

é produtiva a utilização de ferramentas da sociolinguística variacionista, especialmente para explicitar, a partir de elementos quantitativos, aspectos funcionais como os de ordem pragmática, mesmo os alocados nos módulos sintático, semântico e discursivo, que materializam os usos da língua.

REFERÊNCIAS

- BAGNO, M. **Nada na língua é por acaso**: por uma pedagogia da variação linguística. São Paulo: Parábola, 2007.
- BORGES NETO, J. O empreendimento gerativo. In: MUSSALIM, F.; BENTES, A. C. (Orgs.). **Introdução à linguística**: fundamentos epistemológicos. v.3. São Paulo: Cortez, 2004. p. 93-130.
- CAMACHO, R. G. O formal e o funcional na teoria variacionista. In: RONCARATI, C.; ABRAÇADO, J. (Orgs.). **Português brasileiro**: contato linguístico, heterogeneidade e história. Rio de Janeiro: 7Letras/FAPERJ, 2003. p. 55-65.
- CASTANHEIRA, D.; CASEIRA, C. Funcionalismo norte-americano: história e perspectivas. **Revista da ABRALIN**, Aracaju, v. 19, n. 2, p. 1-4, 16 jul. 2020.
- CEZARIO, M. M.; VOTRE, S. Sociolinguística. In: MARTELOTA (Org.). **Manual de linguística**. São Paulo: Contexto, 2008. p. 141-155.
- COELHO, I. L. *et al.* **Sociolinguística**. Florianópolis: LLV/CCE/UFSC, 2010.
- DILLINGER, M. Forma e função na linguística. **D.E.L.T.A.**, São Paulo, v. 7, n. 1, p. 395-407, 1991.
- FORD, C. E.; FOX, B. A.; THOMPSON, S. A. Social Interaction and Grammar. In: TOMASELLO, M. (Ed.). **The new psychology of language**: cognitive and functional approaches to language structure. v. 2. Mahwah, NJ/London: LEA, 2003. p. 119-143.
- FREITAG, R. M. K. *et al.* Avaliação e variação linguística: estereótipos, marcadores e indicadores em uma comunidade escolar. In: FREITAG, R. M. K. *et al.* Sociolinguística e política linguística: olhares contemporâneos. São Paulo: Blucher, 2016.
- FURTADO DA CUNHA, M. A.; TAVARES, M. A. Linguística funcional e ensino de gramática. In: FURTADO DA CUNHA, M. A.; TAVARES, M. A. **Funcionalismo e ensino de gramática**. Natal, RN: EDUFRN, 2016. p. 12-58.
- GÖRSKI, E. M. (Orgs.). **Sociolinguística e Política Linguística**: olhares contemporâneos. São Paulo: Blucher, 2016. p. 139-159.
- GARCÍA, J. S. Sociolinguística y sociología del lenguaje. Niveles de análisis sociolingüístico. Nivel macrolingüístico y microlingüístico. In: PROMOTORA ESPAÑOLA DE LINGÜÍSTICA. **Sociolinguística**. Madri: PROEL, 2004.
- GIVÓN, T. **Functionalism and grammar**. Amsterdam: John Benjamins, 1995.
- GIVÓN, T. **Syntax**. v. 1/2. Amsterdam: John Benjamins, 2001.
- GÖRSKY, E. M.; TAVARES, M. A. Reflexões teórico-metodológicas a respeito de uma interface sociofuncionalista. **Revista do GELNE**, v. 15, p. 75-97, 2013.
- GUY, G. R.; ZILLES, A. **Sociolinguística Quantitativa**: Instrumental de Análise. São Paulo: Parábola, 2007.
- HARRIS, R. Saussure on Language and Thought. In: HARRIS, R.; TALBOT, J. T. **Landmarks in linguistic thought**. The Western Tradition from Socrates to Saussure. London: Routledge, 1989.
- KOELLING, S. B. **A concordância nominal em Porto Alegre (RS)**: Análise variacionista. Santa Cruz do Sul, RS: Edunisc, 2005.
- LABOV, W. **Padrões Sociolinguísticos** [1972]. Trad. Marcos Bagno; Maria Marta P. Scherre; Caroline Rodrigues Cardoso. São Paulo: Parábola, 2008.
- LUCCHESI, D. **Língua e sociedade partidas**: a polarização sociolinguística do Brasil. São Paulo: Contexto, 2015.
- LUCCHESI, D.; ARAÚJO, S. A Sociolinguística Variacionista: Fundamentos Teóricos e Metodológicos. In: **Projeto Vertentes**. Salvador: UFBA/Projeto Vertentes, 2006. Disponível em: <http://www.vertentes.ufba.br/socio.htm>. Acesso em: 12 jan. 2019.

- MARTELOTTA, M. E. **Mudança linguística**: uma abordagem baseada no uso. São Paulo: Cortez, 2011.
- MARTELOTTA; M. E.; AREAS, E. K. A visão funcionalista da linguagem no século XX. *In*: CUNHA, M. A. F.; OLIVEIRA, M. R.; MARTELOTTA, M. E. (Orgs.). **Linguística funcional**: teoria e prática. Rio de Janeiro: DP&A/ Faperj, 2003.
- MAY, G. H. Discutindo o papel do funcional no sociofuncionalismo. **Working papers in Linguística**, Florianópolis, v. 10, n. 2, p. 69-79, jul./dez. 2009.
- NARO, A. J. Modelos quantitativos e tratamento estatístico. *In*: MOLLICA, M. C.; BRAGA, M. L. (Orgs.). **Introdução à Sociolinguística Variacionista**. Rio de Janeiro: UFRJ, 1992.
- NARO, A. J.; VOTRE, S. J. Mecanismos funcionais do uso da língua função e forma. **D.E.L.T.A.**, São Paulo, v. 8, n. 2., p. 285-260, 1992.
- NEVES, M. H. M. Estudos funcionalistas no Brasil. **D.E.L.T.A.**, São Paulo, v. 15, n. esp., p. 71-104, 1999.
- OLIVEIRA, L. A. Formalismo e funcionalismo: fatias da mesma torta. **Sitientibus**, Feira de Santana, n.29, p.95-104, jul./dez. 2003.
- PIRES DE OLIVEIRA, R. Uma história de delimitações teóricas: trinta anos de semântica no Brasil. **Revista D.E.L.T.A.**, v. 15, n. esp., p. 291-321, 1999.
- SAUSSURE, F. **Curso de Linguística Geral**. 28. ed. 2. reimp. São Paulo: Cultrix, 2012.
- SEVERO, C. G. Entre a Sociolinguística e os Estudos Discursivos: O Problema da Avaliação. **Interdisciplinar**, São Paulo, Ano VI, v.14, p. 07-15, jul./dez. 2011.
- SILVA, C. R.; OLIVEIRA, M. J. A variação/mudança numa perspectiva sociofuncionalista. **Miguilim**: Revista Eletrônica do Netlli, v. 6, n. 2, p. 243-264, maio/ago. 2017.
- TARALLO, Fernando. **A pesquisa sociolinguística**. 9. ed. São Paulo: Ática, 2001.
- TAVARES, M. A. **A gramaticalização de E, AÍ, DAÍ e ENTÃO**: estratificação/variação e mudança no domínio funcional da sequenciação retroativo-propulsora de informações – um estudo sociofuncionalista. 2003. 302f. Tese (Doutorado em Linguística) – Universidade Federal de Santa Catarina, Florianópolis.
- TAVARES, M. A. Sociofuncionalismo: um duplo olhar sobre a variação e a mudança linguística. **Interdisciplinar**: Revista de Estudos em Língua e Literatura, v. 17, p. 27-47, 2013.
- TAVARES, M. A.; GÖRSKI, E. M. Variação e sociofuncionalismo. *In*: MARTINS, M. A.; ABRAÇADO, J. (Orgs.). **Mapeamento sociolinguístico do português brasileiro**. São Paulo: Contexto, 2015. p. 249-270.
- WEINREICH, U.; LABOV, W.; HERZOG, M. I. **Fundamentos empíricos para uma teoria da mudança linguística**. Trad.: Marcos Bagno. Revisão técnica: Carlos A. Faraco. São Paulo: Parábola, 2006 [1968].

ABORDAGEM SOCIOFUNCIONAL DE VOCÊ E CÊ NA FALA POPULAR DE SALVADOR - BA

ROSANE BISPO DOS SANTOS
CRISTINA DOS SANTOS CARVALHO

1 INTRODUÇÃO

Nos séculos XX e XXI, distintos têm sido os enfoques dados à descrição das formas *você* e *cê* no português brasileiro: (i) a origem de *você* como pronome de tratamento; e a sua reanálise (com a inserção no quadro pronominal referenciando a 2ª pessoa); (ii) o processo de variação entre *tu*, *você*, *ocê* e *cê*; (iii) a gramaticalização (um processo de mudança categorial através do qual um item lexical adquire um *status* gramatical, ou um item já gramatical passa a exercer uma função mais gramatical (HOPPER; TRAUGOTT, 2003 [1993]), podendo ainda adquirir, nos termos de Martelotta (2011), funções interacionais) e as outras mudanças — fonológicas, morfossintáticas e semântico-pragmáticas — envolvidas nesse processo; (iv) a trajetória de gramaticalização de *você*, que culminou na forma reduzida *cê*, considerada por diversos linguistas (VITRAL, 1996; BAGNO, 2012; CASTILHO, 2019) um pronome que está passando, no português brasileiro, por um processo de cliticização.

Nessa direção, o pronome *você* e a variante *cê* têm sido abordados, em perspectivas sincrônica, diacrônica ou pancrônica, em diversos estudos de diferentes regiões brasileiras, ora à luz do paradigma formal (VITRAL, 1996; VITRAL; RAMOS, 1999, 2006; VITRAL; VIEIRA; OLIVEIRA, 2010, entre outros)¹, ora sob o prisma dos paradigmas funcional e/ou sociolinguístico (PERES, 2006; SOUSA, 2008; GONÇALVES, 2010; LOREGIAN-PENKAL; MENON, 2012; ROCHA, 2017; GUIMARÃES *et al.*, 2018; ROCHA; SOUSA, 2020; SILVA; VITÓRIO, 2020 etc.).

¹ Vale ressaltar que, apesar de as análises apresentadas neste trabalho serem pautadas em pressupostos sociofuncionalistas, alguns textos fundamentados na abordagem formal da gramaticalização (VITRAL, 1996; VITRAL; RAMOS, 1999, 2006; VITRAL; VIEIRA; OLIVEIRA, 2010) serão considerados nesta pesquisa pois: (i) observam o processo de gramaticalização das formas variantes *você* e *cê*; (ii) auxiliam as análises sobre as motivações linguísticas do uso da forma reduzida *cê* no PB; (iii) apresentam diferentes olhares sobre a proposta de cliticização da variante *cê*. Desse modo, esses textos contribuíram significativamente para algumas das reflexões empreendidas neste trabalho.

Neste capítulo, apresenta-se um recorte da pesquisa de mestrado de Santos (2021) sobre o emprego de *você* e *cê* na fala popular de Salvador (BA). Assumindo o viés da variação/ gramaticalização, o presente texto tem por objetivo analisar, na variedade soteropolitana, os usos dessas duas formas de expressão da segunda pessoa do singular na função sintática de sujeito, em posição pré-verbal.

A descrição de *você/cê*, do ponto de vista teórico, orienta-se por postulados do sociofuncionalismo (TAVARES, 2003, 2013; GÖRSKI; TAVARES, 2013, 2017 etc.), que constitui uma interlocução teórico-metodológica entre a corrente norte-americana do funcionalismo (HOPPER, 1991; HOPPER; TRAUGOTT, 2003 [1993] etc.) e a vertente variacionista da sociolinguística (WEINREICH; LABOV; HERZOG, 2008 [1968]; LABOV, 2008 [1972]). Do ponto de vista metodológico, realiza-se uma abordagem qualiquantitativa dos dados, partindo da análise de ocorrências empíricas extraídas de um *corpus* constituído de dezesseis entrevistas sociolinguísticas integrantes do banco de dados do Programa de Estudos do Português Popular de Salvador – PEPP (LOPES; SOUZA; SOUZA, 2009).

Para contemplar o objeto linguístico em tela, inicialmente explicitam-se alguns pressupostos do sociofuncionalismo. Logo após, destacam-se aspectos metodológicos da pesquisa. Em seguida, procede-se a uma análise qualiquantitativa dos usos de *você* e *cê*, citando os resultados referentes às variáveis linguísticas consideradas relevantes pelo programa GoldVarb X (SANKOFF; TAGLIAMONTE; SMITH, 2005) para o fenômeno variável investigado no trabalho de Santos (2021). Na sequência, faz-se uma breve discussão sobre redução fonética, variação e gramaticalização de *você* e *cê* na fala popular soteropolitana. Por fim, expõem-se as considerações finais e as referências do trabalho.

2 O SOCIOFUNCIONALISMO: ENTRECruzando VARIAÇÃO E GRAMATICALIZAÇÃO

A abordagem do sociofuncionalismo (TAVARES, 2003, 2013; GÖRSKI; TAVARES, 2013, 2017, entre outros) consiste no diálogo entre dois modelos teórico-metodológicos: a sociolinguística variacionista (WEINREICH; LABOV; HERZOG, 2008 [1968]; LABOV, 2008 [1972]), que investiga, entre outras questões, formas linguísticas alternativas e seus condicionamentos estruturais e sociais; e o funcionalismo de vertente norte-americana (HOPPER, 1987, 1991; HEINE, 2003; HOPPER; TRAUGOTT, 2003 [1993]), que assume, entre outros pressupostos basilares, a relação entre gramática e discurso, admitindo, que a gramática emerge do uso da língua.

Na proposta conciliadora que ampara a linha sociofuncionalista, são analisadas as motivações para o uso e/ou a emergência de determinadas formas/construções em um mesmo ou em distintos domínios funcionais:

Numa perspectiva sociofuncionalista, os resultados quantitativos e qualitativos obtidos são explicados através de princípios e motivações de natureza cognitivo-comunicativa – cuja fonte principal é o funcionalismo norte-americano ou linguística baseada no uso –, além de princí-

pios e motivações de natureza sociocultural e estilística – cuja fonte principal é a sociolinguística variacionista (TAVARES, 2013, p. 38).

As vertentes envolvidas na interlocução teórico-metodológica apresentam compatibilidades no que concerne a determinados postulados, a saber: (i) primazia da língua em uso; (ii) reconhecimento do caráter inerente da variação; (iii) concepção da mudança como processo contínuo, gradual; (iv) combinação das perspectivas sincrônicas e diacrônicas a partir de um enfoque pancrônico; (v) relevância da frequência de uso; (vi) reconhecimento da atuação do processo de gramaticalização nas línguas; (vii) relação entre a variação morfossintática e o processo de mudança por gramaticalização; (viii) quantificação estatística como evidência para atestar fenômenos de variação e mudança (TAVARES, 2003, 2013; GÖRSKI; TAVARES, 2013, 2017). São esses postulados que, ao permitirem a aproximação entre a sociolinguística variacionista e o funcionalismo norte-americano, têm sido tomados, na literatura da área, como pressupostos sociofuncionalistas.

Embora haja as convergências supracitadas, existem particularidades que diferenciam as vertentes em relação à delimitação do objeto de estudo e a algumas assunções teóricas.² A título de ilustração, pode-se citar o que é considerado, nesses dois modelos teóricos, como ponto de partida no que tange à relação entre os processos de variação e mudança linguísticas. A esse respeito, Rocha (2017) explica:

[...] na Sociolinguística Variacionista, percebemos, primeiramente, que há um estágio de variação e que, posteriormente, caso uma forma suplante a(s) outra(s), dá-se a mudança linguística; por outro lado, ao ancorarmo-nos no Funcionalismo Norte-Americano, a leitura é feita por outra ótica, afinal, há, a priori, uma mudança linguística licenciada pela gramaticalização e, a posteriori, a percepção de um estágio de variação, visto que formas inovadoras, que são frutos da gramaticalização, passam, também, a codificar dados valores que já eram codificados por formas mais conservadoras, estabelecendo-se, assim, a variação linguística. Diante disso, cabe-nos, em poucas palavras, afirmar que, se por um lado, o gatilho, na Sociolinguística, é disparado a partir da variação linguística; por outro lado, no Funcionalismo, ele se dá na mudança (ROCHA, 2017, p. 89).

Adverte-se, então, que a interface mencionada anteriormente não se dá de forma plena, podendo pender mais para uma ou outra abordagem teórica e, por conseguinte, podendo haver distintos enfoques sociofuncionalistas. É o que Tavares (2003) chama de sociofuncionalismo mais variacionista ou sociofuncionalismo mais funcionalista. Assume-se, aqui, o segundo tipo de enfoque. Como já mencionado, este trabalho pretende descrever os usos de *você* e *cê* sob a interface entre variação e mudança por gramaticalização. Nos termos de Görski e Tavares (2017),

[...] essa perspectiva de interface [...] (tomada de um ponto de vista funcionalista) envolve contribuições mútuas: o tratamento da variação linguística pode ser aprimorado com subsídios

² Para mais informações sobre convergências e especificidades entre a sociolinguística variacionista e o funcionalismo norte-americano, conferir o texto “Sociofuncionalismo: ainda uma conversa na diferença?” presente neste livro.

vindos de estudos sobre a gramaticalização, assim como a análise do processo de gramaticalização pode ser enriquecida com informações provenientes de análises variacionistas (GÖRSKI; TAVARES, 2017, p. 35).

Sobre a questão citada, Tavares (2003, p. 146) afirma que a interface variação-gramaticalização “[...] envolve um movimento em direção ao funcionalismo, entendendo-se a gramática como emergente”, conceito postulado por Hopper (1987). Concebe-se, então, a ideia de que a gramática de uma língua existe não por ter um sistema pronto, definido e disponível (em si) para os usos dos seus falantes. Nesse caso, reconhece-se que a interação social — isto é, a comunicação entre os usuários e a materialização da língua — organiza e reorienta a gramática dos seus falantes (FURTADO DA CUNHA, 2013), o que motiva variação e mudança nas línguas humanas.

Nesse contexto, a gramaticalização, termo cunhado por Antoine Meillet para designar “atribuição de caráter gramatical a uma palavra antes autônoma” (MEILLET, 1912, p. 131)³, é um fenômeno que demonstra que a gramática de uma língua não é estática, ou seja, tende a se moldar de acordo com as necessidades comunicativas do falante, segundo os diferentes contextos de uso. Como fenômeno, representa, então, um processo gradual de mudança categorial e, sendo assim, é entendida como um tipo de mudança linguística:

A partir dessa perspectiva, a gramaticalização é geralmente vista como um subconjunto de mudanças linguísticas pelas quais um item ou construção lexical em certos usos assume características gramaticais, ou através das quais um item gramatical se torna mais gramatical (HOPPER; TRAUOGOTT, 2003[1993]), p. 2 tradução nossa).⁴

Segundo o processo de gramaticalização, para explicar as regularidades dos fenômenos que implicam mudança linguística, é imprescindível perceber que os elementos envolvidos no processo deixam de atuar no nível representacional e passam a desenvolver novas funções no nível interpessoal (MARTELOTTA, 2011). Configuram-se como elementos representacionais os itens que fazem referência a elementos desse mundo biossocial como objetos, ações, sentimentos etc., representados pelas classes dos substantivos, verbos e adjetivos. “Funcionam no nível interpessoal os chamados elementos processuais, que não têm exatamente aquela atribuição de caráter referencial, apresentando funções tradicionalmente definidas como gramaticais, funcionais ou interacionais [...]” (MARTELOTTA, 2011, p. 93). Nesse nível, inserem-se, entre outros elementos, os conectivos, verbos auxiliares, os clíticos, marcadores discursivos e modalizadores.

A título de ilustração, pode-se citar a gramaticalização de *você* em *cê* a partir do *cline* de mudança: item gramatical (*você*) > clítico (*cê*). Nesse caso, tem-se uma mudança a partir de um item já gramatical que se torna ainda mais gramatical. Como se verá

³ “L’attribution du caractère grammatical a un mot jadis autonome” (MEILLET, 1912, p. 131).

⁴ “From this perspective, grammaticalization is usually thought of as that subset of linguistic changes whereby a lexical item or construction in certain uses takes on grammatical characteristics, or through which a grammatical item becomes more grammatical” (HOPPER; TRAUOGOTT, 2003[1993], p. 02).

na discussão a ser apresentada na seção 5 deste trabalho, defende-se que a forma *cê* tem passado a se comportar como um clítico, assumindo uma posição mais fixa e regular (posição de sujeito pré-verbal) no português brasileiro e, mais especificamente, na variedade soteropolitana.

Uma observação interessante é que o contexto de sujeito pré-verbal está atrelado não só à gramaticalização, mas também à variação entre *você* e *cê*. Note-se ainda que tanto a gramaticalização como a variação concernentes a essas duas formas de expressão da segunda pessoa do singular são permeadas pela redução fonética: nesse sentido, há alternância entre formas plena e reduzida (*você* ~ *cê*) e gramaticalização de uma forma plena, que sofre erosão fonética, tornando-se uma forma reduzida (*você* > *cê*). Neste estudo, a adoção de postulados sociofuncionalistas se mostra necessária na análise do fenômeno em estudo a fim de verificar os processos de variação e gramaticalização, bem como suas imbricações.

Na próxima seção, expõem-se alguns aspectos metodológicos da pesquisa relativos ao *corpus*, aos procedimentos e às variáveis linguísticas controladas.

3 METODOLOGIA

Este estudo, por centrar sua análise na interface entre os processos de variação e mudança via gramaticalização de *você/cê*, norteia-se metodologicamente por pressupostos da sociolinguística variacionista (LABOV, 2008[1972]) e do sociofuncionalismo (TAVARES, 2003; GÖRSKI *et al.*, 2003 etc.).

Para a investigação dos usos de *você* e *cê* na fala popular soteropolitana, foram examinadas ocorrências empíricas retiradas de dezesseis das quarenta e oito entrevistas sociolinguísticas do acervo do Programa de Estudos do Português Popular de Salvador – PEPP (LOPES; SOUZA; SOUZA, 2009), gravadas no período de 1998 a 2000. Os informantes do banco de dados se encontram estratificados de acordo com as seguintes categorias sociais: (i) sexo (homem/ mulher); (ii) faixa etária (F1 – 15 a 24 anos; F2 – 25 a 35 anos; F3 – 45 a 55 anos; F4 – a partir de 65 anos); (iii) nível de escolaridade (fundamental e médio)

Na pesquisa, os procedimentos metodológicos seguidos para observação e análise dos dados linguísticos, apresentados na fala espontânea soteropolitana foram: (i) levantamento e fichamento dos dados; (ii) audição de dezesseis inquéritos do PEPP; (iii) definição das variáveis independentes da pesquisa; (iv) codificação dos dados; (v) submissão dos dados ao programa de análise estatística GoldVarb X (SANKOFF; TAGLIAMONTE; SMITH, 2005); (vi) análise qualiquantitativa dos dados.

No percurso do levantamento dos dados, observou-se que as formas *você* e *cê*, na fala soteropolitana, têm se comportado de maneira distinta, em especial na função de sujeito. Desse modo, houve a necessidade de estabelecer etapas para a organização e análise dos dados. A primeira fase, referente ao levantamento e fichamento dos dados, possibilitou uma observação inicial dos usos de *você/cê* a partir da apreciação de

dezesesseis inquiridos do PEPP. Inicialmente, foram feitos o levantamento e o fichamento de *você/cê* nas posições pré-verbal e pós-verbal. No entanto, constatou-se que não havia ocorrência de *cê* na posição pós-verbal e, por conseguinte, não havia variação entre *você* e *cê*. Foram excluídas desta análise: (i) ocorrências em que *você* e *cê* estavam em contextos interrompidos e sem a presença de verbo; (ii) dados de *você* em posição pós-verbal. Ademais, houve pouquíssimas ocorrências de *você* nessa posição (1).

(1) Com sua mãe.

[.] Não teve, não teve cas ... outro castigo que eu me lembrasse assim não. Só mais que eu me lembre foi esse mesmo, que eu estava pintando demais, aí minha avó “CÊ daí para casa de sua mãe”, aí eu fui para casa da minha mãe, mas não aguentei e fugi. Ela: “*está vendo VOCÊ*, da próxima vez já sabe”. Mas CÊ sabe que naquele tempo, isto normal, não, não teve demais, não teve nada de mais não... (PEPP Inf. 20, homem, 24 anos, médio, p. 113).

Depois dessa fase, foi realizada a audição dos inquiridos com o intuito de verificar se houve realização da forma plena *você* ou da reduzida *cê*. Na sequência, uma nova triagem foi realizada considerando a função sintática desempenhada pelas variantes, obtendo-se como objeto final de estudo as variantes *você* (2a) e *cê* (2b) na função de sujeito, em posição pré-verbal (variável dependente da pesquisa) por se tratar da função sintática em que a variação e a gramaticalização podem ser verificadas mais amplamente.

(2) a. DOC: São outras leituras, não é?

01: É. Hoje ... A televisão é o que mais eles interessam. Não sei se **VOCÊ** já viu ..., pelo meu neto, eu tiro. Minha neta, ela acorda cedo, pra ir pra ... pro curso, né? Tem que estar lá sete horas (PEPP Inf. 01, mulher, 68 anos, fundamental, p. 330).

b. DOC: Sim, mas quanto tempo você passou lá?

24: Ah, eu passei, **CÊ fala**, o passeio foi de fim de semana, feriadão, essas excursões que fazem aí né, pra Aracaju, Porto Seguro né (PEPP Inf. 24, homem, 51 anos médio, p. 263).

Sendo assim, a variável dependente deste estudo é binária, comportando as variantes *você* e *cê*, que competem entre si: primeiro, com mesmo valor de verdade — já que são formas estratificadas de segunda pessoa do singular, que podem ser empregadas tanto voltadas para o interlocutor como de forma genérica; segundo, pelo fato de exercerem a função de sujeito, em posição pré-verbal. Desse modo, parte-se da suposição de que os usos de *você* e *cê*, conforme o contexto, sejam condicionados pelos fatores de natureza estrutural e social.

Para a análise de *você* e *cê*, foram controlados os seguintes grupos de fatores⁵ linguísticos: natureza semântico-funcional; paralelismo formal; tipo semântico do verbo; o número de sílabas da forma verbal; tipos de sequência linguística. Para a definição de algumas dessas variáveis, foram consideradas pesquisas anteriores, que verificaram a variação de *você* e *cê* em diferentes regiões brasileiras desenvolvidas (SOUSA,

⁵ A pesquisa de Santos (2021) também considerou variáveis extralinguísticas — sexo, nível de escolaridade e faixa etária —, definidas a partir da estratificação social dos informantes do PEPP.

2008; LOREGIAN-PENKAL, 2012; LOREGIAN-PENKAL; MENON, 2012; ROCHA, 2017; SOUSA; ROCHA, 2020, entre outras).

Na próxima seção, apresenta-se, em viés qualiquantitativo, a análise dos usos de *você* e *cê* na fala popular soteropolitana, tendo em vista alguns de seus condicionamentos linguísticos.

4 USOS DE VOCÊ/CÊ NA FALA POPULAR SOTEROPOLITANA: ANÁLISE DOS DADOS

Na amostra examinada por Santos (2021), registrou-se um total de 395 ocorrências de *você/cê* desempenhando a função sintática de sujeito em posição pré-verbal. Dessas 395 ocorrências registradas, 247 (62.5%) foram de *você* e 148 (37.5%), de *cê*. Percebe-se, então, que a forma reduzida *cê* (variante inovadora) ainda ocorre com um percentual menor do que a forma plena *você*.

Como já mencionado, o item *cê* tem motivado discussões a respeito dos processos de variação e mudança por gramaticalização. Um dos princípios de gramaticalização utilizado como parâmetro explanatório dessa relação é o da estratificação (HOPPER, 1991). Sobre esse ele, o autor estabelece que, em um domínio funcional, novas camadas estão sempre emergindo; nesse caso, as formas antigas não são necessariamente descartadas, podendo continuar a coexistir com as novas. Fazendo um paralelo o princípio da estratificação, pode-se entender que *você* e *cê* são formas estratificadas que ocupam o mesmo domínio funcional. Note-se que, com o surgimento de *cê* (camada nova), *você* (camada antiga) não desapareceu.

Vitral e Ramos (2006), em seus estudos, igualmente indicaram que a forma *cê* (inovadora) já estaria exibindo um comportamento distinto da forma *você*. No caso, as explicações apresentadas pelos autores enfatizaram a tendência de a forma reduzida *cê* (em posição pré-verbal) estar apresentando indícios de uma *cliticização*, ou seja, uma mudança categorial operada via gramaticalização. Rocha (2017) também revelou que, no tocante à função sintática de sujeito, a forma reduzida *cê* se comportava de maneira fixa, sendo concorrente da forma plena *você*, como sujeito pré-verbal.

Os resultados obtidos por Santos (2021) se coadunam com os achados de Rocha (2017) e Loregian-Penkall e Menon (2012). Em Rocha (2017), por exemplo, observou-se que, igualmente aos falantes soteropolitanos, os conquistenses utilizavam mais *você* (56%) do que *cê* (44%). O estudo desenvolvido por Loregian-Penkall e Menon (2012) também registrou que os falantes curitibanos preferiam usar mais o pronome *você* (70%) do que a variante *cê* (30%). Inversamente a esses resultados, a pesquisa de Peres (2006) analisou o emprego de *você/cê* pelos falantes de Belo Horizonte. Nas localidades mapeadas pela autora, *cê* apresentava uma frequência de 75.1%, e a forma plena *você*, de 22.4%.

Estabelece-se uma relação entre a distribuição da frequência de *você* e *cê* e o princípio funcionalista da marcação, que, nos termos de Givón (1995), distingue as categorias em *marcadas* e *não marcadas* com base em três critérios: (i) complexidade estrutural: a estrutura marcada tende a ser mais complexa estruturalmente do que a

não marcada; (ii) distribuição de frequência: a estrutura marcada costuma ser menos frequente que a não marcada; (iii) complexidade cognitiva: a estrutura marcada tende a ser cognitivamente mais complexa do que as não marcadas. Considerando-se essa relação, assume-se, com base nas ocorrências gerais, que *cê*, variante inovadora, em função da menor frequência de uso em relação ao pronome *você*, é a forma marcada na fala popular soteropolitana.

Quanto às variáveis linguísticas controladas na pesquisa, três foram selecionadas, em ordem de relevância, pelo programa GoldVarb X como estatisticamente significativas para explicação da variação entre *você* e *cê* na fala popular soteropolitana: natureza semântico-funcional, paralelismo formal e tipo semântico do verbo.

A seguir, exibem-se os resultados referentes a esses grupos de fatores. Vale ressaltar que, na análise estatística, a variante tomada como valor de aplicação foi *cê*, a forma inovadora e reduzida.

4.1 NATUREZA SEMÂNTICO-FUNCIONAL

Nos dados de fala soteropolitana examinados, a variável *natureza semântico-funcional* (valor determinado e indeterminado) foi a primeira selecionada relevante pelo GoldVarb X para mostrar o condicionamento para realização de *cê*. Quanto a essa variável, na amostra, a forma *cê* apresenta os resultados expostos na Tabela 1.

Tabela 1: Condicionamento da variável natureza semântico-funcional na realização de *CÊ* na fala soteropolitana

NATUREZA SEMÂNTICO-FUNCIONAL	Total de ocorrências	%	Peso relativo
Determinado	85/ 233	36,5	.554
Indeterminado	63/162	38.9	.423
TOTAL	148/395	37.5	

Fonte: Santos (2021, p. 100). Significância: 0.048.

Os dados da Tabela 1 mostram que, na fala soteropolitana, a forma de segunda pessoa do singular *cê* é favorecida no contexto semântico-funcional de determinação do sujeito, com o peso relativo correspondente a 0.554. Em outras palavras, esse resultado revela que o item *cê*, nessa variedade do português, é mais condicionado quando o uso se volta para um interlocutor específico (3). Nesse caso, o fato de o valor determinado ser o contexto favorecedor de *cê* parece se justificar por algumas ocorrências de *cê*, nos dados analisados, em marcadores discursivos⁶ voltados para o interlocutor, por exemplo, *cê vê* (3b) e *cê sabe* (3c).

⁶ À esteira de Penhavel (2013, p. 1296), considera-se, aqui, marcadores discursivos como “[...] mecanismos que atuam no nível do discurso (aqui entendido como organização textual-interativa), estabelecendo algum tipo de relação entre unidades textuais e/ou entre os interlocutores”.

(3) a. DOC: Esperta, viu?

19: A gente, a gente era assim, ... (inint) ...o prato de farinha, ela aí botava mais, tia, **CÊ** bota um pouco de caldinho aqui? Aí ela botava mais uma concha de caldo a gente botava mais farinha que era pra poder encher a barriga porque era uma fome danada. Então a gente ficou assim. [...] (PEPP Inf. 19, mulher, 29 anos, fundamental, p. 174).

b. DOC: Será que a cobrança também a nível de direção, coordenação não é maior entre as escolas particulares?

20: Talvez, mas também se ... se eles também cobrassem no colégio estadual, ia se nivelar, porque **CÊ vê**, passou hoje na televisão que o professor agora vai ter que fazer curso universitário para ser professor mesmo, não vai ser mais magistério direto, por causa do ensino, dizem que os alunos não estão aprendendo (PEPP Inf. 20, homem, 24 anos, médio, p. 122).

c. DOC: Ah, sim, seria mais fácil.

01: É, lá ... eh ... seria mais fácil, por causa do ordenado, casa, **CÊ sabe**, interior a vida é um pouco mais barata do que assim na capital, né. Aí nós fomos, eu passei lá seis anos (PEPP Inf. 01, mulher, 69 anos, fundamental, p. 316).

O exemplo (3a) revela o emprego da forma inovadora *cê* com valor determinado, já que o informante a utiliza para se dirigir a uma interlocutora específica, no caso, a própria tia, ao pedir mais um pouco de caldinho da comida. Em (3b), o falante usa *cê vê* como mecanismo para chamar a atenção do documentador sobre a informação exibida na televisão a respeito da exigência de nível superior aos professores. Em (3c), o informante comenta sobre seu ordenado e introduz *cê sabe* para justificar, nesse caso, sua mudança para o interior, alegando ser um local cujo custo de vida é menor que na capital. Nota-se, então, que, nos excertos (3b) e (3c), as construções *cê vê* e *cê sabe* operam como marcadores discursivos, desempenhando funções interacionais.

Ainda com relação aos resultados obtidos, considera-se que, embora, entre os falantes soteropolitanos, o valor determinado seja o elemento condicionador da forma inovadora *cê*, é possível notar que tal item também vem sendo empregado com valor indeterminado, com percentual de 38,9 %, como mostra (4a). No entanto, o peso relativo de 0.423 para *cê* revela que o contexto de indeterminação desfavorece a forma reduzida e, conseqüentemente, se mostra favorável à variante *você* (4b).

(4) a. DOC: Muito grande não é não.

14: É, é grande, em todos, agora recentemente eu fui aí na torre daqui, e o pessoal está todo mundo desgostoso porque não vale a pena **CÊ** ganhar esse dinheiro, nós ganhamos como sargento, nós não ganhamos como controlador de voo, tem uma gratificação pequena, mas ganhamos como sargento e não controlador de voo (PEPP Inf. 14, homem, 73 anos, médio, p. 342).

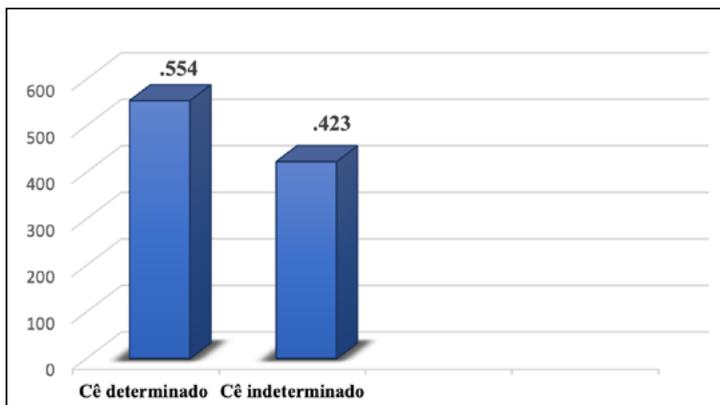
b. DOC: Tênis né?

24: Tênis né. E no meu tempo não existia isso, **VOCÊ** brincava na rua tranqüilo, casa aberta, entrava, entrava e saía na casa do vizinho, o quintal, praia, eu já cansei de sair da Liberdade pra tomar banho na praia do, da Ribeira, na Boa Viagem. Hoje eu não posso fazer isso, hoje ninguém, ninguém, aí oh, hoje mesmo olhe aí oh, né? (PEPP Inf. 24, homem, 51 anos, médio, p. 253).

No caso de (4), verifica-se que *cê* e *você* são utilizados de forma indeterminada (genérica): em (4a), o informante usa essa estratégia de indeterminação ao falar para o documentador sobre o desgosto dos colegas de trabalho com relação ao dinheiro ganho no exercício da profissão. Em (4b), o informante usa *você* de forma indeterminada para falar que, antigamente, as pessoas podiam brincar na rua, com tranquilidade, o que já não pode ser visto hoje.

No que concerne a essa variável, aventou-se a hipótese de que, dentre as formas estratificadas *você* e *cê*, a reduzida *cê* esteja em um estágio de gramaticalização que obedece a um *continuum* de mudança: mais concreto > menos concreto, com um uso menos referencial da segunda pessoa e mais genérico (indeterminado). Pode-se dizer que, na fala popular de Salvador, tal hipótese não se confirmou. No entanto, embora o resultado aqui obtido seja inesperado, é preciso chamar a atenção para o fato de que o peso relativo do valor indeterminado não se encontra tão distante do ponto neutro, como mostra o gráfico 1.

Gráfico 1: Condicionamento da variável natureza semântico-funcional na realização de *CÊ* na fala soteropolitana



Fonte: Santos (2021, p. 102).

Quanto aos resultados referentes à natureza semântico-funcional exibidos nesta pesquisa, verifica-se que as análises aqui empreendidas se mostram semelhantes aos achados de Peres (2006), que examinou o mesmo fenômeno na fala de Belo Horizonte. Nesse trabalho, *cê*, na referência definida, obteve o peso relativo de 0.64, atestando que os falantes da capital mineira utilizam mais a forma reduzida voltada ao seu interlocutor.

4.2 PARALELISMO FORMAL

A variável *paralelismo formal* foi a segunda selecionada no GoldVarb X como relevante para o condicionamento de *cê* na fala popular soteropolitana. Na tabela 2, são exibidos os resultados referentes a essa variável no condicionamento do uso de *cê*.

Tabela 2: Condicionamento da variável paralelismo formal na realização de *CÊ* na fala soteropolitana

PARALELISMO FORMAL	Ocor./total	%	Peso Relativo
Ocorrência isolada	53/113	43.9	.586
Primeiro de uma série	27/93	29.0	.419
Forma precedida de forma igual	48/150	32.0	.440
Forma precedida de forma diferente	20/39	51.3	.669
TOTAL	148/395	37.5	

Fonte: Santos (2021, p. 103). Significância: 0.048.

Como se vê, a forma *cê* obteve o seu uso favorecido em dois contextos: (i) quando sua ocorrência estava em uma série precedida de forma diferente (no caso, de *ocê*), com o peso relativo de 0.669; (ii) quando era uma ocorrência isolada, com o peso relativo de 0,586. Tais contextos são exemplificados, respectivamente, em (5) e (6).

(5) DOC: É a de 10 anos?

36: A de 7. [...] Aí ela fala assim: “eu não vou pegar mais bicicleta, ó quando eu estiver grande”, aí a de dezesseis, desprevenida, pegou a bicicleta, aí se bateu no muro, está aprendendo agora, aí caiu, arrancou o tampão aqui do ombro, ficou em carne viva, aí disse: “estava vendo (inint.), eu falei para você para VOCÊ não brincar de bicicleta que **CÊ** tinha almoçado nagora.” Ela fez: “ah, mainha, mas para aprender tem que cair”, eu disse “mas não é assim, mas quando a mãe fala vocês tem ouvir, que eu tenho experiência,” aí Sheila fez assim: “Está vendo, é por isso que eu não pego mais, mainha disse que não é para eu pegar eu não pegar eu não pego.” Tem uma que chama Milena lá a mãe dela não dá carinho a ela não, só faz pegar ela e bater, “Milena, oh, Milena” (PEPP Inf. 36, mulher, 45 anos, fundamental, p. 247).

(6) DOC: Inventava alguma coisa, ou você dizia não é mentira?

44: Não, ele não inventava não, porque aí quando ele falava que eu batia nele eu falava porque que eu batia, aí minha irmã ela não achava errado, porque ela mesmo falava para mim **CÊ** tá criando ele, ele tem que te obedecer (PEPP Inf. 44, mulher, 16 anos fundamental, p. 108).

Por outro lado, há pouco favorecimento quando *cê* é precedido da mesma forma (7a), com peso relativo de 0.440, e quando é o primeiro de uma série (8a), com peso relativo (0,419), o que significa dizer que, na fala popular soteropolitana, os contextos “forma antecedida pela mesma forma” (7b) e “ocorrência primeira em uma série” (8b) motivam o emprego de *ocê*.

(7) a. DOC: É, é bom.

34: Aqui ela veio, estava trabalhando, ganhando cento, acho que era cento e oitenta no salão de beleza aí, teve que tirar outra carteira, a mulher disse “não vou, com essa carteira aí **CÊ** não trabalha, **CÊ** tira outra, pra eu poder assinar porque eu não vou baixar o seu salário nessa carteira que está aí”, então ela teve que fazer isso (PEPP Inf. 34, homem, 66 anos, fundamental, p. 314).

b. DOC: É, você está, conseguiu ficar no formol esse período.

34: Não, acho, acho que é, quando eu tive doente que o médico disse assim “VOCÊ não pode se preocupar muito, **VOCÊ** pode morrer a qualquer hora”, eu tenho, eu sou cardíaco, tenho

problema no coração, eu digo “é mesmo doutor?”, ele disse “é, é”, “mas o que eu vou fazer, eu vou aliviar mais a minha vida” (PEPP Inf. 34, homem, 66 anos, fundamental, p. 307).

(8) a. DOC: Aí você fez o que não, que não fazia antes...

23: Ninguém mais me segurou, mas fora disso não. É dezoito anos, dezenove, com a igreja, a influência da igreja também, é uma das coisas que eu tenho assim maior gratidão ao próprio Deus é que com tudo isso minha mãe, uma mulher batista né, evangélica, então ela passou toda essa coisa de respeitar Deus. Olha R..., a vida tá aí, tá, tem isso, isso, isso, mas **CÊ** tem um Deus aí que **VOCÊ** vai ter que respeitar. Então, eu com doze anos fui uma das primeiras né, a me batizar nas águas, coisas dos, dos tradicionais, dos batistas, eu levei aquilo sério mesmo, então eu li a Bíblia, aquela coisa, então aquilo me influenciou bastante, qualquer passo errado que eu ia dar eu pensava assim, ih! (PEPP Inf. 23, mulher, 34 anos, médio, p. 208).

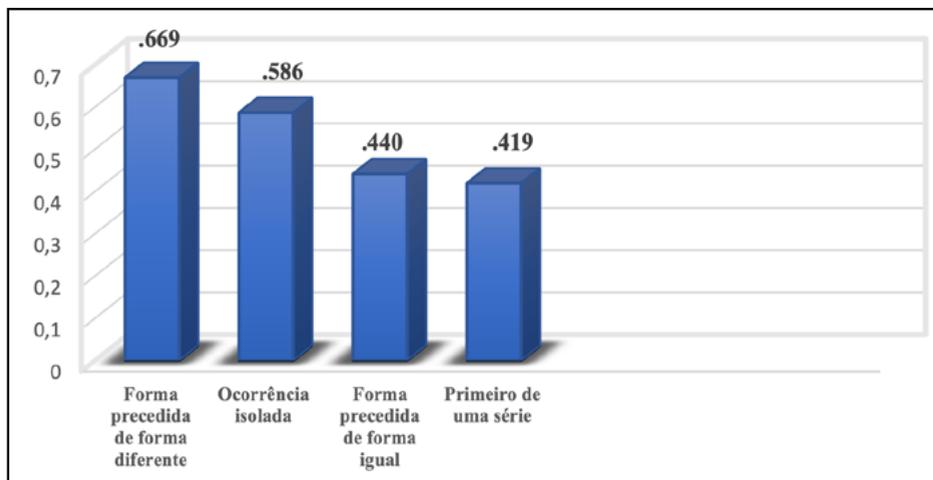
b. DOC: O sexo.

20: O sexo mesmo, então **VOCÊ** não pode, de poder **CÊ** pode, agora depois **CÊ** vai parar para conversar com seus filhos, porque são cenas, muitas cenas pesadas. E isso invoca até na escola que nego fala “pô, **CÊ** viu aquilo que aconteceu na televisão? Isso e aquilo” aí fala, aí fala até a professora “mas menino, numa idade dessa **CÊ** já está sabendo disso e mais aquilo” (PEPP Inf. 20, homem, 24 anos, médio, p. 124).

No que se refere aos resultados sobre essa variável, observa-se que o falante soteropolitano segue tendências distintas no uso de *cê* e *você*: utiliza *cê* quando essa forma é precedida de *você* (5) e em contextos isolados (6); emprega *você* quando essa forma é antecedida por outro *você* (7b) ou é a primeira ocorrência de uma série (8b).

O gráfico 2 mostra os resultados obtidos para essa variável.

Gráfico 2: Condicionamento da variável paralelismo formal na realização de *CÊ* na fala soteropolitana



Fonte: Santos (2021, p. 105).

A hipótese aventada para esse grupo de fatores é a de que, na fala soteropolitana, há uma tendência de o falante manter as mesmas formas ao iniciar uma série com *cê*.

Verifica-se que, para essa variável, a hipótese não se confirma, pois, conforme já mencionado, apenas dois contextos favorecem o uso da forma *cê*: as ocorrências isoladas e quando precedido de *você*. Isso indica que variantes linguísticas, como é o caso de *cê*, podem não ser motivadas pela repetição de formas ou marcas, como prevê o paralelismo formal. Considera-se que a perda de massa fônica seja um fator determinante para essa ausência de paralelismo formal, o que serviria como mais uma evidência da cliticização de *cê*.

A variável paralelismo formal também foi utilizada como parâmetro em outras pesquisas sobre os usos de *você* e *cê* (LOREGIAN-PENKAL, 2012; ROCHA, 2017; ROCHA; SOUSA, 2020, entre outras) e tem sido produtiva para a identificação de contextos que influenciam esses usos. A título de ilustração, Rocha (2017) verificou, na fala conquistense, os seguintes resultados para paralelismo formal: houve favorecimento de *você* quando precedido de uma ocorrência de outro *você* (com peso relativo de 0.63) e quando iniciava uma série (com o peso relativo de 0.58). Nos casos em que *cê* foi observado, houve referência a *cê* em ocorrências isoladas (com o peso relativo de 0.59) e quando *cê* estava precedido de outro *cê* (com peso relativo de 0.73). Esses resultados se mostram parcialmente semelhantes aos encontrados nesta pesquisa. A única diferença foi que, nos dados de fala soteropolitana, houve um maior favorecimento de *cê* antecedido de *você*.

4.3 TIPO SEMÂNTICO DO VERBO

Nas ocorrências soteropolitanas examinadas, a variável *tipo semântico do verbo* foi a terceira selecionada relevante pelo GoldVarb X. Os resultados referentes a essa variável estão expostos na tabela 3.

Tabela 3: Condicionamento da variável tipo semântico do verbo na realização de *CÊ* na fala soteropolitana

TIPO SEMÂNTICO DO VERBO	Total de ocorrências	%	Peso Relativo
Epistêmico/cognitivo	15/38	39.5	.486
<i>Dicendi</i>	1/2	50.0	.749
Perceptivo	38/70	54.3	.733
De ação	74/175	36.6	.492
De movimento	12/42	30.2	.384
Volitivo	5/15	33.3	.340
De estado	8/34	23.5	.330
De posse	3/12	25.0	.293
TOTAL	147/389 ⁷	37,7	

Fonte: Santos (2021, p. 106). Significância: 0.048.

⁷ Santos (2021) ressalta que, nessa tabela, o total geral e o total de ocorrências de *cê* são diferentes das outras tabelas porque, para essa variável, não foram considerados casos de verbos em expressões cristalizadas, tais como “dar bolo”.

Os dados permitem identificar os tipos semânticos verbais que favorecem a realização da forma reduzida *cê* na fala popular de Salvador: são os verbos *dicendi/declarativos* (9) e *perceptivos* (10), com pesos relativos de 0.749 e 0.733, respectivamente. É preciso registrar, todavia, que só houve dois dados na amostra com verbos *dicendi*: um com *cê* e outro com *você*. Sendo assim, os resultados obtidos para esse fator não podem ser considerados confiáveis. Em relação aos verbos perceptivos, o verbo *ver* é o mais frequente na amostra e, nesse caso, ocorre mais na forma *vê*, formando um marcador discursivo (10): das 38 ocorrências de *cê* com verbo perceptivo, 25 são com *vê*.

(9) DOC: L ... é quem?

19: L ... é minha irmã, a caçula, a menor, com dezesseis anos... Foi pra casa ontem. “**CÊ** não disse que ia ficar com L ... lá em casa pra mamãe trabalhar na casa da moça?, ele “Ah mas eu quero ir com a senhora ...”, eu digo “Vá com L ... que daqui a pouco eu ligo pra falar com você.” Que, quando o telefone toca, ele voa em cima do telefone, porque eu não tenho, né? Aí ... (PEPP Inf. 19, mulher, 29 anos, fundamental, p. 176- 177).

(10) DOC: E os professores? Como eram os professores naquela época?

20 Os professores eram ... eram bons professores. Eles passavam as melhores coisas possíveis para a gente. **CÊ vê** que a gente estava falando assim eh... questionando, **CÊ vê** ... a educação daquele... daquele tempo dos professores que passaram para a gente eu acho que é a mesma idéia que todo mundo que estudou junto, a mesma concepção (PEPP Inf. 20, homem, 24 anos, médio, p. 117).

Os verbos de ação e os cognitivos têm comportamentos muito próximos no que diz respeito ao emprego de *cê*: os primeiros (11a) com percentual de 39.5% e peso relativo de 0.492; os últimos (12a) com percentual de 36.6% e peso relativo de 0.486. Sendo assim, pode-se dizer que, quando se trata desses dois tipos semânticos de verbos, os soteropolitanos oscilam entre *cê* e *você*, havendo um ligeiro favorecimento da forma plena (11b) e (12b).

(11) a. DOC: Então você procura pelo menos assim perna, né...

19: Perna ou braço, entendeu, que se **CÊ bate** na cabeça dá um problema, bate nas costas dá um postema aí, amanhã ou depois é que eu vou sofrer com ele, né ... mas meu menino (inint) apanhando não...(inint). Então L ... me falou que ele obedeceu, entendeu, apesar de tudo ele obedeceu ontem, foi, “Como foi, L ...?” “Ficou quietinho, cheguei aqui dei banho, liguei a televisão, ficou quieto, meio-dia dei o almoço dele... ficou quieto ...” (PEPP Inf. 19, mulher, 29 anos, fundamental, p. 177).

b. DOC: Tênis né?

24: Tênis né. E no meu tempo não existia isso, **VOCÊ brincava** na rua tranqüilo, casa aberta, entrava, entrava e saía na casa do vizinho, o quintal, praia, eu já cansei de sair da Liberdade pra tomar banho na praia do, da Ribeira, na Boa Viagem. Hoje eu não posso fazer isso, hoje ninguém, ninguém, aí oh, hoje mesmo olhe aí oh, né? (PEPP Inf. 24, homem, 51 anos, médio, p. 253).

(12) a. DOC: Com sua mãe.

20: [...] Não teve, não teve cas ... outro castigo que eu me lembrasse assim não. Só mais que eu me lembre foi esse mesmo, que eu estava pintando demais, aí minha avó “cê daí para casa de sua mãe”, aí eu fui para casa da minha mãe, mas não aguentei e fugi. Ela: “está vendo você, da próxima vez já sabe”. Mas **CÊ sabe** que naquele tempo, isto normal, não, não teve demais, não teve nada de mais não... (PEPP Inf. 20, homem, 24 anos, médio, p. 113).

b. DOC: Você tem um filho soldado é?

34: Tenho. **VOCÊ sabe**, eh, chamava B..., ele sempre ia na casa de J... (PEPP Inf. 34, homem, 66 anos, fundamental, p. 310)

Com relação aos verbos que menos favorecem o uso da forma inovadora *cê*, registraram-se os *de movimento*, *volitivo*, *de estado* e *de posse*, com os pesos relativos mais baixos. Pode-se inferir, então, que esses tipos semânticos verbais — *de movimento* (13), *volitivo* (14), *de estado* (15) e *de posse* (16) — são os contextos favoráveis aos usos de *você* nos dados examinados, como revelam os exemplos a seguir:

(13) DOC: É isso mesmo

34 [...] mas não era pra empregar, mas apenas pra receber um dinheiro, e aí com isso ela ficou em São Paulo, ela está lá batendo a cabeça de novo, mas eu digo: “menina, **VOCÊ volta** de novo”, ela comprou, essa, até essa casa que ela tem em Castelo Branco que a minha filha está morando é dela, que quando ela veio com o dinheiro da indenização, ela comprou essa casa, [...] (PEPP Inf. 34, homem, 66 anos, fundamental, p. 314).

(14) DOC: (inint) Pode?

19: Estava vendo que esse lenço não estava lá entendeu, e ela queria. Ela disse: “Olhe - me xingou, né - se **VOCÊ** não achar esse lenço eu vou lhe quebrar toda, que o lenço está aí e **VOCÊ quer** procurar”. Aí eu fui procurar e não achei. Ela pegou, me deu um bocado de porrada ... aí me deu uma dentada... (PEPP Inf. 19, mulher, 29 anos, fundamental, p. 166).

(15) DOC: E a escola você começou desde cedo?

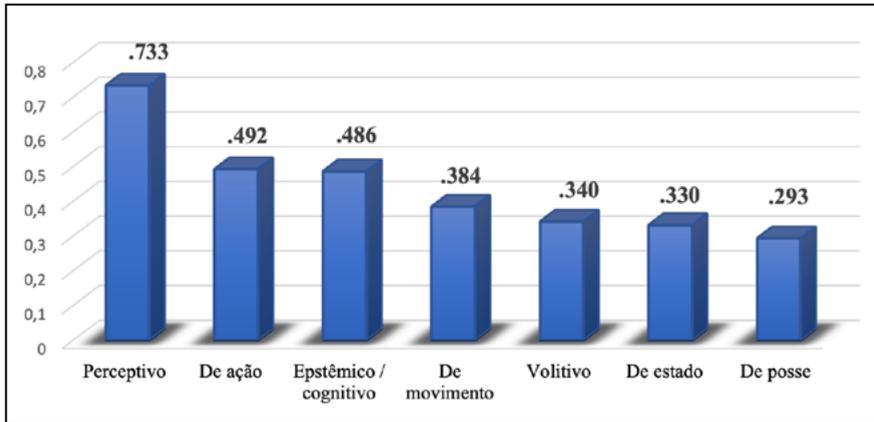
17: Estudava, comecei cedo, comecei a estudar cedo, mas tinha que tomar conta dos irmãos porque todas as empregadas que botavam em casa eu brigava achava que não fazia as coisas certo, aquela, aquela menina mandona, autoritária porque era mais velha, aquele, aquela responsabilidade de meu pai dizer, “**VOCÊ é** a mais velha tome conta de seus irmãos”, eu me achava totalmente dona da casa, minha mãe saía seis horas da manhã chegava onze da noite... (PEPP Inf. 17, mulher 53 anos, médio, p. 275).

(16) DOC: Positivo.

17: É, porque ela disse pra mim **VOCÊ tem** condição de ter três empregos ou quatro, eu disse, “opa! Vou partir pra lá” (PEPP Inf. 17, mulher, 53 anos, médio, p. 280).

No que concerne à distribuição dos dados de *cê* conforme o parâmetro tipo semântico de verbo, o gráfico 3 permite melhor visualização dos resultados supramencionados. Nesse gráfico, devido à baixa frequência de dados dos verbos *dicendi*/declarativos, não serão considerados os pesos relativos obtidos para *cê* diante deles.

Gráfico 3: Condicionamento da variável tipo semântico do verbo na realização de *CÊ* na fala soteropolitana



Fonte: Santos (2021, p. 109).

A hipótese aventada para essa variável foi a de que verbos perceptivos (entre eles *ver*, mais especificamente, em sua forma reduzida *vê*) e cognitivos (entre eles, *saber*) são os que mais favoreceriam o uso do item *cê*. Nesse caso, com base nos resultados obtidos, verifica-se que essa hipótese foi confirmada apenas em relação aos verbos perceptivos. É importante destacar que, já nas etapas de levantamento e fichamento dos dados, começou-se a perceber um uso mais acentuado de *cê* com o verbo perceptivo *ver* na construção *cê vê* como marcador discursivo, o que suscitou a formulação da hipótese no tocante a esse tipo semântico verbal.

Infelizmente, para a realização de uma análise comparativa com outros dados, são raros os estudos que consideram a variável tipo semântico verbal como parâmetro para registrar os casos de variação e mudança das formas *você* e *cê*. No entanto, Guimarães (2014) utilizou esse grupo de fatores para mapear os distintos usos de *tu* / *você* + *cê* na fala popular cearense. O autor obteve os verbos *epistêmicos*, *dicendi* e de *ação* como favorecedores do emprego de *tu*.

4.4 SINTETIZANDO OS RESULTADOS

Como se pôde perceber com base nos resultados obtidos para as variáveis linguísticas selecionadas pelo GoldVarb X, houve, nos dados de fala popular soteropolitana, fatores favoráveis à realização de *cê* e *você*. No Quadro 1, sintetizam-se os resultados referentes a essas duas variantes.

O quadro exhibe um panorama geral dos contextos que, na fala popular soteropolitana, favorecem os usos da forma plena *você* e da reduzida *cê* na função de sujeito pré-verbal. Há condicionamentos linguísticos para uma maior realização de *cê* quando essa forma apresenta ocorrência isolada ou é precedida de *você* e ocorre com verbos *dicendi* e perceptivos. Quanto à forma plena *você*, os contextos que mais a condicionam são os seguintes: estruturalmente, ser primeira de uma série ou estar

precedida da mesma forma; e ocorrer com verbos volitivos, de movimento, de estado e de posse.

Quadro 1: Variáveis e fatores condicionantes de *cê* e *você*

VARIÁVEIS SELECIONADAS	FATORES CONDICIONANTES DAS VARIANTES	
	CÊ	VOCÊ
Natureza semântico-funcional	Valor determinado	Valor indeterminado
Paralelismo formal	Contexto isolado e precedido de forma diferente	Primeiro de uma série e precedido da mesma forma.
Tipo semântico do verbo	Verbos perceptivos	Verbos epistêmicos, volitivos, de movimento, de estado e de posse

Fonte: Adaptado de Silva (2021, p. 115).

Pode-se dizer que, mesmo a variante mais conservadora *você* ainda se mostrando a preferida na fala de Salvador, a forma inovadora *cê* já exhibe características distintas quanto ao seu funcionamento na sentença. Por exemplo, tal comportamento pode estar sendo motivado principalmente por conta da perda de material fonético, o que estaria resultando no funcionamento de *cê* apenas como sujeito sentencial, na sua fixação na posição pré-verbal e na sua ocorrência em contextos isolados ou precedidos da forma *você*.

Na seção a seguir, discutem-se os processos de redução fonética, variação e gramaticalização associados às formas *você* e/ou *cê*, estabelecendo, sempre que possível, uma relação com os dados empíricos da fala popular soteropolitana analisados.

5 VOCÊ E CÊ: REDUÇÃO FONÉTICA, VARIAÇÃO E GRAMATICALIZAÇÃO

A redução da variante *você* à forma *cê* tem sido descrita por pesquisas de distintas abordagens da linguística em uma perspectiva mais formal ou nas vertentes variacionista ou sociofuncional. Retomam-se, aqui, brevemente, algumas discussões acerca do processo de redução fonética até a chegada da forma *cê*, mais reduzida, em funcionamento no português brasileiro.

Segundo Câmara Jr. (1979), o pronome *você* surgiu a partir da forma de tratamento *vossa mercê*, sendo que houve um estágio intermediário de variação *vosmecê*, variante abandonada na medida que o uso de *você* se cristalizou. Ainda sobre esse processo, Gonçalves (2008) explana que percursos, tais como o considerado por Nascentes (1956), Lapa (1991), entre outros (17) e o citado em Amaral (1955) (18), fazem parte das reflexões sobre a redução de *você* e *cê*.

(17) *vosmecê* > *vossemecê* > *vossecê* > *você* > *ocê* > *cê*

(18) *vossuncê* > *vassuncê* > *mecê* > *vancê* > *vacê* > *vosmincê*

Vitral, Vieira e Oliveira (2010) esclarecem que, em um recorte diacrônico, a forma *vossa mercê* seja a fonte do processo de gramaticalização que gerou o item gramatical *você*, que, por sua vez, em contínuo processo de mudança, gerou a forma inovadora *cê*, que se encontra em adiantado percurso de cliticização. Nos termos de Vitral, Vieira e Oliveira (2010), enfatiza-se que:

Ocorre que *você* e *cê*, na acepção de segunda pessoa do discurso, passam, portanto, a ter o mesmo valor de verdade, tornando-se como previsto, formas concorrentes. Nesta visão dos fatos, certos fenômenos de concorrência de formas são esperados na análise de inovações linguísticas prevista pela noção de gramaticalização o que abre a possibilidade de aplicação da metodologia variacionista, com as consequências teóricas que conhecemos (VITRAL; VIEIRA; OLIVEIRA, 2010).

Sousa (2008), ao explicar a trajetória de *você* na fala paraibana, destaca a redução fonética como um fator atinente ao processo de gramaticalização. Parte-se, aqui, desse entendimento para traçar comentários acerca da trajetória de *você* e *cê* na fala soteropolitana e, mais especificamente, sobre os indícios de cliticização de *cê*. A redução fonética é descrita como perda de massa fônica ocorrida nos vocábulos em detrimento dos usos, a exemplo *você* > *cê*, em que há a passagem de um vocábulo dissílabo para uma forma monossílabo.

Com a redução fonética, em um primeiro momento, passa a haver uma estratificação (HOPPER, 1991) entre *você* e *cê* empregados, no português brasileiro, no mesmo domínio funcional, ou seja, como referência à segunda pessoa do discurso na função de sujeito. A estratificação, princípio relacionado ao conceito de variação, estabelece, nos termos de Hopper (1991), que, em um domínio funcional, novas camadas estão sempre emergindo; no caso, as formas antigas não são necessariamente descartadas, podendo continuar a coexistir com as formas novas. Em um segundo momento, começa a haver uma divergência (HOPPER, 1991), mediante a atuação da gramaticalização, entre *você* (forma fonte) e *cê* (forma gramaticalizada), que passa a se comportar como clítico, na função de sujeito.

A análise, com base nos registros empíricos da fala soteropolitana, revelou que a função de sujeito pode ser um critério sintático adotado para a identificação e distinção de *você* e *cê*. Na amostra, não foram encontradas ocorrências de *cê* como complemento verbal, mas apenas de *você* (19), o que fez com que houvesse uma delimitação do fenômeno linguístico aqui investigado para emprego de *você/cê* na função de sujeito (20), (21).

(19) DOC: Mente?

20: É. desenvolve mais, aque... como é que se fala? Eh... ele vai se sentir mais livre, mais solto, fora de outros meios. Então o esporte ensina **VOCE** até a viver, a desenvolver na sua vida (PEPP Inf. 20, homem, 24 anos, médio, p. 127).

(20) DOC: Oh meu Deus...

17: Eu digo, “tá mainha, pode falar”, chego a me sentar, as vezes eu me deito, fico as mãos dormentes, mudo o telefone de lugar, e fico, mas não tem nada pra falar, nada, fala todas as besteiras que **CÊ pode imaginar** (PEPP Inf. 17, mulher 53 anos, médio, p. 290).

(21) DOC2 :Agora porque eles não queriam (...inint...)?

17: Não é que não queriam o namoro, eu não sei o que era, **VOCÊ acredita** que até hoje eu não sei o que era? V... diz que é porque a minha era de ouro e a das minhas irmãs era de bata (PEPP Inf. 17, mulher, 53 anos, médio, p. 287).

Em relação à posição do sujeito na sentença, constatou-se, nos dados examinados, que a forma reduzida *cê* se apresenta em uma posição apenas pré-verbal (20), diferentemente de *você*, cuja ocorrência igualmente se dá não só nessa posição (21) mas também na pós-verbal (22). No que concerne à descrição da reduzida *cê*, considera-se que a posição do sujeito seja mais uma evidência da sua cliticização. Nesse caso, pode-se estabelecer uma relação entre a posição pré-verbal de *cê* e a próclítica, típica posição dos clíticos no português brasileiro.

(22) Com sua mãe.

[..]Não teve, não teve cas ... outro castigo que eu me lembrasse assim não. Só mais que eu me lembre foi esse mesmo, que eu estava pintando demais, aí minha avó “CÊ daí para casa de sua mãe”, aí eu fui para casa da minha mãe, mas não aguentei e fugi. Ela: “*está vendo VOCÊ*, da próxima vez já sabe”. Mas CÊ sabe que naquele tempo, isto normal, não, não teve demais, não teve nada de mais não... (PEPP Inf. 20, homem, 24 anos, médio, p. 113).

Sousa (2008) e Rocha, e Santos e Sousa (2016), baseados em outras variedades do português brasileiro — paraibana e conquistense, respectivamente —, também mostraram que, no tocante à função sintática de sujeito, a forma reduzida *cê* se comporta de maneira fixa, anteposta ao verbo. Assim, o fato de a forma plena *você* ter perdido, ao longo do tempo, substância fonética, tem sido tomado como argumento para considerar *cê* (forma reduzida) como palavra fraca (átona), o que justificaria a sua exclusividade na posição pré-verbal.

Segundo um dos princípios de gramaticalização, o da decategorização (HOPPER, 1991), há, na forma gramaticalizada, perda ou neutralização de marcas morfossintáticas da forma fonte. Hopper e Traugott (2003[1993]) e Heine (2003) destacam que, nesse processo, quando ocorre uma cliticização ou afixação de um elemento, há perda da sua autonomia lexical, da sua independência lexical, o que parece ilustrar o que está acontecendo com *cê* funcionando como clítico na fala dos soteropolitanos e em outras variedades do português brasileiro.

A gramaticalização é vista como um processo unidirecional, mostrando uma tendência de seguir trajetórias de mudança: *lexical > gramatical, representacional > interpessoal, concreto > abstrato, menos gramatical > mais gramatical*, entre outras. Com a gramaticalização de *você* e *cê* na fala popular de Salvador, percebeu-se também um comportamento que sugere um processo de abstratização semântica das formas. Partindo-se dos distintos usos de *você*, forma fonte representativa da segunda pessoa do singular, verificou-se que esse item apresenta comportamento linguístico genérico, mais abstrato, caracterizado nesta pesquisa como valor indeterminado. Seguindo tal entendimento, observou-se que a forma reduzida *cê* igualmente se comporta, muitas

vezes, com sentido mais concreto e menos concreto (abstrato), sendo um comportamento herdado da sua forma fonte *você*. É interessante relembrar que, na amostra, como se viu na tabela 1 da subseção 4.1, as diferenças percentuais entre o valor determinado (36.5%) e o indeterminado (38.9%) de *cê* não foram muito altas. No entanto, levando-se em conta o peso relativo obtido, foi o valor determinado que mais favoreceu a forma reduzida *cê*, o que refutou a hipótese aventada na pesquisa.

Ainda sobre os usos de *cê* encontrados no *corpus*, pode-se dizer que a aquisição de *status* de mais gramatical do item reduzido *cê* parece estar motivando seus usos como marcador discursivo diante de determinadas formas verbais. Conforme o resultado quantitativo desta pesquisa, verificou-se que a variante *cê*, na função de sujeito pré-verbal, tende a ser favorecida por verbos perceptivos. Nesse caso, destaca-se o verbo *ver*, na forma *vê*: na amostra, das 38 ocorrências com verbo perceptivo, 25 foram com *vê*, o que permite dizer que, nessa classe semântica verbal, *vê* tem sido a forma eleita pelos soteropolitanos para ser a hospedeira de *cê*. Nos dados analisados, a ligação de *cê* ao hospedeiro verbal *vê* (perceptivo) sugere uma nova função, a de marcador discursivo, baseada na construção *cê vê* (23).

(23) DOC: E... e negativos assim? Como você rotula os negativos? Dá alguns exemplos?

20 Negativos são, **CÊ VÊ** começa a partir das novelas, né? **CÊ VÊ** muitas cenas, eu mesmo, eu na minha idade mesmo, eu fico olhando assim, eu estou assistindo a televisão com minha avó, eu vejo aquelas cenas assim aí eu fico assim, eu fico enver... eu fico envergonhado. “Mas rapaz, como é que pode?” Imagine sua filha assistindo, sua filha ou filho. Tudo bem, tem uns pais que aceitam, né? [...] (PEPP Inf. 20, homem, 24 anos, médio, p. 125).

O exemplo (23) mostra que o informante usa a construção *cê vê* para exemplificar, para o seu interlocutor (no caso, o documentador), o que está sendo dito sobre aspectos negativos da televisão. Nesse caso, *cê vê* é empregado pelos falantes com a função exemplificativa, atrelada à macrofunção de chamada de atenção do ouvinte (ROST SNICHELOTTO, 2009; CARVALHO; GOMES, 2017).

Com base na discussão aqui empreendida e nos resultados obtidos na análise quantitativa dos dados, é possível alocar os usos de *você* e *cê* registrados no *corpus* em três possíveis trajetórias de gramaticalização.

- (i) Os itens perdem a liberdade e passam a ser mais fixos e regulares de forma unidirecional, como ilustra, por exemplo, o *cline* de mudança categorial: item gramatical (*você*) > clítico (*cê*).
- (ii) A referência genérica ou indeterminada, motivada a partir de uma referência determinada, sinaliza uma tendência à abstratização semântica na gramaticalização de *você/cê*, que pode ser evidenciada no *cline* de mudança: + concreto > - concreto.
- (iii) O comportamento da forma *clítica* diante de determinadas formas verbais com mais um percurso de gramaticalização ilustrado pelo *cline* de mudança: nível representacional > nível interpessoal. Nesse caso, como já mencionado, é toda a construção — *cê vê* — que se gramaticaliza. Essa construção passa a funcionar

como marcador discursivo, desempenhando uma função interpessoal, e deixa de atuar no nível representacional, não mais fazendo referência a elementos do mundo bio-psíquico-social.

A Figura 1 sintetiza melhor as considerações feitas sobre as trajetórias citadas.

Figura 1: Trajetórias de gramaticalização de *você* a *cê*

Item gramatical.....	Clítico
Mais concreto	Menos concreto
Clítico + verbo perceptivo.....	Marcador discursivo

Fonte: Santos (2021, p. 124).

Exibe-se, no primeiro *continuum*, no extremo à esquerda, o uso de *você* como item gramatical e, no extremo à direita, apenas da forma *cê* como clítico, entendendo-se que *cê* (clítico) tem sua origem a partir do *você* (item gramatical). No segundo percurso, ocorrem tanto *você* e *cê* com valor mais concreto (determinado, voltado a um interlocutor) e valor menos concreto (indeterminado, genérico) mostrando seu processo de abstratização. No terceiro *continuum*, registra-se a gramaticalização de *cê* em uma construção verbal: no extremo à direita, está o resultado dessa gramaticalização, um marcador discursivo.

Neste sentido, no que concerne às formas *você* e *cê*, as análises revelam a migração de valores gramaticais para valores mais gramaticais, perpassam pelo processo de abstratização e, por conseguinte, adentram a esfera interpessoal por via da gramaticalização. Discussões mais recentes como as de Rocha (2017) e Rocha e Sousa (2020) verificaram, entre os falantes conquistenses, a redução de *cê*, o processo de cliticização e os deslizamentos semânticos em detrimento dos usos menos concretos do item. Este trabalho, além dos parâmetros mencionados por pesquisas que a antecedem, registrou o comportamento da variante reduzida *cê* diante de um verbo específico como uma construção que funciona como marcador discursivo, além de evidências de mais uma via de gramaticalização do item *cê* no português brasileiro.

6 CONSIDERAÇÕES FINAIS

Este trabalho investigou, à luz do enfoque sociofuncionalista, os usos de *você* e *cê* na função sintática de sujeito, em posição pré-verbal, na fala popular soteropolitana. A pesquisa atestou que *você* (forma conservadora), quando comparado a *cê* (forma inovadora), apresenta um elevado índice de preferência entre os soteropolitanos. Ao correlacionar esse resultado a um dos critérios de distinção entre categorias marcadas e não marcadas, a frequência de uso, constatou-se que *você* constitui a forma não marcada e *cê*, a marcada na variedade em questão.

Os resultados da análise estatística mostraram que três grupos de fatores linguísticos — natureza semântico-funcional, paralelismo formal e tipos semânticos dos verbos — condicionam a variação entre *você* e *cê* na fala popular de Salvador. Sobre a variável natureza semântico-funcional, a análise revelou que a forma *cê* é favorecida quando se apresenta com valor determinado. Esse condicionamento pode ser explicado pelo fato de haver, na amostra, alguns usos de *cê* em marcadores discursivos voltados para o interlocutor como, por exemplo, *cê vê* e *cê sabe*.

No que concerne ao paralelismo formal, os resultados apontaram elevado uso de *cê* quando ocorre isoladamente e é precedido de forma diferente (no caso, de *você*), o que sinaliza que não há atuação do paralelismo formal na escolha de *cê*. Em relação a *você*, essa forma foi favorecida quando ocorre: como primeiro de uma série e precedido de forma igual, atestando a presença de paralelismo conforme também já visto em outros estudos (LOREGIAN-PENKAL, 2012; PERES, 2006; ROCHA; SOUSA, 2020, entre outros).

Quanto à variável tipo semântico do verbo, observou-se que os verbos perceptivos são os que favorecem o uso da forma reduzida *cê* e os volitivos, de movimento, de estado e de posse condicionam o uso de *você*. Como mencionado, verificou-se que *cê* e o verbo perceptivo *ver*, na construção *cê vê*, se comportam como um marcador discursivo, com a função de chamar a atenção do interlocutor para alguma porção do texto. Esse caso ilustra uma mudança da função representacional dessa construção para a função interpessoal.

Quanto à intersecção variação-gramaticalização, constatou-se que, nos dados de fala popular soteropolitana examinados, há uma estratificação entre a forma plena *você* e a forma inovadora *cê* é exercendo a função de sujeito sentencial, na posição pré-verbal. Ademais, a análise empreendida revelou que o item *cê*, via gramaticalização, tem adquirido um comportamento como clítico, ocorrendo apenas na posição pré-verbal de sujeito. Quanto ao processo de gramaticalização, assume-se, um *continuum* de mudança categorial: item gramatical (*você*) > clítico (*cê*). Além dele, considerou-se um *cline* de mudança, conforme os valores semântico-funcionais de *você* e *cê*: mais concreto (*determinado*) > menos concreto (*indeterminado*). Sobre esse *cline*, pode-se dizer que *você* e *cê* apresentam uma tendência à abstratização semântica ou, nos termos de Heine (2003), à dessemantização.

O presente estudo permitiu evidenciar, a partir de dados empíricos, as motivações linguísticas para os empregos de *você* e *cê* entre os falantes soteropolitanos. Nesse sentido, entender os contextos estruturais que condicionam os usos de *você* e *cê* tende a colaborar para a descrição dos processos de variação e mudança pelos quais esses itens passam, ressaltando os impactos no quadro pronominal e as novas funções assumidas pelos itens. Espera-se que a descrição aqui apresentada para os usos de *você* e *cê* possa contribuir para a realização de outras pesquisas sobre esse mesmo fenômeno em outras variedades do português brasileiro.

REFERÊNCIAS

- AMARAL, A. O **Dialeto Caipira**. Gramática Vocabulário. São Paulo: Anhembi, 1955.
- ANDRADE, A. L. S. de. **A variação você, cê e ocê no português brasileiro falado**. 2004. 136f. 2004. Dissertação (Mestrado em Linguística) – Universidade de Brasília, Brasília.
- BAGNO, M. **Gramática pedagógica do português brasileiro**. São Paulo: Parábola, 2012.
- CÂMARA JR., J. M. **História da linguística**. Petrópolis: Vozes, 1979.
- CARVALHO, C. dos S.; GOMES, J. C. C. *Olha, olhe e oh*: gramaticalização do verbo olhar na fala popular soteropolitana. **Estudos linguísticos e literários**, Salvador, n. 57, p. 297-318, 2017.
- CASTILHO, A. T. de. **Nova Gramática do Português Brasileiro**. São Paulo: Contexto, 2019.
- FREITAG, R. M. Estratégias gramaticalizadas de interação na fala e na escrita: marcadores discursivos revisitados. **ReVEL**, v. 7, n. 13, p. 01-05, 2009.
- FURTADO DA CUNHA, A. Funcionalismo. In: MARTELOTTA, M. E. (Org.). **Manual de linguística**. São Paulo: Contexto, 2013. p. 157-176.
- GIVÓN, T. **Functionalism and grammar**. Amsterdam: John Benjamins, 1995.
- GONÇALVES, C. R. **Uma abordagem sociolinguística dos usos das formas você, ocê e cê no português**. 2008. 349f. Tese (Doutorado em Linguística) – Universidade de São Paulo, São Paulo.
- GONÇALVES, C. R. De *Vossa Mercê a Cê*: caminhos, percursos. **Cadernos do CNLF**, vol. XIV, n. 4, p. 2535-2550, 2010.
- GÖRSKI, E. M.; TAVARES, M. A. O objeto de estudos na interface variação-gramaticalização. In: BAGNO, M.; CASSEB-GALVÃO, V. C.; REZENDE, T. F. (Orgs.). **Dinâmicas funcionais da mudança linguística**. São Paulo: Parábola, 2017. p. 35-63.
- GÖRSKI, E. M.; TAVARES, M. A. Reflexões teórico-metodológicas a respeito de uma interface *sociofuncionalista*. **Revista do GELNE, Natal**, v. 15, p. 79-101, 2013.
- GÖRSKI, E. M. *et al.* Fenômenos discursivos: resultados de análises variacionistas como indícios de gramaticalização. In: RONCARATI, C.; ABRAÇADO, J. (Org.). **Português brasileiro: contato linguístico, heterogeneidade e história**. Rio de Janeiro: 7Letras: FAPERJ, 2003. p. 106-122.
- GUIMARÃES, T. A. A. S. **Tu é doido, macho!** A variação das formas de tratamento no falar de Fortaleza. 2014. 237f. Dissertação (Mestrado em Linguística Aplicada) – Universidade Estadual do Ceará, Fortaleza.
- GUIMARÃES, T. A. A. S.; ARAÚJO, A. A.; PEREIRA, M. L. de S. Os pronomes *você(s)* e *cê(s)* no falar de Fortaleza-CE. **Letras escreve**, Macapá, v. 8, p. 75 -10, 2018.
- HOPPER, P. J. Emergent grammar. **Berkeley Linguistics Society**, v. 13, p. 139-57, 1987.
- HOPPER, P. J. On some principles of grammaticalization. In: TRAUGOTT E. C.; HEINE, B. (Eds.). **Approaches to grammaticalization**. v. 1. Philadelphia: John Benjamins, 1991. p. 17-35.
- HOPPER, P. J.; TRAUGOTT, E. C. **Grammaticalization**. Cambridge University Press, 2003 [1993].
- LABOV, W. **Padrões sociolinguísticos**. São Paulo: Parábola, 2008 [1972].
- LAPA, M. R. **Estilística da língua portuguesa**. 3. ed. São Paulo: Fontes, 1991.
- LOPES, N. da S.; SOUZA, C. M. B. de; SOUZA, E. H. P. M. de. (Orgs.). **Um estudo da fala popular de Salvador**: PEPP. Salvador: Quarteto, 2009.
- LOREGIAN-PENKAL, L. **(Re)análise da referência de segunda pessoa na fala da região Sul**. 2004. 260f. Tese (Doutorado em Letras) – Universidade Federal do Paraná, Curitiba.
- LOREGIAN-PENKAL, L.; MENON, O. P. da S. *Você, ocê, cê* em Curitiba, Paraná. **SIGNUM: Estudos da Linguagem**, Londrina, n. 15/1, p. 223-243, 2012.
- MARTELOTTA, M. E. **Mudança linguística: uma abordagem baseada no uso**. São Paulo: Cortez, 2011.
- MEILLET, A. L'évolution des formes grammaticales. *Scientia*, 12, v.6, n. 26, 1912. Reimpr. In: MEILLET, A. **Linguistique historique et linguistique générale**, 1. Paris: Champion, 1948. p. 130-148.
- NASCENTES, A. **O tratamento de você no Brasil**. Curitiba: Companhia das Letras, 1956.
- PERES, E. P. De "vossa mercê" a "cê": os processos de uma mudança em curso. **Revista (Con)Textos Linguísticos**, Vitória, v. 1 n. 1, p. 155-168, 2007.

- PERES, E. P. **O uso de “você”, “ocê” e “cê” em Belo Horizonte**: um estudo em tempo aparente e em tempo real. 2006. 235f. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte.
- ROCHA, W. J. C. **Você e Cê**: um estudo sociofuncional em uma comunidade do Sertão da Ressaca. 2017. 199f. Dissertação (Mestrado em Linguística) – Universidade Estadual do Sudoeste da Bahia, Vitória da Conquista.
- ROCHA, W. J. C.; SANTOS, L. O. dos; SOUSA, V. V. O pronome você e sua variante cê: Um Estudo (socio) funcional. **Interdisciplinar**, Itabaiana, Ano XI, v. 24, p. 143-158, 2016.
- ROST-SNICHELOTTO, C. A. **“Olha” e “vê”**: caminhos que se entrecruzam. 2009. 408 f. Tese (Doutorado em linguística) – Universidade Federal de Santa Catarina, Florianópolis.
- RUMEU, M. C. de B. Vestígios da pronominalização de Vossa Mercê > Você em missivas cariocas e mineiras: uma incursão pelo português brasileiro escrito nos séculos XIX e XX. **Veredas online – Atemática**, Juiz de Fora, 2012.
- SANKOFF, D.; TAGLIAMONTE, S. A.; SMITH, E. **Goldvarb X**: a variable rule application for Macintosh and Windows. Department of Linguistics, University of Toronto, 2005. Disponível em: <http://individual.utoronto.ca/tagliamonte/goldvarb.html>. Acesso em: 13 jul. 2021.
- SANTOS, R. B. dos. **Usos de você/cê na fala popular soteropolitana**: uma análise sociofuncionalista. 2021.136f. Dissertação (Mestrado em Estudo de Linguagens) – Universidade do Estado da Bahia, Salvador.
- SILVA, S. de O. P.; VITÓRIO, E. G. de S. L. A. A variação *você* e *cê* no sertão alagoano. **Leitura**, Maceió, v. 2, n. 59, p. 122-142, 2017.
- SOUSA, V. V. **Os (des)caminhos do você**: uma análise sobre a variação e mudança na forma, na função e na referência do pronome *você*. 2008. 223f. Tese (Doutorado em Letras) – Universidade Federal da Paraíba, João Pessoa.
- TAVARES, M. A. **A gramaticalização de e, aí, daí e então**: estratificação/variação e mudança no domínio funcional da sequenciação retroativo-propulsora de informações um estudo sociofuncionalista. 2003. 307f. Tese (Doutorado em Linguística) – Universidade Federal de Santa Catarina, Florianópolis.
- TAVARES, M. A. Funcionalismo e gramáticas do português brasileiro. In: SOUZA, E. R. de (Org.). **Funcionalismo linguístico**: análise e descrição. São Paulo: Contexto, 2012. p. 33- 52.
- TAVARES, M. A. Sociofuncionalismo: um duplo olhar sobre a variação e a mudança linguística. **Interdisciplinar**, Itabaiana, ed. esp. ABRALIN/SE, Ano VIII, v. 17, p. 27-48, jan./jun. 2013.
- VITRAL, L. A forma *cê* e a noção de gramaticalização. **Estudos da Linguagem**, Belo Horizonte, v.4, n. 1, p. 115-124, 1996.
- VITRAL, L.; RAMOS, J. Gramaticalização de “você”: um processo de perda de informação semântica. **Filologia e Linguística Portuguesa**, São Paulo, n. 3, p. 55-63, 1999.
- VITRAL, L.; RAMOS, J. **Gramaticalização**: uma abordagem formal. Belo Horizonte: Faculdade de Letras FALE/ UFMG, 2006.
- VITRAL, L. T.; VIEGAS, M. do C.; OLIVEIRA, A. J. de. Inovação *versus* mudança: a intersecção gramaticalização/teoria da variação e mudança. In: VITRAL, L. T.; COELHO, S. (Orgs.). **Estudos de processos de gramaticalização em português**: metodologias e aplicações. Campinas: Mercado de Letras, 2010. p. 201-228.
- WEINREICH, U.; LABOV, W.; HERZOG, M. I. **Fundamentos empíricos para uma teoria da mudança linguística**. São Paulo: Parábola, 2006 [1968].

PESQUISAS EM SOCIOLINGUÍSTICA HISTÓRICA NA BAHIA: ALGUNS PERCURSOS E POSSIBILIDADES

HUDA DA SILVA SANTIAGO

ROSANA CARVALHO BRITO

1 INTRODUÇÃO

Situada na interface entre a sociolinguística e a linguística histórica, a sociolinguística histórica (SH) (ROMAINE, 1982; CONDE SILVESTRE, 2007) tem despontado como um viés para a aproximação aos fenômenos linguísticos de sincronias pretéritas, através de trabalhos que contribuem para uma melhor descrição da constituição histórica do português brasileiro, ao tempo que também são estudos que evidenciam os desafios relacionados a essa tarefa, instituindo um perfil próprio a esse campo¹. Hernández-Campoy e Schilling (2012) pontuam que, habitualmente, se questionam a representatividade e a validade dos estudos nessa área. De fato, como, em sua tarefa, a SH depende diretamente de textos escritos do passado, os dados que analisa podem resultar incompletos ou fragmentários, da mesma forma que é difícil reconstituir o contexto sociocultural do passado das línguas.

Na Bahia, muitas vezes sob o rótulo de *linguística sócio-histórica*, principalmente pela influência de Mattos e Silva (1998, 2004), estudos na perspectiva de uma sociolinguística histórica têm enfrentado desafios, por exemplo, de localização e constituição de *corpora* escritos representativos, bem como de identificação do perfil dos escreventes e de estabelecimento dos fatores sociais relacionados à comunidade da qual faziam parte.

Em sua apresentação no I Seminário do Projeto para a História do Português Brasileiro, em 1997 — *Ideias para a história do português brasileiro: fragmentos para uma composição posterior* —, texto publicado pela primeira vez em 1998, Mattos e Silva entretê, para recuperar a história do português brasileiro, a necessidade de reconstruir uma história social linguística do Brasil: “[...] uma sociolinguística histórica (ou só-

¹ Agradecemos ao amigo Pedro Daniel dos Santos Souza, pelos longos diálogos acerca das questões discutidas neste texto e pelo convite para participarmos da mesa-redonda com o mesmo tema, no XI Encontro de Sociolinguística: a Sociolinguística no Nordeste, em dezembro de 2021.

cio-história linguística) e uma história linguística, ou seja, a história das mudanças linguísticas que fizeram e fazem o português brasileiro apresentar as características que tem, o seu perfil próprio, a sua gramática” (MATTOS E SILVA, 1998, p. 39).

O objetivo, aqui, é apresentar alguns desses enfrentamentos no campo, principalmente no que se refere ao tratamento metodológico dos *corpora* utilizados, através de pesquisas como as desenvolvidas no âmbito do *Programa para a História da Língua Portuguesa* (PROHPOR/UFBA)² e do *Corpus Eletrônico de Documentos Históricos do Sertão* (CE-DOHS/UEFS)³.

Perseguindo esse fim, apresenta-se, na próxima seção, uma breve revisão bibliográfica sobre a pesquisa em SH e os desafios enfrentados por essa disciplina. Logo após, comentam-se alguns estudos desenvolvidos por pesquisadores baianos, que seguem o viés da SH, embora não assumam explicitamente, em seus textos, a filiação a esse modelo.

2 SOCIOLINGÜÍSTICA HISTÓRICA: ASPECTOS METODOLÓGICOS E DESAFIOS

A SH se constitui como um campo híbrido, entre a linguística histórica e a sociolinguística, no intento de aplicar, a dados do passado das línguas, o arcabouço metodológico fundado por este último modelo. A publicação de *Socio-Historical Linguistics. Its Status and Methodology* (1982), de Suzanne Romaine, é pioneira no estabelecimento dos métodos e práticas da SH, a que a autora denominou linguística sócio-histórica. Consoante Romaine, uma vez que a língua escrita, tal qual a oral, está sujeita a processos de variação e mudança estruturados, os métodos e técnicas utilizados na análise da fala podem ser aplicados ao estudo da variação em textos escritos do passado. Nos termos de Conde-Silvestre (2007, p. 33), a proposta de Romaine é

[...] la imbricación de ambas disciplinas [Linguística Histórica e Sociolinguística] en una *lingüística socio-histórica* que investigue y explique las diversas manifestaciones de la variación lingüística en una comunidad a lo largo del tiempo y también el modo en que esa variación – sus distintos usos, funciones y tipos – se desarrolla en lenguas y comunidades específicas, en el seno de ciertos grupos y redes sociales o en individuos concretos [...]. Se trata, por consiguiente, de una importante aplicación de los objetivos de una sociolingüística que era histórica en tanto que se aplicaba al estudio de los fundamentos de los cambios lingüísticos – tal como había sido implícitamente concebida por Weinreich, Labov y Herzog – hacia la reconstrucción efectiva de las relaciones de covariación entre determinadas variables lingüísticas y ciertos factores sociales en estadios pasados de evolución de los idiomas (CODE-SILVESTRE, 2007, p. 33).⁴

² Coordenado pela Professora Doutora Juliana Soledade Barbosa Coelho. Cf. <https://www.prohpor.org/>. Acesso em: 31/05/2023.

³ Coordenado pelas Professoras Doutoras Zenaide Carneiro e Mariana Lacerda. Cf. <http://www5.uefs.br/cedohs/view/home.html>. Acesso em: 31/05/2023.

⁴ “O entrelaçamento de ambas as disciplinas [Linguística Histórica e a Sociolinguística] em uma linguística sócio-histórica que investigue e explique as diversas manifestações da variação linguística em uma comunidade ao longo do tempo e também a forma como essa variação – seus diferentes usos, funções e tipos – se desenvolve em línguas e comunidades específicas, dentro de determinados grupos e redes sociais ou em indivíduos específicos

A aplicação dos métodos que a sociolinguística desenvolveu para estudos de processos de variação e mudança em línguas contemporâneas a dados do passado busca a reconstrução histórica da língua em sua relação com o contexto social. Nesse processo, por um lado, a investigação de mudanças em progresso pode lançar alguma luz à compreensão de fenômenos linguísticos do passado, auxiliando a inter-relação desses fenômenos com os fatores que os condicionam. Por outro lado, os processos de mudança em curso podem ser aclarados pelos resultados alcançados pela SH sobre mudanças linguísticas consumadas.

Um dos princípios centrais que orientam o trabalho em SH é o princípio da uniformidade, segundo o qual, partindo da percepção de que a língua do presente é inerentemente variável e que os processos de variação e mudança ocorrem de forma estruturada em correlação com variáveis independentes, não há razão para não se supor que também a realidade linguística do passado passou por estágios de variação e mudança ordenados e sistemáticos.

Uma vez que, em seu trabalho, depende diretamente do acesso a textos escritos — única via de aproximação à língua de comunidades pretéritas —, se impõe, para a SH, uma série de dificuldades. Hernández-Campoy e Schilling (2012) elencam sete desafios que se apresentam ao pesquisador na área:

- i. representatividade: como alguns textos alcançam os dias atuais por acaso, as amostras podem apresentar lacunas que comprometem sua representatividade. Ademais, apenas uma parcela reduzida dos estilos, registros ou variedades do passado é recuperada nos textos, o que se agrava ao se considerar que a escrita não foi praticada por todos os grupos sociais do passado;
- ii. validade empírica: as amostras frequentemente são limitadas e fornecem uma quantidade reduzida de dados, constituindo um entrave para o desenvolvimento de análises quantitativas. Somam-se a isso as limitações de acesso a informações socioculturais de comunidades pretéritas;
- iii. invariação: o caráter conservador da escrita pode restringir a probabilidade de variação. Possivelmente, a língua falada no passado apresentou mais processos de variação do que os escritos sobreviventes permitem conhecer;
- iv. autenticidade: não é difícil de as formas legadas pelos manuscritos conterem hipercorreção, mistura de dialeto e erro do escriba, ou mesmo serem cópia de outros documentos que se perderam com o tempo;
- v. autoria: a mão que escreve pode não ser a mesma que assina. Alguns documentos podem ter sido copiados ou ditados por uma segunda pessoa;
- vi. validação social e histórica: a reconstrução do perfil sociocultural do escrevente e da estruturação social da comunidade é tarefa de grande complexidade e só pode

[...]. Trata-se, portanto, de uma importante aplicação dos objetivos de uma sociolinguística que era histórica na medida em que era aplicada ao estudo dos fundamentos das mudanças linguísticas – como havia sido implicitamente concebida por Weinreich, Labov e Herzog – para a reconstrução das relações de covariação entre determinadas variáveis linguísticas e certos fatores sociais em estágios anteriores da evolução das línguas” (Tradução nossa).

ser realizada por via indireta, aspecto que se reflete na dificuldade de controle de variáveis independentes sociodemográficas em estudos variacionistas da língua do passado;

- vii. ideologia padrão: sob a égide da ideologia padrão, os estudos, por vezes, estabelecem uma correspondência direta entre os documentos de indivíduos considerados ilustres e a variedade padrão, o que nem sempre é efetivamente observado na história da língua. Analogamente, pode acontecer de os dialetos não padrão serem tomados como derivação ou perversão do padrão.

Acerca da representatividade, conquanto isso não seja a regra, alguns textos alcançam os dias atuais através de atos conscientes de preservação e são, por isso, menos lacunares. Apesar disso, em geral, as fontes disponíveis ao linguista são uma parte de uma produção escrita bem maior. Daí o caráter, por vezes, fragmentário das amostras. Tal dificuldade tem sido amenizada por meio do esforço conjunto de diversos estudiosos no estabelecimento de bancos de dados: uma amostra reduzida, disponibilizada por uma equipe somada ao *corpus* igualmente condensado de outro projeto, resulta menos lacunar. Consequentemente, o problema da validade empírica dos dados também se reduz.

A linguística de *corpus* também colabora para enfrentar essas questões, mediante a preparação, compilação e informatização de bancos de textos históricos, facilitando a manipulação, a seleção dos materiais por períodos e a comparação entre variedades de diferentes épocas, o que, como nota Conde-Silvestre (2007, p. 48), “[...] ayuda, en parte, a resolver el problema del aislamiento y el carácter fragmentario de los materiales históricos”.⁵

Para contornar o problema do conservadorismo da língua escrita, os pesquisadores têm priorizado gêneros que permitam alguma aproximação ao vernáculo, embora isso nem sempre seja possível. Sobre as fontes para investigação em SH, pontua Conde-Silvestre (2007, p. 45):

[...] los textos más útiles para el investigador son aquellos que trasladan al medio escrito intercambios comunicativos que han ocurrido o podrían haber ocurrido en el medio oral pues, en principio, este tipo de textos debería manifestar un grado mayor de variación y, por otro lado, facilitar la correlación de las variables lingüísticas con las circunstancias personales de sus emisores y receptores (CONDE-SILVESTRE, 2007, p. 45).⁶

As fontes mais requisitadas pelo pesquisador em SH são as que, além de indícios de usos linguísticos menos monitorados, consequentemente mais propícios à variação, fornecem pistas sobre o autor, o interlocutor e o contexto social.

⁵ “[...] ajuda, em parte, a resolver o problema do isolamento e do caráter fragmentário dos materiais históricos” (Tradução nossa).

⁶ “[...] os textos mais úteis para o pesquisador são aqueles que transferem para o meio escrito trocas comunicativas que ocorreram ou poderiam ter ocorrido no meio oral, pois, em princípio, esse tipo de texto deveria apresentar maior grau de variação e, por outro lado, facilitar a correlação das variáveis linguísticas com as circunstâncias pessoais de seus emissores e receptores” (Tradução nossa).

O problema da autenticidade é mais geral e se impõe, inclusive, para os estudos sociolinguísticos de variedades contemporâneas, tendo em conta que qualquer comunidade de fala está sujeita a episódios de contato interdialeto e fenômenos como variação estilística, que dificultam o acesso ao vernáculo de um indivíduo. É por essa razão que, acertadamente, Hernández-Campoy e Schilling (2012, p. 63-64) argumentam que “[...] we cannot hold historical sociolinguistics to standards with which sociolinguistics itself cannot comply”⁷.

A fixação da autoria configura um dos maiores desafios da SH. As possibilidades para enfrentamento desse aspecto não são muitas. Em amostras maiores, com muitos textos assinados pelo mesmo indivíduo, a margem de acerto na atribuição da autoria pode ser maior, por se contar com diferentes registros para uma comparação *escriptográfica*.

Na validação social e histórica da amostra, o diálogo entre SH e a história social tem se mostrado produtivo. Ocupando-se do estudo de grupos sociais pretéritos, da forma como se organizavam e interagiam, dos conflitos estabelecidos entre eles e dos aspectos que determinaram as relações humanas no passado — as relações familiares, as relações entre sexo, as relações de poder e solidariedade, o desenvolvimento de mentalidades e ideologias etc. — a história social fornece à SH parte importante da informação necessária à reconstrução das variáveis históricas independentes que podem ter interagido com os processos de variação e mudança na língua do passado. A reconstrução desses fatores, porém, deve evitar a transferência, para o passado, de modelos extraídos de circunstâncias atuais, a fim de que não se incorra em anacronismo (CONDE-SILVESTRE, 2012).

Muito se tem também em relação ao problema da ideologia padrão. Sabe-se, por exemplo, que indivíduos pertencentes aos segmentos privilegiados da sociedade podem não ter domínio total das técnicas da escrita padrão. A investigação do estilo, discutida por Conde-Silvestre (2012), colabora para uma compreensão adequada do binômio padrão *versus* não padrão e, por conseguinte, para a reconstrução das variáveis independentes relevantes à interpretação dos processos de variação e mudança em perspectiva histórica.

Se o pesquisador em SH consegue extrair dos textos características dos estilos históricos, ele pode, a partir dessas características, alcançar algum conhecimento sobre a estratificação social e o contexto social que, no passado, interagiram com os fenômenos de variação e mudança linguística. Em contrapartida, as investigações sobre o estilo demandam uma série de cuidados metodológicos para evitar associações equivocadas. É central considerar que formalidade e informalidade são fatores multidimensionais — não se relacionam simplesmente ao contexto, de um lado, e ao tema da comunicação, de outro. Ademais, o estilo pode trazer alguma informação útil

⁷ “[...] não podemos manter a sociolinguística histórica em padrões que a própria sociolinguística não pode cumprir” (Tradução nossa).

quando não se dispõe de mais informações contextuais, mas a adequada aproximação ao contexto social perpassa pela consideração de aspectos de difícil acesso para as investigações históricas, como o prestígio das formas envolvidas, a variação dessas formas ao longo da história e o acesso que os diferentes grupos tiveram aos diferentes textos (CONDE-SILVESTRE, 2012).

Até aqui, procurou-se mostrar que, apesar das dificuldades, a SH tem desenvolvido métodos e técnicas para correlacionar as circunstâncias sociais aos usos linguísticos do passado a partir de bases de dados mais confiáveis, com o auxílio de áreas como a linguística de *corpus* e a história social. É nesse esforço que tem contribuído para interpretação da mudança e da variação em perspectiva histórica.

3 MOTIVAÇÕES E TENDÊNCIAS DE ESTUDOS DESENVOLVIDOS POR PESQUISADORES BAIANOS

Mattos e Silva (1998, p. 40), ao propor uma agenda de trabalho, no texto já mencionado, estabelece quatro campos de pesquisa, dos quais o segundo, campo *b*, é aquele que “[...] se moverá na reconstrução de uma sócio-história linguística ou de uma sociolinguística histórica”⁸.

Alguns estudos desenvolvidos por pesquisadores baianos atendem a essa agenda, pondo em diálogo os campos propostos, com o estabelecimento de relações entre fatos linguísticos e fatores sociais, mobilizando esforços com o objetivo de tentar captar a variação histórica do português brasileiro no âmbito do encaixamento na sociedade (*embedding problem*), da difusão de variantes na sociedade (*transition*) e, ainda, da captação das avaliações que favorecerão ou não a mudança (*evaluation*).

A tentativa de aproximar esse tipo de abordagem para o tempo real de longa duração, conforme Mattos e Silva (1998), é um trabalho árduo, porém não impossível, como já comentado na seção anterior. Para a questão do encaixamento na sociedade, segundo a pesquisadora, é preciso um conjunto significativo de documentação representativa, desde os estilos mais informais até os mais formais. Nessa perspectiva, *os estilos de texto* serão o análogo às classes sociais dos estudos sincrônicos. Para se captar a difusão na sociedade, a documentação deverá estar distribuída seriadamente no tempo do português, ou seja, a difusão das mudanças na sociedade brasileira ao longo do seu tempo histórico. E, para a avaliação social, torna-se preciso recorrer a fontes indiretas (como gramáticas, epistolografia dos jesuítas, narrativas de viajantes etc.).

Como trabalhos pioneiros, podem-se citar a tese de Tânia Lobo (2001), defendida na Universidade de São Paulo, cuja relação com o campo fica explícita já no título: *Para uma sociolinguística histórica do português do Brasil: edição filológica e análise linguística*

⁸ Os outros três campos são: a. o campo que se moverá na reconstrução de uma história social linguística do Brasil; c. o campo que se moverá na reconstrução diacrônica no interior das estruturas da língua portuguesa em direção ao português brasileiro; e d. o campo que se moverá no âmbito comparativo entre o português europeu e o português brasileiro.

de cartas particulares do Recôncavo da Bahia, século XIX, assim como a tese de Zenaide Carneiro (2005), intitulada *Cartas brasileiras (1809-1904): um estudo linguístico-filológico*, defendida da Universidade Estadual de Campinas. São pesquisas que consideram fatores sociais e históricos na descrição das ocorrências do fenômeno linguístico — ainda que a interpretação dos dados se apoie em uma tendência gerativista — e disponibilizam um conjunto significativo de documentação, editada com critérios filológicos precisos.

Nessa perspectiva, Mattos e Silva orientou alguns trabalhos — mesmo que nem sempre os autores mencionem diretamente o aporte metodológico da SH —, por exemplo os estudos de Costa (2002) e Souza (2005), para citar apenas dois. Ambos desenvolvem estudos variacionistas com dados coletados em documentos de variedades linguísticas pretéritas — textos portugueses do século XIII ao XVI, em Costa (2002), e dos séculos XIII e XIV, em Souza (2005). Aquela descreveu a presença variável do artigo definido antes de pronomes possessivos e de nomes próprios de pessoas a partir da quantificação dos contextos que poderiam influenciar a presença ou ausência do artigo definido — com nomes próprios, a autora identifica ausência praticamente categórica do artigo. Este estudou a concordância verbal e, com a análise quantitativa dos dados, notou que a variação identificada nos textos é condicionada por fatores linguísticos (de natureza morfofônica, sintática e semântica) e social (o tipo de texto).

Perspectiva semelhante à dos trabalhos de Costa (2002) e Souza (2005) é observada em estudos realizados pela equipe do Projeto CE-DOHS. Além do investimento na constituição de *corpora* para estudos linguísticos históricos, esses pesquisadores têm realizado análises gramaticais a partir das amostras constituídas. São estudos que adotam os pressupostos teórico-metodológicos da sociolinguística quantitativa (WEINREICH; LABOV; HERZOG, 2006 [1968]; LABOV, 2008 [1972]) para observar, a partir dos textos, fenômenos em variação e mudança na língua do passado. Apesar de não explicitarem em seu aporte metodológico, os trabalhos desenvolvem as técnicas da SH e enfrentam os desafios impostos pela investigação nessa área.

O CE-DOHS, desde 2012, tem investido no trabalho de prospecção, edição e disponibilização de *corpora* para estudos linguísticos. Em língua escrita, o projeto já soma mais de um milhão de palavras, em textos de diferentes gêneros — com destaque para documentação epistolar —, organizados em 36 acervos que abrangem as duas fases do projeto: a primeira, que reúne textos escritos entre 1823 e 2000, por indivíduos nascidos no Brasil a partir de 1724; e a segunda, iniciada em 2019 e ainda em preparação, com manuscritos produzidos entre 1640 e 1822, por brasileiros nascidos a partir de 1590. À etapa atual do projeto, paulatinamente, têm-se incorporado textos produzidos no Brasil por portugueses, nos primeiros 150 anos de colonização.

Nos documentos reunidos, a equipe realiza a edição filológica, em conformidade com as normas de edição do Projeto Para História do Português Brasileiro — PHPB — e a edição em linguagem XML, com auxílio do eDictor (PAIXÃO DE SOUZA; KEPLER; FARIA, 2009). Alguns acervos estão sintaticamente anotados, outros em fase de

anotação sintática, a partir da parceria firmada entre o CE-DOHS e o Projeto *Corpus* Histórico do Português *Tycho Brahe*.

O levantamento de informações sobre os escreventes e os documentos é uma preocupação central no Projeto. Além da edição fidedigna dos textos, almeja-se o controle de *onde, quando, por quem e para quem* o texto foi escrito (SANTIAGO *et al.*, 2021). Tem sido dada atenção especial à possibilidade de se reconhecerem os índices de habilidade/inabilidade dos escreventes, na tentativa de disponibilizar *corpora* em que se garanta essa melhor caracterização da *mão* que escreve.

Os materiais que compõem o banco de dados do CE-DOHS estão disponíveis na rede mundial de computadores para acesso livre e gratuito. Ao acessar a plataforma, o pesquisador pode recortar sua própria amostra, segundo o seu interesse, observando, em relação ao escrevente, a etnia (indígenas, brancos, negros do Brasil, mulatos, mamelucos e pardos), a escolaridade, o sexo, a profissão, o local de nascimento e a estratificação social; e em relação ao documento, no caso dos materiais escritos, a data e local de escrita e para que e para quem o texto foi destinado.

Além de suas contribuições ao processo de localização, preparação e disponibilização de materiais para os estudos linguísticos históricos, a equipe do CE-DOHS tem investido na análise de fenômenos gramaticais a partir dos *corpora* constituídos. Para exemplificar, sumarizam-se resultados de cinco dissertações desenvolvidas por pesquisadores dessa equipe, no âmbito do Programa de Pós-Graduação em Estudos Linguísticos (PPGEL/UEFS),⁹ que seguem essa linha investigativa.

São trabalhos que utilizaram como *corpus* a parte 1 do acervo *Cartas em Sisal*, reunido e editado por Santiago (2012, 2019). Trata-se de uma amostra constituída por cartas pessoais, escritas entre 1906 e 2000, por sertanejos baianos pouco escolarizados, caracterizados como escreventes inábeis na técnica de escrita. As missivas tratam de assuntos do cotidiano, da lida no campo, de confidências de namorados e são também um instrumento para envio e recebimento de notícias. São correspondências trocadas entre familiares e amigos em relações marcadas pela simetria de poder e que sugerem um uso mais espontâneo da escrita, configurando-se em um *corpus* de especial valor para os trabalhos em SH.

Mascarenhas (2016) descreveu as estratégias de relativização utilizadas nas cartas. A autora analisou 121 construções relativas, sendo 55 de relativas restritivas, 48 de não restritivas e 18 de relativas livres. Notou, para as relativas restritivas e não restritivas, que a função de sujeito é a posição mais relativizada (68,6%), seguida da função de objeto direto (26,5%). Não se encontraram relativas restritivas e não restritivas nas posições objeto indireto, oblíquo e genitivo (as chamadas *posições baixas*). Relativas não

⁹ Além dos cinco trabalhos citados neste texto, há a dissertação de Lemos (2019), também desenvolvida no âmbito do PPGEL/UEFS a partir do acervo *Cartas em Sisal*, que não é citada em maiores detalhes aqui porque a autora não associa seu trabalho ao modelo variacionista, embora, em termos metodológicos, não se diferencie substancialmente das dissertações aqui citadas, já que também procura interpretar o tema em análise, no caso de Lemos (2029) a sintaxe dos clíticos, a partir do cotejo de aspectos linguísticos e sociais, lançando mão, para tanto, da quantificação dos dados.

padrão (cortadora) e padrão (*pied piping*) não ocorreram em posições baixas em construções de relativas restritivas e não restritivas. Essas estratégias ocorreram apenas na função de adjunto, totalizando cinco dados, todos em relativas restritivas (4 construções com a cortadora e uma com *pied piping*). Nas relativas livres, localizaram-se um caso de estratégia com pronome lembrete (em objeto direto) e um caso de *pied piping* (em objeto indireto). As estratégias com lacuna ocorreram nas posições mais altas da hierarquia (sujeito e objeto direto).

Santos (2017) analisou a variação na concordância nominal de número em perspectiva mórfica (ou atomística), analisando 710 termos, dos quais 420 (59%) apresentaram marcas de plural; e numa abordagem sintagmática, descreveu 318 sintagmas nominais, que apresentaram 31% de marcação formal de pluralidade. Na investigação ao nível do sintagma, encontrou maior incidência de marcas de plural em constituintes à direita do verbo e em sintagmas compostos por dois elementos. Na análise atomística, a pesquisadora observou que o núcleo em primeira posição e o elemento à esquerda adjacente ao núcleo retêm mais a marca de plural, e a saliência fônica não revelou uma atuação clara da maior ou menor oposição da substância fônica entre singular e plural e marcação de plural no elemento. Foi possível verificar também que o processamento paralelo não é relevante para a marcação da concordância nominal de número nas cartas. Os fatores sociais analisados foram os mesmos nas duas análises e exibiram resultados semelhantes: os indivíduos que estudaram até a quarta série tendem a marcar mais a concordância e há taxas superiores de concordância quando o elemento não é realizado em fórmula.

Oldack Barbosa (2018) considerou o uso dos pronomes possessivos de segunda pessoa *teu* e *seu* e observou que *seu*, como paradigma de segunda pessoa, já está plenamente implementado nas cartas. Das 185 ocorrências de pronomes possessivos analisadas, somente 17 foram da forma *teu* (9.2%). E, das 168 realizações de *seu*, apenas uma foi em referência à terceira pessoa. Sobre os condicionamentos do uso das formas possessivas de segunda pessoa no *corpus*, Oldack Barbosa (2018) observou, em relação aos fatores estruturais, que as formas *seu* e *teu* ocorreram, sobretudo, em contexto de posse alienável. O pronome *seu* também foi amplamente utilizado com os traços [+animado +humano] e [-animado]. Na dimensão extralinguística, constatou-se que, somente na década de 1980, os usos de *teu* e *seu* exibiram certo equilíbrio, sendo a forma *seu* favorecida nos demais períodos abrangidos pelas cartas. Homens e mulheres preferiram a forma *seu*, mas as poucas realizações de *teu* foram mais empregadas por mulheres (11 dos 17 dados). Ademais, concluiu-se que a maior intimidade entre os missivistas incide na taxa mais elevada de uso de *seu*, do mesmo modo que o uso dessa forma é favorecido nas cartas marcadas por relações simétricas [+solidárias].

Brito (2019) analisou a combinação do artigo definido com pronome possessivo. A autora levantou, nas cartas, 334 sintagmas possessivos em que é possível a alternância presença/ausência do artigo definido, sendo o artigo realizado em 56,60% dos casos (189/334). A partir da análise quantitativa dos dados, constatou-se que o fenômeno

em análise é condicionado, sobretudo, por fatores de ordem linguística. A presença de uma preposição antes do pronome possessivo, quando se considera a possibilidade de contração do artigo com a preposição, foi um dos principais condicionantes da omissão ou realização do artigo definido, além de incidir na atuação da variável função sintática. No que concerne aos aspectos semânticos da variação, destacam-se os grupos: tipo de possuído, que relaciona a presença do artigo à ideia de posse expressa no elemento possuído; animacidade, que aponta para uma reduzida presença do artigo em constituintes cujo núcleo é caracterizado pelo traço [+humano]; e tipo de substantivo, que indica que a presença do artigo é favorecida quando o núcleo do sintagma possessivo é um nome abstrato. Os fatores sociais (naturalidade, data de escrita da carta, local de escrita das cartas e gênero do remetente), embora selecionados pelo programa de processamento estatístico dos dados, exibiram uma distribuição pouco regular em Brito (2019).

Santos (2019) realizou o estudo do sistema de tratamento em referência à segunda pessoa do singular na posição de sujeito, correlacionando-o às formas empregadas na função de acusativo, dativo e complemento oblíquo. Foram consideradas 93 formas de tratamento com função de sujeito, sendo *você* a mais frequente (50 ocorrências). A forma *vosmecê* aparece exclusivamente nas cartas da década de 1900 (25 ocorrências), e a forma *o/a senhor/a* foi realizada 17 vezes como sujeito. A presença escassa do pronome *tu* com função de sujeito nas cartas (um único dado) mostra que a forma *você* é majoritária desde a primeira década do século XX — sendo empregada, nesse momento, através da forma de origem nominal *vosmecê*. Como complemento verbal, foram analisadas 226 formas pronominais de segunda pessoa nas cartas (129 como dativo; 71 como acusativo e 26 como oblíquo). Ao subsistema tratamental de *você* exclusivo, preponderante nas cartas, correlacionam-se os complementos verbais *te*, para os acusativos; *lhe*, para os dativos, e de *você*, como oblíquo.

Embora nenhum dos estudos supracitados assumam diretamente filiação à SH, todos deixam claro, em sua metodologia, que recorreram aos métodos sociolinguísticos para análise dos fenômenos que estudam no acervo *Cartas em Sisal*:

- “Este trabalho [...] Apoiar-se também nos pressupostos da teoria Sociolinguística variacionista (WEINREICH, LABOV E HERZOG, 1968), na direção do que é defendido por Labov (1972) [...]” (MASCARENHAS, 2016, p. 45);
- “A pesquisa toma como escopo teórico a Linguística Histórica *stricto sensu*, nos termos definidos por Mattos e Silva (2008) e a Sociolinguística Laboviana (LABOV, 1972) [...]” (SANTOS, 2017, s./p.);
- “Os dados, em sua análise qualitativa e quantitativa, tiveram como base o aporte teórico da sociolinguística variacionista [...]” (OLDACK BARBOSA, 2018, s./p.);
- “A partir do aparato teórico-metodológico da Sociolinguística Quantitativa, nesta pesquisa, analisa-se a variação no uso do artigo definido diante de possessivos [...]” (BRITO, 2019, p. 74);
- “[...] através do aparato teórico-metodológico da sociolinguística variacionista laboviana (WEINREICH, LABOV E HERZOG, 2006 [1968]; LABOV, 2008 [1972]) apresenta um estudo sobre o sistema de tratamento em cartas pessoais [...]” (SANTOS, 2019, s./p.).

As dissertações de Mascarenhas (2016), Santos (2017), Oldack Barbosa (2018), Brito (2019) e Santos (2019)¹⁰ adotam os fundamentos gerais do modelo da sociolinguística, assumindo, a partir de Weinreich, Labov e Herzog (2006 [1968]) e Labov (1972), a heterogeneidade ordenada do sistema linguístico e a necessidade de correlacionar aspectos da estrutura linguística e do contexto mais amplo de uso da língua para análise e interpretação dos processos de variação e mudança. A partir do levantamento, classificação e quantificação dos dados do fenômeno em análise em cada dissertação, os autores descrevem os grupos de fatores estruturais e sociais que condicionam a realização das formas variantes.

As dificuldades para o trabalho com textos históricos se fazem notar nessas pesquisas, especialmente na quantidade dos dados analisados e em sua distribuição temporal. As cinco dissertações trabalharam com um número reduzido de dados, o que, além de ser um problema para a quantificação das variáveis controladas, é uma barreira para uma interpretação confiável. Os resultados alcançados sugerem o comportamento do fenômeno no *corpus*, mas dificilmente se pode chegar a uma conclusão mais geral sobre o tema a partir deles. Os autores parecem atentar para esse fato e não almejam maiores generalizações a partir dos dados encontrados nas cartas.

Quanto à localização temporal, as 91 missivas foram escritas entre 1906 e 2000, mas estão distribuídas de modo bem irregular ao longo do período¹¹:

— 1ª metade do século XX: década de 1900 - 6 cartas.

— 2ª metade do século XX: década de 1950 – 16 cartas; década de 1960 – 24 cartas; década de 1970 – 20 cartas; década de 1980 – 1 carta; década de 1990 – 4 cartas; 2000 – 1 carta.

— Sem indicação de data: 19 cartas.

Da primeira metade do século XX, como se vê, só há cartas da década de 1900 e em quantidade reduzida. Para o final do século, também há poucos documentos. Os resultados dos estudos a partir dessas missivas, considerando a datação, representam um intervalo de tempo muito curto, basicamente as décadas de 1950, 1960 e 1970, que concentram a maioria das cartas.

Apesar dessas dificuldades, os trabalhos em questão se beneficiam de uma detalhada caracterização sócio-histórica dos documentos e dos escreventes, disponibilizada por Santiago (2012, 2019), o que auxiliou os pesquisadores a superarem um dos maiores percalços da investigação em sociolinguística histórica: a reconstrução do contexto social. Santiago (2012, 2019) elenca informações sobre o perfil biográfico dos remetentes, a partir de entrevistas realizadas com os destinatários e com os remetentes e seus familiares, da consulta a documentos pessoais de alguns escreventes e do

¹⁰ As dissertações de Mascarenhas (2016), Santos (2017) e Oldack Barbosa (2018) foram orientadas pela Profa. Dra. Zenaide Carneiro, e as de Brito (2019) e Santos (2019), pela Profa. Dra. Mariana Lacerda.

¹¹ O acervo completo das *Cartas em Sisal*, parte 1 e 2, é constituído por 131 cartas.

conteúdo das próprias cartas. Através desse trabalho, é possível saber a profissão, a idade quando da escrita da carta, a naturalidade, a filiação e o estado civil da maioria, além de notícias de como se deu o contato dos escreventes com as primeiras letras e os materiais de leitura a que tinham acesso na região. O contexto no qual essas cartas circularam também foi objeto de interesse, com a apresentação de aspectos históricos, políticos e sociais dos municípios de origem dos missivistas, no semiárido baiano.

4 CONSIDERAÇÕES FINAIS

Conquanto o trabalho com documentos do passado, sobretudo de períodos mais recuados, lance ao pesquisador em SH uma série de desafios, as pesquisas na área têm avançado significativamente no que concerne à confiabilidade dos resultados. Conde-Silvestre (2012) assinala que a própria sociolinguística tem desenvolvido ferramentas para lidar com as dificuldades impostas pela pesquisa com textos históricos — como o princípio da uniformidade e a aplicação das relações entre os contínuos estilístico e sociodialetal —, e a SH tem se beneficiado, ainda, do desenvolvimento de disciplinas como a linguística de *corpus* e a história social. Ademais, como pontuam Hernández-Campoy e Schilling (2012), limitações teóricas e procedimentais se impõem, inclusive, para a sociolinguística, que trabalha com dados recentes. Há, portanto, cada vez menos razões para se questionarem a representatividade e a validade empírica dos trabalhos em SH.

Pesquisas nesse campo têm sido desenvolvidas, por exemplo, por pesquisadores baianos: embora a denominação da disciplina nem sempre apareça nos trabalhos, aplicam a dados extraídos de documentos do passado a metodologia desenvolvida pela sociolinguística e lidam, conseqüentemente, com as dificuldades decorrentes dessa transposição.

REFERÊNCIAS

- BRITO, R. C. **Uso variável dos artigos definidos antes de possessivos em cartas pessoais do sertão baiano (século XX)**. 2019. 553f. Dissertação (Mestrado em Estudos Linguísticos) – Programa de Pós-Graduação em Estudos Linguísticos, Universidade Estadual de Feira de Santana, Feira de Santana.
- CONDE SILVESTRE, J. C. Problemas y principios. In: CONDE SILVESTRE, J. C. **Sociolinguística histórica**. Madrid, Gredos, 2007. p. 19-72.
- COSTA, Iraneide. O uso do artigo definido diante de nome próprio de pessoa e de possessivo do século XIII ao século XVI. In: MATTOS E SILVA, Rosa Virgínia; MACHADO FILHO, Américo Venâncio Lopes (Org.). **O português quinhentista: estudos linguísticos**. Salvador: EDUFBA; Feira de Santana: UEFS, 2002. p. 284-306.
- HERNÁNDEZ-CAMPOY, J. M.; SCHILLING, N. The Application of the quantitative paradigm to historical sociolinguistics: Problems with the generalizability principle. In: HERNÁNDEZ-CAMPOY, J. M.; CONDE SILVESTRE, J. C. (Orgs.). **The handbook of historical sociolinguistics**. Oxford: Wiley-Blackwell, 2012. p. 63-94.
- LABOV, W. **Padrões Sociolinguísticos**. Tradução: Marcos Bagno, Maria Marta Pereira Scherre e Caroline Rodrigues Cardoso. São Paulo: Parábola Editorial, 2008 [1972].

- LEMOS, M. da S. **A colocação dos clíticos em sentenças finitas: um estudo sócio-histórico das vertentes do PB em cartas do sertão baiano (século XX)**. 2019. 830f. Dissertação (Mestrado em Estudos Linguísticos) – Programa de Pós-Graduação em Estudos Linguísticos, Universidade Estadual de Feira de Santana, Feira de Santana.
- LOBO, T. C. F. **Para uma sociolinguística histórica do português do Brasil: edição filológica e análise linguística de cartas particulares do recôncavo da Bahia, século XIX**. 2001. 808f. Tese (Doutorado em Filologia e Língua Portuguesa) – Universidade de São Paulo, São Paulo.
- MASCARENHAS, J. de O. C. **Sentenças relativas em cartas de inábeis**. 2016. 232f. Dissertação (Mestrado em Estudos Linguísticos) – Programa de Pós-Graduação em Estudos Linguísticos, Universidade Estadual de Feira de Santana, Feira de Santana.
- MATTOS E SILVA, R. V. **Ensaio para uma sócio-história do português brasileiro**. São Paulo: Parábola Editorial, 2004.
- MATTOS E SILVA, R. V. Ideias para a história do português brasileiro: fragmentos para uma composição posterior. In: CASTILHO, A. T. de. (Org.). **Para a história do português brasileiro**. Vol. I: primeiras ideias. São Paulo: Humanitas, 1998. p. 21-52.
- OLDACK BARBOSA, G. M. **O uso dos pronomes possessivos teu e seu em cartas pessoais de sertanejos baianos do século XX**. 2018. 242f. Dissertação (Mestrado em Estudos Linguísticos) – Programa de Pós-Graduação em Estudos Linguísticos, Universidade Estadual de Feira de Santana, Feira de Santana.
- PAIXÃO DE SOUSA, M. C.; KEPLER, F.; FARIA, P. E-dictor: Novas perspectivas na codificação e edição de corpora de textos históricos. In: VIII Encontro de Linguística de Corpus, 2009. Rio de Janeiro, **Anais do VIII Encontro de Linguística de Corpus**. Rio de Janeiro: UERJ, 2009. p. 69-105. Disponível em: http://www.tycho.iel.unicamp.br/~tycho/pesquisa/artigos/PAIXAO-DE-SOUZA_etal-ECL2009.pdf. Acesso em: 20 fev. 2021.
- SANTIAGO, H. da S. **A escrita por “mãos inábeis”**: uma proposta de caracterização. 2019. 722f. Tese (Doutorado em Língua e Cultura) – Universidade Federal da Bahia, Salvador.
- SANTIAGO, H. da S. **Um estudo do português popular brasileiro em cartas pessoais de “mãos cândidas” do sertão baiano**. 2012. 2v. ?f. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Estadual de Feira de Santana, Feira de Santana.
- SANTIAGO, H.; LACERDA, M.; BRITO, R.; CARNEIRO, Z. CE-DOHS: um banco de dados sociolinguísticos para a história do português brasileiro. **LaborHistórico**, Rio de Janeiro, v. 7, n. Especial, p. 311-329, 2021.
- SANTOS, E. S. e. **Para a história do português popular brasileiro: o sistema de tratamento em cartas baianas do século XX**. 2019. 378f. Dissertação (Mestrado em Estudos Linguísticos) – Programa de Pós-Graduação em Estudos Linguísticos, Universidade Estadual de Feira de Santana, Feira de Santana.
- SANTOS, L. E. R. **A variação da concordância nominal de Número em cartas de inábeis do sertão baiano (1906-2000)**. 2017. 239f. Dissertação (Mestrado em Estudos Linguísticos) – Programa de Pós-Graduação em Estudos Linguísticos, Universidade Estadual de Feira de Santana, Feira de Santana.
- SOUZA, P. D. dos S. **Concordância verbal em português: o que nos revela o período arcaico?** 2005. 95f. Dissertação (Mestrado em Letras e Linguística) – Instituto de Letras, Universidade Federal da Bahia, Salvador.
- WEINREICH, U.; LABOV, W.; HERZOG, M. **Fundamentos empíricos para uma teoria da mudança linguística**. São Paulo: Parábola, 2006 [1968].

ASSENTOS DE CASAMENTOS DA FREGUESIA DE SANTO ANTÔNIO DO URUBU DE BAIXO DO RIO SÃO FRANCISCO DO SÉCULO XVIII: ANÁLISE DA HABILIDADE DOS SCRIPTORES

LÉCIO BARBOSA DE ASSIS
JORGE AUGUSTO ALVES DA SILVA
VERA PACHECO

1 INTRODUÇÃO

Os registros paroquiais são fontes primárias normalmente utilizadas para o estudo da história social, representando um campo de variadas possibilidades de investigação. Aqui, busca-se contribuir para a reconstituição da história da língua e de um povo, em determinado tempo e espaço, utilizando um *corpus* documental ainda pouco explorado nos estudos linguísticos. Os manuscritos produzidos no sertão do Rio São Francisco do século XVIII têm o potencial de trazer informações sobre o contexto sócio-histórico marcado pelos conflitos dos períodos históricos da ortografia da língua portuguesa (COUTINHO, 1976), revelando a cultura linguística da época.

Nesse sentido, o *corpus* extraído dos assentos de casamentos do primeiro livro da Freguesia de Santo Antônio do Urubu de baixo do Rio São Francisco (1719-1753) tem o intuito de mostrar alguns traços da escrita em uso no século XVIII, com o objetivo de analisar a habilidade dos *scriptores*, através dos aspectos físico-gráficos e gráficos e fonéticos, como o traçado das letras, substituições de segmentos, segmentação vocabular, alçamento de vogais, substituições consonantais, formas etimologizadas, entre outros.

Trata-se de uma atividade que recorre ao labor filológico para o “estudo rigoroso dos documentos escritos antigos e de sua transmissão para estabelecer, interpretar e editar esses textos” (CAMBRAIA, 2005, p. 14) e ao trabalho paleográfico, imprescindível para “se entender como se constituíram sócio-historicamente os sistemas de escrita” (CAMBRAIA, 2005, p. 23). Os documentos manuscritos contribuem para o conhecimento das práticas sociais e culturais de determinada época, permitindo o acesso ao contexto sócio-histórico por meio dos dados linguísticos registrados nos textos. Recorrer aos preceitos da sociolinguística histórica para “reconstruir as informações extralinguísticas sobre os autores dos documentos que são objeto de estudo

e tentar estabelecer correlações com o andamento histórico de certas mudanças, plenamente atestadas na história das línguas” (CONDE-SILVESTRE, 2007, p. 14), torna-se importante para compreender a comunidade de fala e de prática do sertão de baixo do período em análise.

No século XVIII, a língua portuguesa herdou as contradições ortográficas do passado, causadas pela indecisão dos ortógrafos, divididos entre a tradição, o costume ortográfico, a etimologia e as realidades fonéticas da língua (MOREL PINTO, 1988). A publicação da *Ortografia Nacional* de Gonçalves Viana, em 1904, foi um marco para o início do período simplificado, com o acordo convencional na padronização das normas, depois de longos períodos marcados pela arbitrariedade e capricho dos escribas e copistas.

Ao analisar os traços da escrita em uso nos assentos de casamentos do século XVIII, busca-se obter uma compreensão mais profunda dos processos diacrônicos que determinam as mudanças implementadas ou em curso, refletindo os conflitos entre inovação e conservação nos diversos aspectos linguísticos analisáveis (SILVA, 2011).

Os assentos de casamentos, fonte documental deste estudo, foram escritos pelos sacerdotes responsáveis pela freguesia. Esses homens pertenciam a um grupo social que, de forma geral, tinha alto nível de letramento e seguia um modelo orientado pela igreja, através das Constituições¹ Primeiras do Arcebispado da Bahia (1719), com base no Concílio de Trento (1545-1563) para lavrar os registros paroquiais. A língua oficial eclesiástica se sobressai nos assentos de casamentos, no entanto alguns traços de variação gráfica são encontrados no documento, desencadeando a seguinte reflexão: por que a escrita dos assentos de casamentos revela traços específicos da oralidade na grafia, se os *scriptores* pertenciam a uma elite instruída na região do baixo São Francisco no referido período? Através da observação dos aspectos físico-gráficos e gráficos e fonéticos retratados pela pena dos *scriptores*, é possível delinear a hipótese de que, no século XVIII, eles eram influenciados pela oralidade e pelo princípio etimológico, reflexo do conflito existente entre os períodos da história da ortografia da língua portuguesa em que o manuscrito foi produzido (1719-1753).

No que diz respeito à metodologia adotada, a análise desenvolvida com os 231 assentos de casamentos, trilharam os seguintes passos: (i) exploração das particularidades do suporte material do documento; (ii) leitura e transcrição do manuscrito a partir das reproduções fac-similares; (iii) sistematização e classificação dos aspectos físico-gráficos e gráficos e fonéticos com auxílio da ferramenta computacional AntCont² (ANTHONY, 2005; 2011) e (iv) análise das características gráficas de cada *scriptor*.

O intuito de pesquisar os arquivos paroquiais da Cúria Diocesana de Bom Jesus da Lapa (BA) se deu pelo interesse em realizar análises linguísticas em textos produ-

¹ As Constituições foram promulgadas em 1707 e só foram impressas em 1719.

² A ferramenta computacional AntConc é utilizada para a exploração e análise de *corpora*. O programa permite a sistematização e classificação de dados através das diferentes opções de análise, como lista de palavras (*Word List*), lista de concordâncias (*Concordance*) e palavras-chave (*KeyWord In Context – KWIC*) (ANTHONY, 2005; 2011).

zidos no contexto sócio-histórico do século XVIII, tendo em vista que “cada registro de um texto escrito constitui um testemunho” (CAMBRAIA, 2005, p.63), e dessa forma, contribuir para o conhecimento mais detalhado de um legado cultural que registra as características pretéritas da língua portuguesa no sertão do São Francisco.

2 PRINCÍPIOS TEÓRICO-METODOLÓGICOS DA SOCIOLINGÜÍSTICA HISTÓRICA

A sociolinguística histórica tem como objetivo a reconstrução dos fatores linguísticos e extralinguísticos nos diferentes períodos da história das línguas. Os dados do passado são correlacionados com fatores sociodemográficos como idade, *status* social, formação profissional e acadêmica, gênero, domicílio, histórico de migração e rede de relacionamentos pessoais. A questão básica é a necessidade de definir e apresentar a difusão das mudanças históricas do uso da língua, tanto no tempo real ou aparente (aspecto histórico), quanto no espaço (aspecto social) (CONDE-SILVESTRE; HERNÁNDEZ-CAMPOY, 2012).

De fundamental importância para os estudos sociolinguísticos históricos, a interface multidisciplinar facilita o acesso às informações específicas de determinado recorte temporal através das circunstâncias sócio-históricas que motivaram a variação e mudança linguística no passado. Para Bergs (2005), a sociolinguística histórica evidencia outras áreas do conhecimento sobrepostas a ela, refletindo a natureza integrativa das abordagens dessas ciências, como a história social, a filologia, a paleografia, a linguística de *corpus*, entre outras.

O autor afirma, ainda, que a sociolinguística histórica deve ser encontrada na intersecção de três diferentes campos: história, ciências sociais e linguística. Ela está focada na variação linguística no passado e tenta descrevê-la de forma socialmente motivada em um ponto particular nesse passado, ou seja, o estado sincrônico de uma língua e suas variedades, independentemente de serem dinâmicas ou estáticas.

É importante situar que a concepção dos estudos sociolinguísticos históricos teve início com a obra *Empirical Foundations for a Theory of Language Change* de Uriel Weinreich, William Labov e Marvin I. Herzog (1968). Os autores argumentaram que a linguística histórica tradicional deveria ser repensada por um viés sociolinguístico e apresentaram a necessidade de envolver os fatores externos em uma teoria da mudança da linguagem, numa abordagem histórica e social.

Em 1982, com a obra *Socio-Historical Linguistics — its status and methodology*, Suzanne Romaine apresentou os fundamentos para a teoria da linguística sócio-histórica ao investigar os marcadores de orações relativas em textos de diferentes níveis estilísticos do inglês do período escocês médio (1530-1550). A autora utilizou, em seu estudo, como variáveis externas, os diferentes tipos de textos e estilos. Em relação aos textos, referiu-se a prosa (oficial, literária, epistolar) e verso (religioso, moralizante e cômico), havendo, assim, a possibilidade de analisar duas categorias de estilo, o literário e o não literário, que influenciaram a ocorrência do uso dos marcadores relativos.

A pesquisadora propôs a junção dos métodos de análise da sociolinguística variacionista aos da linguística histórica a fim de usar o passado para explicar o presente, no sentido de confrontar a teoria sociolinguística com evidências de variação através do tempo, e o modo como funções, usos e tipos se desenvolvem em determinadas línguas e comunidades de fala.

Francisco Gimeno Menéndez (1995), com a obra *Sociolingüística histórica (siglos X-XII)*, tentou redefinir a linguística histórica com base na realização de pesquisas diacrônicas utilizando o modelo sociolinguístico para explicar o processo da mudança linguística incorporando os fatores sociais, com o objetivo de reconstruir a história social da língua espanhola do século X ao XII. De acordo com o autor, a sociolinguística histórica utiliza a metodologia sociolinguística para investigar os estágios iniciais de uma língua, com vistas a entender os processos de mudança e explicar o presente ou a mudança linguística, estabelecendo controles sobre os processos de reconstrução da língua do passado dentro do contexto social de uma comunidade de fala (GIMENO MENÉNDEZ, 1995, p. 15).

Outro nome de destaque nos estudos sociolinguísticos históricos foi Juan Camilo Conde Silvestre (2007), cuja obra, *Sociolingüística Histórica*, destaca os princípios e as metodologias da sociolinguística aplicados aos textos históricos, de modo que a investigação do presente se aplica à elucidação do passado. Do mesmo modo, Conde-Silvestre e Hernández-Campoy (2012) postulam que a sociolinguística histórica tem sido uma contribuição enriquecedora para o diálogo do presente e do passado, em conexão com o contexto sócio-histórico da evolução dos sistemas linguísticos de determinada língua.

Na percepção de Labov (1972), o diálogo entre o passado e o presente é intercambiável para a aplicar o princípio do uniformitarismo, que utiliza o presente para explicar o passado. A esse fato, Labov (1994) chama de paradoxo histórico. Assim, é necessário saber que o passado foi diferente, embora não seja fácil precisar o quanto. O teórico também menciona a questão de fazer o melhor uso de dados ruins, relacionando-a com o problema em relação aos dados históricos que, muitas vezes, não estão disponíveis nos documentos pesquisados.

Em pesquisas com dados históricos, a sociolinguística histórica tem guiado a investigação a partir da relação dos fatores linguísticos e extralinguísticos na mudança da língua ao longo do tempo. Esses fatores têm se mostrado bastante relevantes e, neste estudo, a respeito da habilidade de escrita dos *scriptores* do sertão de baixo, o contexto sócio-histórico e o perfil social vêm desempenhar certo efeito no uso da língua, principalmente nos desvios das normas de escrita compartilhadas pela comunidade de prática.

Por meio dos assentos de casamentos, houve a garantia da “recuperação do patrimônio cultural escrito” (CAMBRAIA, 2005, p.19) do sertão do Rio São Francisco, permitindo a possibilidade de análise da escrita em uso no século XVIII, no referido cenário, e apresentando amostras representativas de um processo histórico da língua

portuguesa. A análise do *corpus* diacrônico permite a caracterização do ato de criação, viabilizando descobertas por meio da peculiaridade dos diferentes punhos identificados, através dos aspectos físico-gráficos e gráficos e fonéticos, como o traçado das letras, as substituições de segmentos, a segmentação vocabular, o alicamento de vogais, as substituições consonantais, as formas etimologizadas, entre outros.

Para a coleta dos dados e a constituição da amostra a ser analisada, realizou-se a exploração das particularidades do suporte material do documento, a leitura e a transcrição de 231 assentos de casamentos a partir da reprodução fac-similar, pertencente ao Projeto³ de Pesquisa Educação Patrimonial: mapeando acervos históricos e culturais de Bom Jesus da Lapa, vinculado ao Departamento de Ciências Humanas e Tecnologias – *Campus XVII*, da Universidade do Estado da Bahia – UNEB. A fonte documental está alocada no acervo da Cúria Diocesana de Bom Jesus da Lapa (BA). Em seguida, foi realizado o mapeamento sistemático das unidades de análise, com o auxílio da ferramenta computacional AntConc (2011), e por último, o cotejamento com obras coetâneas, a exemplo de Bluteau (1728) e Feijó (1734).

A análise dos dados, embasada em dados empíricos, recorreu à abordagem quantitativa, utilizando os modelos estatísticos, ao mesmo tempo que se combinou com a abordagem qualitativa de cunho filológico e paleográfico, fornecendo indícios que documentam a habilidade dos *scriptores*. Sob a ótica da aplicação dos métodos da sociolinguística histórica, a interpretação dos materiais do passado deve ser refinada através do tempo e do espaço social e geográfico, imprimindo a capacidade de compreender a natureza e estrutura da língua como uma faculdade humana.

3 O GÊNERO ASSENTOS DE CASAMENTOS

Antes do século XVI, nos países da Europa, já havia a prática de registrar os batizados e casamentos. Após o Concílio de Trento (1545-1563), estas anotações tiveram o uso normatizado, reforçado pelas Constituições Primeiras do Arcebispado da Bahia (1719), que regulamentaram a obrigatoriedade dos livros paroquiais nas igrejas (cf. Título LXXIII, Art. 318, 1719, p. 130), seguindo o estilo de linguagem padrão oficial do período, sendo “exarados com toda a exação, clareza e aceyo⁴”.

Os assentos de casamentos, documentos selecionados para este estudo, constituem uma espécie documental que, segundo Bellotto (2002, p.48), se configura como um “documento diplomático testemunhal de assentamento ou comprobatório, horizontal. Registro de um compromisso passado em livro próprio (códice)”. Os livros

³ O Projeto de Pesquisa Educação Patrimonial: mapeando acervos históricos e culturais de Bom Jesus da Lapa, coordenado pela Prof.^a Ma. Ádma Bernardino Magalhães, Professora Auxiliar — DCHT *Campus XVII*, Universidade do Estado da Bahia (UNEB) —, iniciou suas atividades em 2015, com o objetivo de identificar e catalogar acervos documentais, saberes e práticas tradicionais e manifestações culturais de Bom Jesus da Lapa e também do Território de Identidade Velho Chico.

⁴ Cf. Livro de Registros de casamentos da Freguesia de Santo Antônio do Urubu de Baixo do Rio São Francisco (1773-1802, f. 37 r.).

paroquiais eram a única fonte dos mais variados registros da população, cabendo ao vigário as principais funções sociais, políticas e religiosas na freguesia.

Os arquivos paroquiais “preservaram escassa documentação para os séculos XVI e XVII, começando a ser mais expressivos a partir do século XVIII” (BACELLAR, 2008, p. 40). Por meio dos assentos em questão, houve a garantia do testemunho da história e da cultura da sociedade do Sertão do São Francisco do século XVIII, possibilitando diferentes focos de estudos, como a análise do uso da escrita, revelando os padrões característicos da língua, mediante as ocorrências ortográficas, o vocabulário, a historicidade do texto e as análises discursivas, entre outros aspectos.

O primeiro livro da Freguesia de Santo Antônio do Urubu de baixo do Rio São Francisco (1719-1753) consiste em um manuscrito de 138 fólios, que contém assentos de casamentos e óbitos. Os primeiros são os mais detalhados, com informações variadas, como regulamentam as *Constituições Primeiras do Arcebispado da Bahia* (1719, p.130):

Aos tantos de tal mez, de tal anno pela manhã, ou de tarde em tal Igreja de tal Cidade, Villa, Lugar, ou Freguezia, feitas as denunciações na fóрма do Sagrado Concilio Tridentino nesta Igreja, onde os contraentes são naturaes, e moradores, ou nesta, e tal, e taes Igrejas, onde N. contrahente é natural, ou foi, ou é assistente, ou morador, sem se descobrir impedimento, ou tendo sentença de dispensação no impedimento, que lhe sahio, como consta da certidão, ou certidões de banhos, que ficão em meu poder, e sentença que me apresentarão, ou sendo dispensados nas denunciações, ou diferidas para depois do Matrimonio por licença do Senhor Arcebispo, em presença de mim N. Vigario, Capellão, ou Coadjutor da dita Igreja, ou em presença de N. de licença minha, ou do Senhor Arcebispo, ou Provisor N., e sendo presentes por testemunhas N. e N. , pessoas conhecidas, (nomeando duas, ou três das que se acharão presentes) se casarão em face da Igreja solemnemente por palavras N. filho de N., e de N., natural, e morador de tal parte, e freguez de tal Igreja, com N. filha de N., ou viúva que ficou de N. natural, e morador de tal parte, e Freguezia desta, ou de tal Parochia: (e se logo lhe der as bençãos acrescentará) e logo lhe dei as bençãos conforme os ritos , e ceremonias da Santa Madre Igreja, do que tudo fiz este assento no mesmo dia, que por verdade assignei (Constituições Primeiras do Arcebispado da Bahia, 1719, p.130).

A organização textual dos assentos de casamentos está baseada na macroestrutura constituída pelas seguintes partes: data, local, nome dos contraentes, filiação e naturalidade, diligências, impedimentos (ou a falta deles, ou a sua dispensa), testemunhas, ato do casamento, registro e assinatura(s). As informações neles contidas são resultantes das “diligências necessárias” feitas pelos sacerdotes responsáveis pela lavratura dos registros.

Segundo as orientações das *Constituições* (1719), os assentos deveriam ser lavrados no ato do casamento, no entanto pode-se observar que muitos foram lançados no livro depois de algum tempo, especialmente os realizados nos atos de desobriga do vigário em visitação às diversas localidades da freguesia ou aqueles realizados por sacerdotes auxiliares nas capelas filiais, nas missões dos aldeamentos indígenas ou oratórios particulares. Nesse caso, os registros eram feitos em papéis avulsos ou ca-

ternos de anotações, o que, muitas vezes, comprometia a totalidade — ou até mesmo facilitava o extravio — dos registros.

A seguir, apresenta-se a transcrição de um assento de casamento referente ao *corpus* estudado, localizado no fôlio 35 v.

Transcrição

Aleixo | Gomes com | Agueda da | Costa

Aos dous dias domes de Mr.co doanno de mil Setecentos, e | trinta, equatro nesta Igreja Matris de Santo Ant.º Santo | Antonio do Urubû debxº do Rio de Sam Fr.co de manhãá, | ena mesma feitas as denunciaçõins dos Contrahentes na | forma do sagr. Conc. Trid. onde Sam moradores, Sem | Sedescubrir impedimento em prezença de mim o Vigr.º | Joam Gomes de Souza, Sendo presentes p testemunhas oad | judante Pedro Gomes de Abreu de Lima, eo Alfrz' | Pedro de Abreu de Lima pessoas Conhecidas, Se Caza | raõ em face de Igreja Solemnemente por pa | lavras Aleixo Gomes escravo do Coronel Alex.e | Gomes Ferraõ Castelbr.co Com Agueda da Costa | Crioulla forra viuva q ficou de Mathias de Oli | veyra, do que tudo fiz este assento no mesmo dia, | q p verdade meaSigney.

OVigr.º Joam Gomes deSouza

A transcrição do primeiro livro de registros paroquiais da Freguesia de Santo Antônio do Urubu de baixo do Rio São Francisco trouxe à tona costumes e fatos sócio-históricos do sertão do São Francisco, além da observação de fenômenos linguísticos, entre eles, os aspectos físico-gráficos e gráficos e fonéticos, peculiaridades do curso da escrita singular dos diferentes *scriptores*, ao longo do período em que o documento foi produzido.

4 O PERFIL SOCIAL DOS *SCRIPTORES*

A Freguesia de Santo Antônio do Urubu de baixo do Rio São Francisco foi fundada em 1718, desmembrada da Freguesia de Santo Antônio de Vila Nova Real do Rio São Francisco, que, por sua vez, foi fundada em 1679, localizada à margem direita do Rio São Francisco, pertencente ao Arcebispado da Bahia e à Capitania de Sergipe *del Rey*. Exercia uma estrutura organizacional política, administrativa e religiosa para as populações do sertão de baixo, constituída da igreja matriz de Santo Antônio, da capela filial dedicada à Nossa Senhora do Amparo, da igreja de Nossa Senhora do Desterro, na Fazenda Jaboatão, dos Padres da Companhia de Jesus, e da igreja da Missão de São Pedro do Porto da Folha dos capuchinhos italianos. Em 1757, a comunidade era composta por 1357 pessoas de comunhão, que participavam das celebrações litúrgicas, e pelos sacerdotes responsáveis pela orientação religiosa do rebanho e pela lavratura dos registros paroquiais.

Os *scriptores* dos assentos de casamentos que aparecem no primeiro livro (1719-1753) são vigários e coadjutores responsáveis pela freguesia, identificados na fonte documental, pertencentes a um grupo social com habilidade na escrita, do sexo masculino, de faixa etária adulta e que adotava um estilo de escrita formal, com uso

de fórmulas pré-estabelecidas. Mesmo sendo pessoas de alto nível socioeducativo, deixaram evidências de traços da língua falada ao longo do manuscrito, por exemplo, características fônicas refletidas na escrita (descubrir, dezaceti, Pernambuco, riu, gentiu, molher, natorais, orubu, despensa, deligencias, empedimento, imcomendado), além de ver uma mesma palavra sendo grafada de maneiras diferentes (por exemplo: manháá, manhá, manhá , manháa, manhâ, manhã, manham, minha, menhã), entre outros aspectos paleográficos e linguísticos.

A habilidade com a pena pode proporcionar investigações linguísticas, por meio da “mostra regular da variação nos símbolos ortográficos condicionada pelo contexto, da mesma forma que a língua falada apresenta variação alofônica” (ROMAINE, 1982, p.16). Os costumes da escrita dependem das práticas da comunidade de determinado período, estabelecida pela tradição diplomática, além de considerar os fatores sociais, textuais e estilísticos como condicionantes da variação na escrita e da atitude e hábitos dos *scriptores* dentro da comunidade de prática.

Esta última é formada por um conjunto de pessoas agregadas em razão do engajamento mútuo, que compartilham modos de fazer as coisas, modos de dizer, crenças, valores, relações de poder e que aprofundam o conhecimento interagindo com os comportamentos sociais de forma duradoura (MEYERHOFF, 2004 [2002]; ECKERT; MCCONNELL-GINET, 2010). Entre os *scriptores* do sertão de baixo do São Francisco, é possível perceber, através dos padrões comuns do modo de dizer e da forma de escrever as palavras e expressões características da tradição diplomática, o uso vernáculo da língua e sua variação de acordo com a prática social, afetando a prática linguística.

A investigação dos fatores sócio-históricos possibilita um delineamento do perfil sociocultural para adequar a realidade histórica, levando em consideração as relações entre a língua e o contexto sócio-histórico, por exemplo, os reflexos dos conflitos dos períodos históricos da ortografia do português e da influência renascentista, trazendo consequências para os padrões do uso da língua.

Em Silva (2011, p. 260), colhe-se a informação de que “os paradoxos vividos no século das luzes português devem indicar em que elementos a língua está se alterando por pressão não só do sistema interno, mas pela influência das ciências e artes [...]”. Ademais, esse período histórico consistia em recuperar o latim clássico e atribuir dignidade às línguas vernáculas, aplicando os mesmos ideais de correção (BORGES NETO, 2009, p. 43).

No século XVIII, agravou-se a polêmica entre a ortografia fonética e etimológica, produzida pelas tendências naturais e pelos hábitos fonéticos espontâneos. As línguas vernáculas eram influenciadas pelo reflexo das gramáticas do latim, pois não havia uma literatura canônica a ser usada como modelo de língua padrão, muito menos uma norma ortográfica estabelecida (BORGES NETO, 2009). Coutinho (1976) descreveu três períodos históricos da ortografia da língua portuguesa: o fonético (dos primeiros escritos até o século XVI), o pseudoetimológico (do século XVI até 1904) e

o simplificado (iniciado com a publicação de Gonçalves Viana, em 1904, e estendido até os nossos dias).

O autor afirma que, no período fonético, a ortografia portuguesa não era uniforme, pois buscava-se aproximar a oralidade da escrita. Williams (1975) também menciona que, nesse período, os escribas tentavam representar foneticamente os sons das palavras que escreviam, adaptando velhas grafias ou inventando novas, gerando inconsistências, a exemplo de representar o mesmo som de diferentes maneiras e diferentes sons da mesma maneira.

Acerca do período pseudoetimológico, Coutinho (1976) relata que, nos primeiros tempos, o pseudoetimologismo era marcado por uma grande influência fonética, até considerar a necessidade de maior conhecimento do latim — influenciado pelo movimento da Renascença. Caracterizava-se pela grande quantidade de complicações gráficas, devido à tendência etimológica.

Nesse período histórico, a etimologia dependia da imaginação de cada escritor e era muito comum o uso de consoantes geminadas, de encontros consonantais impróprios e substituições consonantais. Ao observar o uso da grafia variada nos assentos de casamentos, constata-se a presença de diversas palavras influenciadas por esses períodos históricos da língua portuguesa escrita.

O terceiro período, identificado por Coutinho (1976) como simplificado, é marcado pela publicação da *Ortografia Nacional*, em 1904, do foneticista luso Gonçalves Viana, na qual se apresentam os princípios em que se deveria basear a simplificação ortográfica, orientados pela pronúncia, sem descuidar da etimologia e do elemento histórico.

Embora haja escassas informações sobre os *scriptores* que desempenharam o seu trabalho pastoral na freguesia, a partir dos arquivos paroquiais da Cúria Diocesana de Bom Jesus da Lapa (BA), podem-se identificar alguns dados históricos indispensáveis para reconstituir o perfil social de cada um deles:

Quadro 1: Perfil social dos *scriptores*

<i>Scriptores</i>	Cargo/função	Ordem/ Formação	Tempo de atuação	Naturalidade	Filiação
Joam da Silva Ribeiro	Vigário colado	Presbítero do hábito de São Pedro	1719-1728	-	-
Joachim de Souza	Vigário encomendado	-	1728-1729	-	-
Gonçallo Vellozo de Souza	Coadjutor	Licenciado	1729-1735	Freguesia de Santo Antônio do Urubu de baixo do Rio São Francisco	Capitão Manoel Gomes de Souza e Catarina Vellozo

João Gomes de Souza	Vigário colado	Licenciado	1733-1746	Freguesia de Santo Antônio do Urubu de baixo do Rio São Francisco	Capitão- mor Dom Pedro de Souza e Dona Bernarda Maria da Silva
José da Silva Gomes	Vigário encomendado	Presbítero do hábito de São Pedro	1746-1751	—	—
Manoel de Oliveira Cabo	Coadjutor	Licenciado	1749-1757	—	—
Joachim Marques de Oliveira	Vigário encomendado	Licenciado Presbítero do hábito de São Pedro	1751-1753	Villa do Penedo - Pernambuco	—

Fonte: elaboração própria.

A investigação nos registros paroquiais do sertão de baixo certamente permitiu uma aproximação do perfil social dos *scriptores* que atuaram durante uma fase de contradições ortográficas da língua portuguesa. No caso da estratificação social do século XVIII, é preciso levar em consideração a centralização da educação no projeto colonial, como menciona Anísio Teixeira (1989, p. 62):

Sobre esta sociedade pairava a elite, cuja cultura intelectual era claramente a da Idade Média e comum com a da Metrópole, compreendendo a formação teológica e de letras latinas do clero e dos leigos, canonistas e médicos, que se iriam graduar na Universidade de Coimbra, que era tanto a universidade da Metrópole quanto da Colônia.

Os *scriptores* da Freguesia de Santo Antônio do Urubu de baixo do Rio São Francisco eram profissionais que atendiam às demandas da paróquia, exercendo seus trabalhos pastorais conforme as exigências do Concílio Tridentino⁵, que visava à formação de um clero ilustrado e santo, composto por uma verdadeira elite cultural religiosa, capaz de promover a instrução do povo pela catequese e pela escolarização. Acerca do clero no século XVIII, Silva (2011) ressalta que ele tinha direito a foro privativo, constituindo grande força organizacional sedimentada no direito canônico e nas regras das Ordens.

No caso específico dos *scriptores* do sertão de baixo, não foi possível localizar, até o momento, informações sobre como se teria dado a formação religiosa. Contudo, é possível conjecturar que os sacerdotes que ostentavam o título de “licenciado” foram buscar a formação acadêmica na Universidade de Coimbra, uma vez que “até os começos do século XIX, a universidade do Brasil foi a Universidade de Coimbra, onde iam estudar os brasileiros, depois dos cursos no Brasil nos reais colégios dos jesuítas” (TEIXEIRA, 1989, p. 65). A história social e a sociolinguística histórica devem auxiliar

⁵ O Sacrosanto, e Ecumenico Concílio de Trento em Latim, e Portuguez. Lisboa: Francisco Luiz Ameno, 1781.

a construção das circunstâncias do passado a partir da informação procedente de documentos e dados coetâneos (CONDE-SILVESTRE, 2007), representando o melhor testemunho do passado.

De acordo com Cambraia (2005, p.20) “os textos escritos utilizados como *corpus* para o conhecimento da língua só poderão ser adotados como fonte de dados se espelharem o emprego efetivo da língua, ainda que apenas na sua modalidade escrita”. Dessa forma, contribuem para a reconstrução do contexto social, pois figuram como importantes indícios para compreender os conflitos entre inovação e conservação pertinentes a cada fenômeno linguístico analisado.

No caso dos assentos de casamentos do sertão de baixo, apesar das orientações estabelecidas pelas *Constituições Primeiras do Arcebispado da Bahia* (1719), a frequência de uma grafia variável e com traços da modalidade oral, bem como a escolha de determinados signos, a exemplo do uso de abreviaturas, não recomendado pelas *Constituições* (1719), mostram um distanciamento do modelo a ser seguido e a partilha de atitudes e hábitos dentro da comunidade de prática.

Os punhos, identificados no documento, apresentam determinado grau de habilidade de escrita, revelando uma prática familiarizada com as normas diplomáticas no processo de elaboração dos registros paroquiais, transparecendo aspectos físico-gráficos característicos do contexto histórico e também específicos de cada *scriptor*. Sobre alguns desses aspectos é que trata a próxima seção.

5 RESULTADOS E DISCUSSÃO: ASPECTOS FÍSICO-GRÁFICOS

No século XVIII, a escrita humanística era utilizada na tradição documental em Portugal, e por conseguinte, nas colônias. Por sua vez, esse modelo de escrita apresentava características peculiares, como a clareza (ACIOLI, 1994). Com base nos trabalhos de Acioli (1994), Cambraia (2005) e Berwanger e Leal (2008), busca-se entender de um modo mais claro as características específicas do *uso scribendi* de cada *scriptor*. Para isso, foi necessário investigar as habilidades de escrita, bem como o contexto sócio-histórico de produção do manuscrito analisado.

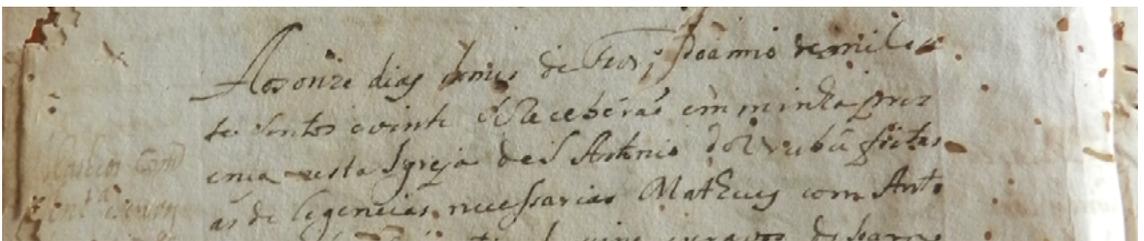
Mediante a isso, os critérios adotados por Cambraia (2005) foram imprescindíveis para conhecer a habilidade de escrita, a partir da análise das seguintes informações:

[...] a morfologia das letras (sua forma), o seu traçado ou *ductus* (ordem de sucessão e sentido de traços de uma letra), o ângulo (relação entre os traços verticais das letras e a pauta horizontal da escrita), o módulo (dimensão das letras em termos de pauta) e o peso (relação entre os traços finos e grossos das letras; descrição sucinta do sistema dos sinais abreviativos empregados na referida escrita; descrição dos outros elementos não alfabéticos existentes e de seu valor geral: números, diacríticos, sinais de pontuação, separação vocabular intralinear, paragrafação, etc.; descrição de pontos de dificuldade na leitura e as soluções adotadas) (CAMBRAIA, 2005, p. 24).

Ancoradas nessas orientações, as particularidades a serem investigadas na prática da escrita — embora nem todos os aspectos sejam levantados no presente estudo — possibilitam novos olhares sobre os sujeitos cujas mãos escrevem os textos, ressaltando a importância do documento histórico para a reconstrução do passado, das mudanças e inovações de uma língua. Higounet (2003) diz que a escrita faz parte da civilização e poderia servir para defini-la, e ainda acrescenta que a primeira é a expressão gráfica da linguagem.

Por meio dos caminhos trilhados nessa empreitada filológica-paleográfica, esquadriharam-se os traços deixados pelos diferentes punhos identificados nos assentos, que permitiram transparecer características de mãos hábeis, como se pode verificar a seguir:

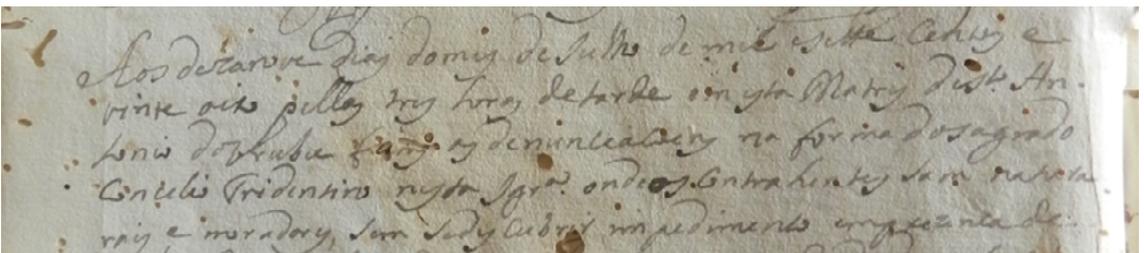
Figura 1: *Scriptor* Joam da Silva Ribeiro — trecho do fólio 2v



Fonte: Livro de Registros de casamentos da Freguesia de Santo Antônio do Urubu de Baixo (1719-1753).

O primeiro *scriptor* identificado foi o padre Joam da Silva Ribeiro, vigário colado. Sua escrita respeita às margens do fólio; as letras cursivas apresentam regularidade e homogeneidade, são arredondadas, de módulo pequeno, levemente inclinadas para a direita e com traçados firmes. O perfil do traçado das letras desse *scriptor* apresenta as tendências registradas por Higounet (2003) sobre a escrita humanística cursiva, descrita como erudita, suave, geralmente traçada com penas pontudas, fortemente inclinada para a direita, clara e inteligível.

Figura 2: *Scriptor* Joaquim de Souza — trecho do fólio 11v

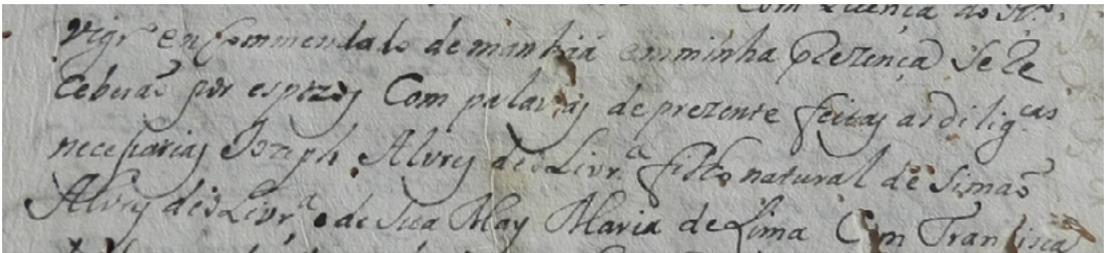


Fonte: Livro de Registros de casamentos da Freguesia de Santo Antônio do Urubu de Baixo (1719-1753).

O segundo *scriptor* diz respeito ao padre Joaquim de Souza, vigário encomendado, cuja escrita cursiva oferece uma série de traços recorrentes por mãos hábeis e que,

portanto, podem ser usados para sua caracterização: escrita regular, módulos pequenos, traçado firme, claro e levemente inclinado para a direita. Conforme esclarecem Berwanger e Leal (2008), a escrita humanística cursiva liga as letras umas às outras, devido ao desejo de escrever mais depressa e à preguiça de se levantar a mão. A escrita respeita as margens dos fólhos, não há rasuras e nem segmentação vocabular. O uso de abreviaturas é frequente, mas não exagerado.

Figura 3: *Scriptor* Gonçallo Vellozo de Souza — trecho do fólio 13v



Fonte: Livro de Registros de casamentos da Freguesia de Santo Antônio do Urubu de Baixo (1719-1753).

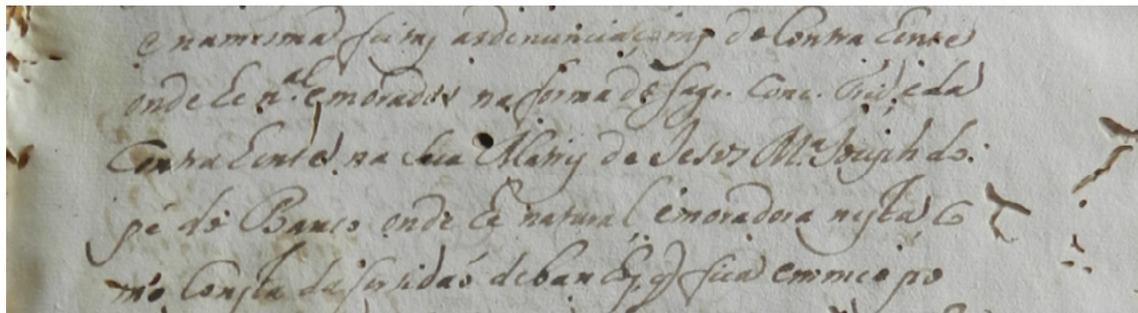
O terceiro *scriptor* refere-se ao padre licenciado Gonçallo Vellozo de Souza, coadjutor, que apresenta uma escrita com regularidade e homogeneidade; as letras são levemente arredondadas, os módulos são pequenos e, mesmo as letras que apresentam hastes, estão dentro da dimensão correspondente às pautas do fólio. Quanto ao ângulo do traçado, ele é perpendicular aos traços verticais e à horizontalidade da pauta. Os traçados são firmes, e o peso é fino e claro. As letras maiúsculas apresentam-se quase ornamentadas.

Acioli (1994) destaca que, nos documentos brasileiros dos séculos XVIII e XIX, era comum utilizar, como instrumento de escrita, a pena de ave e a tinta ferrogálica. É o que se pode constatar nos assentos do sertão de baixo do século XVIII, de acordo com a análise codicológica realizada por Assis, Silva e Pacheco (2021a):

O suporte é composto de um papel de textura fina e de alta qualidade, sem pauta, apenas com piques equidistantes nas margens direita e esquerda para guiar os traços da justificação, provavelmente feito por um objeto que lembra um bico de pena, um compasso ou estilete. [...] O livro é escrito com tinta ferrogálica, de coloração amarronzada, com variações mais escuras ou mais claras ao longo do documento e em alguns casos, a coloração aproxima-se do preto. Em alguns fólhos, há manchas e corrosão deixadas pela tinta, onde a concentração foi maior, além da migração que transferiu as imagens para o verso e para outros fólhos do documento (ASSIS; SILVA; PACHECO, 2021a, p. 233-234).

O exame empreendido pelos autores baseou-se na proposta de Cambraia (2005, p.25), que considera “a decifração e a reprodução [...] realizadas com mais segurança e propriedade quando se tem consciência de como eram produzidos os documentos, em que classes se distribuíam e como se estruturaram internamente”, a fim de conhecer a materialidade do documento em questão, revelando suas características culturais de produção.

Figura 4: *Scriptor* Joam Gomes de Souza — trecho do fólio 39r

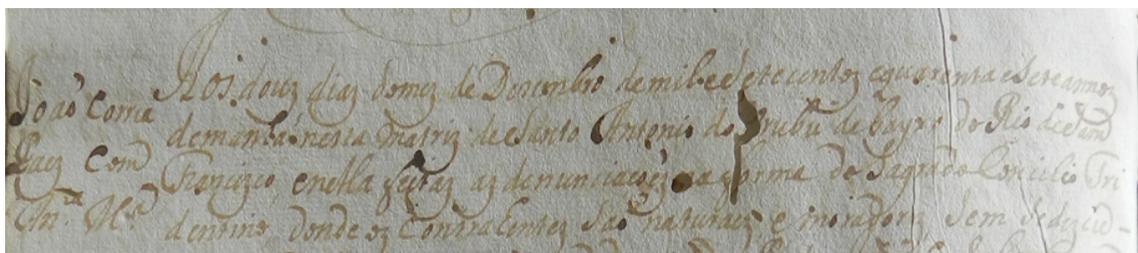


Fonte: Livro de Registros de casamentos da Freguesia de Santo Antônio do Urubu de Baixo (1719-1753).

Na sequência, o documento revela o quarto *scriptor*. O padre licenciado Joam Gomes de Souza, vigário colado, registrou assentos de casamentos de 1733 a 1746. O tipo de escrita do seu texto é cursivo, cujas letras são arredondadas, de módulo pequeno, levemente inclinadas para a direita e de fácil leitura. Os traços são leves, regulares, firmes e legíveis. Não há sinais de borrões ou rasuras. O *scriptor* respeitou o espaço em relação às margens, e sua escrita apresenta poucas abreviaturas (debx^o, R.do, L.do, q, Vigr.^o, p, S.to, Ant.^o, Pr.^a, Fr.co, freg.^a, N. Snr.^a).

Em relação ao uso de abreviaturas, Assis, Silva e Pacheco (2021a) listaram a frequência dos tipos de abreviaturas utilizadas pelos *scriptores*, a saber: letras sobrescritas, siglas, suspensão ou apócope e contração ou síncope. Os autores concluíram que o uso das abreviaturas, nos registros paroquiais do sertão de baixo do século XVIII, refletia o princípio da economia (ACIOLI, 1994) e a influência da tradição documental da época, embora o emprego de formas abreviadas fosse contrário às normas das *Constituições Primeiras do Arcebispado da Bahia* (1719), como já dito.

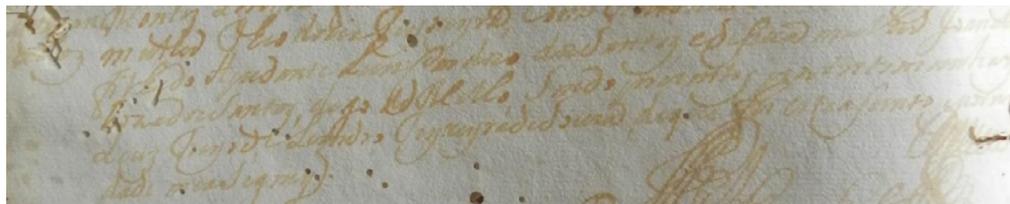
Figura 5: *Scriptor* José da Silva Gomes — trecho do fólio 67v



Fonte: Livro de Registros de casamentos da Freguesia de Santo Antônio do Urubu de Baixo (1719-1753).

O quinto *scriptor* refere-se ao padre José da Silva Gomes, vigário encomendado. É possível dizer que o traçado de suas letras é firme e, com respeito ao uso das margens do fólio, o módulo utilizado é pequeno, e o peso da escrita, leve; apresenta um uso reduzido de palavras abreviadas — apenas três: Fran.^o, q, S.^a — e não há variações relacionadas à segmentação vocabular.

Figura 6: Manoel de Oliveira Cabo — trecho do fólio 74v.



Fonte: Livro de Registros de casamentos da Freguesia de Santo Antônio do Urubu de Baixo (1719-1753).

O sexto *scriptor* remete ao padre licenciado Manoel de Oliveira Cabo, coadjutor. A tendência do traçado de sua escrita é firme, e o uso dos módulos é pequeno, respeitando o uso das margens do fólio. O peso da escrita é leve, ao qual se soma o traçado seguro, claro e inteligível. O texto é produzido com ausência de abreviaturas e rasuras, e não há incidência de segmentação vocabular.

Figura 7: Joaquim Marques de Oliveira — trecho do fólio 79r



Fonte: Livro de Registros de casamentos da Freguesia de Santo Antônio do Urubu de Baixo (1719-1753).

O sétimo *scriptor* é o padre doutor Joaquim Marques de Oliveira, vigário encomendado. Sua escrita cursiva apresenta regularidade, e as letras são levemente inclinadas à direita. Há homogeneidade no tamanho das letras, e a escrita das maiúsculas é bem elaborada. O texto é alinhado e margeado. Há poucas abreviaturas (M.^a, Roiz', S., Vigr.^o, Olivr.^a).

No que se refere à pontuação, esse aspecto tem características semelhantes a todos os *scriptores*. Os sinais de pontuação vírgula, ponto e vírgula e ponto são os mais frequentes. O ponto (.) foi utilizado como sinal abreviativo e ponto final; a vírgula (,) foi empregada, na maioria dos casos, precedendo a conjunção <e>; e o ponto e vírgula (;) substituiu o emprego da vírgula. Os demais sinais de pontuação foram utilizados com menos frequência, a saber: os traços duplos (=) e o hífen (-) foram utilizados para separar as palavras no final da linha; a barra vertical (|), foi usada como parênteses; e o apóstrofo ('), como sinal abreviativo.

A partir da prática de escrita e da habilidade dos *scriptores* do sertão de baixo, foi possível verificar os traços gráficos característicos de cada um deles, qualificando-os como hábeis na escrita, embora, durante esta análise, foi comum encontrar palavras apresentando variações gráficas, muito comuns no período.

A próxima seção aborda diferentes aspectos gráficos e fonéticos encontrados nos assentos de casamentos.

6 RESULTADOS E DISCUSSÃO: ASPECTOS GRÁFICOS E FONÉTICOS

Os aspectos gráficos e fonéticos encontrados nos registros paroquiais do sertão de baixo do século XVIII caracterizam-se por hesitações históricas, relativas aos períodos da ortografia da língua portuguesa, ainda que produzidos por punhos treinados para a prática da escrita. As características do nível gráfico e fonético, evidenciadas no manuscrito, descrevem o estado da língua portuguesa marcado pelo contexto sócio-histórico em que o documento foi produzido.

A seguir, são apresentados alguns desses aspectos:

6.1 SISTEMA VOCÁLICO

No que diz respeito ao sistema vocálico, avista-se a possibilidade de compreender a instabilidade do emprego das vogais através dos processos fonológicos recorrentes na prática de escrita dos *scriptores* do sertão de baixo. Para a descrição do sistema vocálico, optou-se por analisar as amostras expostas na tabela 1, a seguir:

Tabela 1: Sistema vocálico

	Ocorrências	Ocorrências
As grafias <e>, <i>		
Posição não acentuada final	dezaceti (2)	dezacete (2)
Em posição inicial absoluta	imcomendado (1) incomendado (1) empedimento (8)	encomendado (43) - impedimento (130)
Em posição pretônica	despensa (1) deligencias (17) despensados (2) despensado (1)	dispensa (1) diligencias (9) dispensados (7) dispensados (1)
As grafias <o>, <u>		
Posição não acentuada final	Pernambucu (3)	Pernambuco (31)
Em posição inicial absoluta	Orubû (5) Orubu (3)	Urubû (199) Urubu (129)
Em posição pretônica	descubrir (70) Sedescubrir (44)	descobrir (51) Sedescobrir (40)

Fonte: elaboração própria.

Como se pode observar, no caso de dezaceti / dezacete, a variação ocorre na posição não acentuada final, grafada com <i>, ilustrada pelo exemplo: “Aos dezaceti dias domes de Fevr.º demile Sete cen | tos edous [...] (f. 4r). Conforme aponta Mattos e

Silva (2006, p. 58-59) “nos exames das grafias a partir do século XIII, observa-se, esporadicamente, a variação <i> e <e> final, convencionada, posteriormente, na ortografia do português [...]”. O *corpus* analisado contém a variação do <e> e <i> final apenas em dezaceti / dezacete, com 02 ocorrências em cada forma.

Em posição inicial absoluta, o uso da vogal alta <i> no lugar da vogal média <e>, diante de nasais, foi encontrado em “Imcomendado / incomendado/ encomendado”, caracterizando um alteamento vocálico. Já as ocorrências empedimento / impedimento, o uso da vogal média <e> no lugar da vogal alta <i> representa um caso de abaixamento vocálico. Mattos e Silva (2006, p. 59) ressalta que a variação <e> / <i> é usual em sílabas iniciais, nas quais a vogal é travada por nasal ou sibilante.

Os vocábulos Despensa / dispensa / deligencias / diligencias / despensados / dispensados / despensado / dispensados apresentam variações da grafia em relação ao uso da vogal média <e> no lugar da vogal alta <i>. De acordo com Williams (1973, p. 17-18) a frequente modificação do <e> em <i> em documentos do latim medieval é puramente ortográfica e provavelmente decorre de confusão entre as duas letras, por causa da impressão de que ambas representavam o mesmo som.

No caso dos exemplos apresentados, especialmente quando a vogal é travada por uma sibilante, eles parecem motivados por hipercorreção, configurando casos de abaixamento vocálico. Bluteau (1728, p. 443) e Madureira Feijó (1734, p. 79) registram, em suas obras, as formas “dispensa”, “dispensação” e “diligencia”.

No exemplo “Pernambucu”, encontrado nos dados analisados, Paul Teyssier (2007) declara que a passagem do <o> final para <u>, observada em documentos da primeira metade do século XVIII, pode ser encontrada em épocas anteriores. Segundo o autor, a realização de <u> onde deveria ser escrito <o> final era comum no território português da Europa e também no Brasil.

Quanto à variação gráfica entre <o> e <u> da palavra Urubu, que aparece grafada como Urubu, Orubû, Orubu, Urubú, Mattos e Silva (2006, p. 60) esclarece que a vogal que tem representação gráfica variável está seguida de vogal alta na sílaba vizinha, motivando um abaixamento. Entretanto, ao que tudo indica, seria resultado de uma hipercorreção.

Em relação ao verbo descobrir / descobrir, a variação entre <o> e <u>, em posição pretônica, é motivada pela vogal alta <i> na sílaba seguinte, que induz ao alçamento.

Coutinho (1976) aponta que não havia um padrão uniforme na transcrição das palavras, e um mesmo vocábulo aparecia em um documento grafado de modo diferente. O autor defende que, embora houvesse a pequena influência do latim, a tendência fonética vigorava no sistema em uso.

Ao analisar o sistema vocálico, observa-se como os manuscritos produzidos em um momento de conflitos históricos da ortografia da língua portuguesa mantêm pontos que se aproximam da oralidade representada na escrita. Além disso, foi possível perceber que a falta de normatização igualmente contribuiu para as oscilações na escrita. Apesar de os *scriptores* serem homens letrados e seguirem um modelo pré-es-

tabelecido de escrita, foram influenciados pelo contexto sócio-histórico e cultural no qual o documento foi elaborado.

6.2 SISTEMA CONSONANTAL

No tocante ao emprego dos usos gráficos do sistema consonantal, há fenômenos registrados na escrita devido à falta de normatização gráfica, revelando características específicas de cada um dos períodos históricos da ortografia da língua portuguesa. Conforme enfatizado por Mattos e Silva (2006), sendo a documentação escrita uma representação convencional da fala, os documentos refletem conclusões até certo ponto seguras no nível fônico-mórfico, já que a falta de normatização ortográfica favorece a variação na escrita, apresentando indícios para alguma percepção da voz.

A seguir, apresentam-se alguns dos fenômenos investigados.

6.2.1 Emprego do <h>

O tom etimologizante do emprego do <h> em documentos do século XVIII têm, como Silva (2011) apontou, uma influência do Renascimento, tentando aproximar o vernáculo da língua latina. Segundo o autor, a influência não seguiu o rigor etimológico, sendo constatada, muitas vezes, em formas sem correspondência com o latim.

Conforme atestam os exemplos documentados por Silva (2011, p. 269), o uso do <h> estaria registrado em três situações: influência etimológica (“reprehensaõ”, “reprehendia”); forma de marcar o hiato (“distrahida no interior”, “esteja distrahida”); e casos especiais (“fazendo que hum, e outro seja”, “he certo que e huns, e outros”).

Considerando os assentos de casamentos escritos no século XVIII, semelhantes aos achados de Silva (2011) e de Assis, Silva e Pacheco (2021b), observa-se a incidência de três situações do emprego do <h>, a saber: a influência etimológica (horas, homem, hábito, parochia, parochial, parochio); a pseudoetimológica (honde, hú, hum, huá, he, sachristaõ); e a forma de marcar o hiato (contrahentes, sahio, Bahya, Carahibas, Sucurihû, Arahujo), resultantes das tendências de escrita em que a influência da tradição latina vigorava nas grafias etimológicas presentes na natureza conservadora dos registros paroquiais (ASSIS; SILVA; PACHECO, 2021b).

Nesse sentido, na análise da grafia do emprego do <h>, as tendências do uso antes de vogais iniciais, em determinadas situações, indicava-se o hiato com a vogal final da palavra precedente, como em “na | freg.^a de Jezus M.^a Jozê do Pe do banco **honde** he | natural eaSistente” (f. 64r.). Em outras situações, a letra <h> foi usada para marcar o hiato entre duas vogais diferentes, como em **contrahente**, **Arahujo** e, no que diz respeito ao uso na posição inicial dos numerais “um” e “uma”, tanto Bluteau (1728, p. 686) quanto Madureira Feijó (1734, p. 71) registram, em suas obras, a grafia hum e huá.

A terceira pessoa do singular do presente do indicativo do verbo ser, “he”, é explicada por Madureira Feijó (1734, p. 70) como forma de diferenciar o “e” conjunção do “he” terceira pessoa do verbo ser no português. Coutinho (1976) assinala que, no

período arcaico da língua portuguesa, o <h> se apresentava no início da palavra por analogia ao latim, estendendo-se a outros vocábulos sem justificativa etimológica.

As variações no emprego do <h> estão inclinadas à atuação da latinização, por influência do Renascimento e, como já dito, pela ausência de normatização ortográfica da língua portuguesa no referido período, o que intensificava a oscilação gráfica no momento da escrita.

6.2.2 Emprego do <y> em sequências vocálicas orais

A variação da grafia dos ditongos <ai>, <ay>; <ei>, <ey>; <oi>, <ou> e <ui>,<uy> é uma notável característica da escrita dos *scriptores* do sertão de baixo e constitui um comportamento linguístico exagerado na busca do ideal de perfeição da língua, assim como constatado por Silva (2011, p. 271) nos exemplos “mayormente”, “leys”, “não foy com”, “tornemos a falar do muy forte gênero de Monges”, “A terceira maneira de monges he muy abominável”, retirados do *Livro das Monjas*, do Padre Gabriel Talbot (século XVIII).

O uso do <y> nos ditongos era prática comum aos escritores quinhentistas e seiscentistas (ALI, 1964) e, de acordo com Madureira Feijó (1734, p. 24), tanto na pronúncia quanto na escrita, não havia diferença alguma, a exemplo de “Moreira, Pereira, Teixeira e Moreyra, Pereyra, Teyxeira”, pois o som é o mesmo. O autor acrescenta, ainda, que o uso frequente do <y> se deve à prática que o introduziu e, nos ditongos, a utilização do <y> ou <i> é equivalente.

A tabela 2 apresenta os resultados, listando os dados extraídos dos assentos de casamentos do sertão de baixo (1719-1753):

Tabela 2: Emprego do <y> em sequências vocálicas orais

Ditongo ay / ai	
bayxo (85)	baixo (26)
Mayo (23)	Maio (01)
Payva (03)	Paiva (03)
Cayetano (09)	Caetano (16)
Ditongo ey / ei	
Oliveyra (21)	Oliveira (11)
Frey (10)	Frei (02)
Freyre (03)	Freire (15)
Fevereyro (16)	Fevereiro (10)
Solteyro (04)	Solteiro (14)
Primeyro (03)	Primeiro (12)

Ditongo iy	
Miyrim (01)	Mirim (03)
Ditongo oy / oi	
foy (03)	foi (19)
Oitubro (07)	Oytubro (01)
	Outubro (15)

Fonte: elaboração própria.

Os resultados apresentam tendências registradas na tradição documental, a partir do acesso aos textos sagrados, através da iniciação à leitura e à escrita do latim nos cursos de formação dos clérigos. Nota-se que os *scriptores* seguiam as tendências da escrita do seu tempo, e as variações nas sequências vocálicas orais mostram a influência do “resultado da aplicação da teoria gramatical latina aos dados das línguas vernáculas” (BORGES NETO, 2009, p. 44), revelando o desejo de erudição e, muitas vezes, produzindo ultracorreção.

6.2.3 Uso de formas etimologizadas

O Renascimento marca a volta dos clássicos greco-latinos, influenciando os *scriptores* quanto ao uso da grafia etimológica de forma exacerbada, a fim de remeter a uma erudição e à busca de perfeição da língua. Coutinho (1976) pondera sobre o grande influxo de vocábulos novos, além das formas vulgares, revestidas de aspecto gráfico alatinado.

As motivações históricas relacionadas ao período pseudoetimológico são caracterizadas pelo “emprego de consoantes geminadas e insonoras, de grupos consonantais impropriamente chamados gregos, de letras como o y, k e w, sempre que ocorriam nas palavras originárias” (COUTINHO, 1976, p. 72). Por conseguinte, refletiam no uso das consoantes presentes nos assentos de casamento do sertão de baixo.

De acordo com os fenômenos registrados, quanto ao emprego das consoantes, os segmentos foram agrupados em três categorias diferentes: consoantes geminadas, encontros consonantais impróprios e substituições consonantais, conforme observado no quadro 2 a seguir:

Quadro 2: Formas etimologizadas

consoantes geminadas	occidental, Affonsequa, capella, villa, commissario, anno, Joanna, cappitam, supposto, capittam, rittos, rritos
encontros consonantais impróprios	Baptista, septembro, acto, Victoria, asigney, Ignacio, solemnemente, Martha, christhans, Jozeph, parochia, Marhcos
substituições consonantais	Asumpsão, denunciassõins, prezenssa, fasse, sinco, serimonias, asento, acento, assistente, crus, Felis, fregezia, certaõ

Fonte: elaboração própria (grifo nosso).

Sobre as consoantes geminadas, Madureira Feijó (1734, p. 32) relata a dificuldade das regras para dobrá-las no meio das palavras, originada da pronúncia, uma vez que o som é o mesmo. O autor dá exemplos de palavras latinas que, escritas com as letras dobradas ou não, apresentam o mesmo som: *affinitas*, *aggravo*, *abbrevio*, *fallo*, *pello*, *tollo*. Madureira Feijó (1734, p. 32) ainda revela que as regras são confusas ou incertas, e a regra geral seria observar as palavras latinas e escrever, de forma semelhante, as palavras portuguesas derivadas.

Coutinho (1976) aponta que o uso das consoantes geminadas latinas, no interior das palavras em português, é reduzido a consoantes simples, sendo que, no latim vulgar, essa simplificação já havia ocorrido. Segundo Said Ali (1964, p. 44), o uso das consoantes geminadas era simplesmente “fantasia e capricho” dos escritores. No que diz respeito aos *scriptores* do sertão de baixo, pareceu conveniente grafar as consoantes de forma geminada e utilizar encontros consonantais impróprios com o intuito de dar aos vocábulos aspecto mais elegante, deparando-se frequentemente com a ultracorreção.

Nota-se, ainda, que, entre todos os *scriptores*, é possível encontrar substituições consonantais nos vocábulos grafados com <s>, <c> e <z>, gerando incerteza quanto ao uso desses grafemas, possivelmente influenciados pela oralidade. Optava-se, então, pelo uso da representação gráfica que não provocasse dúvidas na pronúncia.

Ao analisar a representação gráfica das consoantes presentes no manuscrito, constata-se que os *scriptores* procuravam seguir, à sua maneira, o princípio etimológico, mostrando preferência pelas formas etimologizadas, o que refletia a admiração pela corrente renascentista em voga. No entanto, o uso de substituições consonantais indica que havia também a influência da oralidade, além da ausência de normatização ortográfica da língua portuguesa no período estudado, que levava à hesitação gráfica no momento da escrita.

6.2.4 As grafias do <aõ> e do <am>

No que diz respeito ao uso do ditongo <aõ> nos assentos de casamentos da Freguesia de Santo Antônio do Urubu de baixo do Rio São Francisco, verifica-se que o diacrítico, marca da nasalização, ora aparece na semivogal, ora é grafado com <-am>.

É importante mencionar que, na análise empreendida por Silva (2011, p. 268), o ditongo grafado <aõ> aparece nas formas do presente do indicativo: “se ajuntar ao coro das que cantaõ”; “saõ os do presente”; “declaraõ o reconhecimento”; “quando entraõ”. O Padre Gabriel Talbot, no *Livro das Monjas*, segue a tendência de marcar a nasalidade no segundo elemento do ditongo. Silva (2011, p. 268) encontra, ainda, construções em que o presente do indicativo é grafado com <-am> como em: “com que falam”.

No estudo atual, verifica-se a falta de uniformidade ortográfica também em relação ao uso dos sinais diacríticos, refletindo a hesitação na escolha, no momento da escrita. Em estudo recente sobre os sinais diacríticos agudo, circunflexo e til em assentos de casamentos do sertão do São Francisco (1719-1753), Assis e Silva (2021), registraram os seguintes resultados no que se refere ao emprego do <aõ> e do <am>:

Tabela 3: Vocábulos com <aõ> e <am>

Vocábulo com <aõ>	Ocorrências	Vocábulo com <am>	Ocorrências
Joaõ	82	Joam	222
Saõ	69	Sam	225
São	3		
Capitaõ	24	Capitam	4
Cappitaõ	1	Cappitam	15
		Cappittam	1
missaõ	20	Missam	9
		Misam	1
Sebastiaõ	5	Sebastiam	7
Aragaõ	1	Aragam	3
certidão	25	Certidam	4
Certidão	8		
Sertidaõ	3		
Capelaõ	1	Capellam	2
Leitaõ	2	Leitam	2
Cristovaõ	2	Christovam	1
Conceiçaõ	20	Conseisam	1
Conceyção	1		
Manhaã	15	Manham	1
Manhã	8		
minha	1		
Povoação	1	Povoaçam	1
Simaõ	4	Simam	1
Constituiçaõ	2	Constituiçam	1
Vizitaçaõ	1	Vizitaçam	1

Fonte: Assis e Silva (2021, p. 383).

Cabe destacar que o uso do diacrítico til sobre a segunda vogal do ditongo nasal se alterna ao emprego da consoante <m> para representar a nasalidade. Nesse caso, é importante observar como os *scriptores* refletem, nos textos, uma oscilação de escrita, em grande parte condicionada pela tradição documental e pelo estilo adotado pelo próprio *scriptor*.

6.2.5 Ausência de contração em “em o” e “em a”

Em estudo realizado por Silva (2011, p. 272), sobre fenômenos linguísticos no *Livro das Monjas* (século XVIII), de autoria do Padre Gabriel Talbot, foi encontrada a ausência de contração em “em o” e “em a”, em construções do tipo: “aprenderão **em a** companhia de muitos a pelear”; “Da reverencia, que se dever ter **em a** Oração” e “bem ensinados **em o** Mosteiro”.

O autor revisitou alguns estudiosos que mostraram o uso comum, na fase arcaica da língua portuguesa, da contração da preposição “em” com os artigos definidos. Entre eles, destacam-se Leite de Vasconcelos (1922), Mário Pereira de Souza Lima (1937), D. Carolina Michaelis (1946), Marques da Cruz (1954) e Mário Barreto (1986).

A recusa em fazer a contração de “em o” e “em a”, nos assentos de casamentos do sertão de baixo do século XVIII, aparece em 16 ocorrências deixadas por quatro diferentes *scriptores*, expostas no Quadro 3 a seguir:

Quadro 3: ausência de contração em “em o” e “em a”

Contexto	Ano / fólio	Scriptor
pello meio dia em a igreia da Missão desam Pedro	1728, 11 v.	Joachim de Souza
pela manhã em a Igreja de Nossa Senhora daConceição	1743, 50 r.	Joam Gomes de Souza
em a Missão deSam Pedro doPorto da folha	1747, 67 v.	José da Silva Gomes
em a Capella de Nossa Senhora da Conceição do Urubû Miyrim	1748, 70 v.	José da Silva Gomes
morador em o Citio da Borda da Mata desta freg. ^a	1730, 13 v	Gonçallo Vellozo de Souza

Fonte: elaboração própria (grifo nosso).

Para explicar o fenômeno em discussão, a hipótese proposta foi aventada por Barreto (1986) — tal qual o fez Silva (2011) —, que esclarece que a recusa da contração poderia estar relacionada aos cacófatos. Nas palavras de Barreto (1986), “em alguns casos a construção contrata poderia produzir uma série de sílabas e iniciais repetidas, enfraquecendo, assim, a beleza do estilo” (apud SILVA, 2011, p. 273).

Em relação às ocorrências retiradas dos assentos de casamentos, assim como nos achados de Silva (2011), não se percebe a possibilidade de cacófatos, tampouco de casos em que a contração se daria com o artigo determinante do sujeito, configurando um preciosismo desnecessário dos *scriptores*.

7 CONSIDERAÇÕES FINAIS

A partir de um *corpus* de 231 assentos de casamentos, foram investigados os usos físico-gráficos e gráficos e fonéticos de *scriptores* do baixo sertão do Rio São Francisco do século XVIII. Para o presente estudo, levaram-se em consideração as práticas de escrita e os dados empíricos da fonte documental, com o intuito de evidenciar a

contextualização de espaço e tempo do documento histórico, bem como de ampliar as possibilidades de reconstrução da história social da língua, estabelecendo uma possível correlação entre os dados linguísticos do passado.

Traçar o perfil sociolinguístico torna-se útil para interpretar a atuação de fatores linguísticos e extralinguísticos na prática de escrita, revelando que determinados caminhos e escolhas eram influenciados pelos conflitos referentes aos períodos históricos da ortografia da língua portuguesa.

Os resultados mostraram que, apesar de os *scriptores* fazerem parte de uma elite instruída — uma vez que as habilidades de ler e escrever seriam fundamentais para desempenhar os trabalhos exigidos em sua paróquia de atuação —, a escrita era influenciada pelos conflitos referentes aos períodos históricos da ortografia da língua portuguesa, como se observa nas grafias etimológicas, principalmente nas variações das consoantes geminadas ou dobradas, nos encontros consonantais impróprios, nas substituições consonantais e nas variações no uso do <h>. Nota-se que existem grafias que podem ser decorrentes da influência da oralidade, especialmente nas variações vocálicas das grafias <e>, <i> e <o>, <u>, nas posições não acentuada final, inicial absoluta e pretônica.

Por meio dos dados linguísticos registrados nos assentos de casamentos, é possível constatar que as oscilações na grafia eram frequentes na época em que o manuscrito foi produzido. Os *scriptores* seguiam um modelo de escrita conforme a tradição documental — dicionários, gramáticas e manuais ortográficos — do período, apesar da falta de uma normatização oficial vigente, gerando divergências gráficas na escrita.

Por meio desta pesquisa, espera-se contribuir para a compreensão dos aspectos sócio-históricos relacionados ao processo de produção de escrita dos arquivos paroquiais do século XVIII, bem como para a reconstrução da história social e linguística do português brasileiro, sobretudo do sertão do Rio São Francisco.

REFERÊNCIAS

Fontes manuscritas:

Livro de Registros de casamentos e óbitos da Freguesia de Santo Antônio do Urubu de Baixo do Rio São Francisco (1719-1953). Arquivo da Cúria Diocesana de Bom Jesus da Lapa – Bahia.

Fontes impressas e eletrônicas:

- ACIOLI, V. L. C. **A escrita no Brasil Colônia: um guia para leitura de documentos manuscritos**. Recife: FUNDAJ, Editora Massangana; UFPE, Editora Universitária, 1994.
- ALI, M. S. **Gramática histórica da língua portuguesa**. 5.ed. São Paulo: Melhoramentos, 1964.
- ANTHONY, L. AntConc: Design and development of a freeware corpus analysis toolkit for the technical writing classroom. **IEEE International Professional Communication Conference Proceedings**, 2005. p. 729-737.
- ANTHONY, L. **Lawrence Anthony Website (AntConc)**. Disponível em: <http://www.antlab.sci.waseda.ac.jp/index.html>. Acesso em: 08 nov. 2021.

- ASSIS, L. B.; SILVA, J. A. A.; PACHECO, V. Uso de abreviaturas em um livro de assentos de casamentos do século XVIII do sertão de baixo do Rio São Francisco. **Revista Prolíngua**, v. 16, n. 1, p. 229–243, 2021a.
- ASSIS, L. B.; SILVA, J. A. A.; PACHECO, V. “Honde he o contrahente natural emorador”: emprego do <h> em assentos de casamentos do século XVIII. **Revista Philologus**, Ano 27, a. 79 Suplemento, Rio de Janeiro: SENEFIL, 2021b.
- BACELLAR, C. Uso e mal uso dos arquivos. In: PINSKY, C. B. (Org.). **Fontes Históricas**. 2.ed. São Paulo: Editora Contexto, 2008. p. 23-79.
- BAIA, Arcebispado. Constituições primeyras do Arcebispado da Bahia feytas, & ordenadas pelo Illustríssimo, e Reverendíssimo Senhor D. Sebastião Monteyro da Vide, Arcebispo do dito Arcebispado, & do Conselho de Sua Magestade, propostas, e aceytas em o sínodo Diecesano que o dito Senhor celebrou em 12 de junho de 1707. Lisboa Occidental: na Oficina de Pascoal da Sylva, Impressor de Sua Magestade, 1719. Disponível em: <http://purl.pt/24092>. Acesso em: 18 nov. 2021.
- BELLOTTO, H. L. **Como fazer análise diplomática e análise tipológica de documento de arquivo**. São Paulo: Arquivo do Estado, Imprensa Oficial, 2002.
- BERGS, A. **Social networks and historical linguistics: studies in morphosyntactic variation in the Paston letters (1421- 1503)**. Berlin; New York: Mouton de Gruyter, 2005.
- BERWANGER, A. R.; LEAL, J. E. F. **Noções de Paleografia e Diplomática**. Santa Maria: Editora UFSM, 2008.
- BLUTEAU, R. **Vocabulário português & latino: aulico, anatomico, architectonico ...** Coimbra: Collegio das Artes da Companhia de Jesus, 1728. 8 v. Disponível em: <http://purl.pt/13969>. Acesso em: 10 jan. 2022.
- BORGES NETO, J. **A Teoria da Linguagem de Fernão de Oliveira**. In: ABAURRE, M. B.; PFEIFFER, C.; AVELLAR, J. (Orgs.) **Fernão de Oliveira**. Um gramático na História. Campinas: Pontes Editores, 2009. p. 43-51.
- CAMBRAIA, C. N. **Introdução à crítica textual**. São Paulo: Martins Fontes, 2005.
- CONDE SILVESTRE, J. C. **Sociolinguística Histórica**. Madrid: Editorial Gredos, 2007.
- COUTINHO, I. L. **Pontos de gramática histórica**. 7.ed. Rio de Janeiro: Ao Livro Técnico, 1976.
- ECKERT, P.; McCONNELL-GINET, S. Comunidades de práticas: lugar onde co-habitam linguagem, gênero e poder. In: OSTERMANN, A. C.; FONTANA, B. (Org.). **Linguagem. Gênero. Sexualidade**. Clássicos traduzidos. São Paulo: Parábola Editorial, 2010, p. 93-107.
- GIMENO MENÉNDEZ, F. **Sociolingüística histórica (siglos X-XII)**. Madrid: Visor, 1995.
- HERNÁNDEZ-CAMPOY, J. M.; CONDE SILVESTRE, J. C. **The Handbook of Historical Sociolinguistics**. Oxford: Wiley-Blackwell, 2012.
- HIGOUNET, C. **História concisa da escrita**. Tradução de Marcos Marcionilo. São Paulo: Parábola, 2003.
- LABOV, W. **Padrões sociolingüísticos**. Tradução de Marcos Bagno, Maria Marta Pereira Scherre e Caroline Rodrigues Cardoso. São Paulo: Parábola Editorial, 2008 [1972].
- LABOV, W. **Principles of linguistic change: internal factors**. Oxford-UK; Cambridge-MA: Wiley-Blackwell, 1994.
- MATTOS E SILVA, R. V. **O português arcaico**. Fonologia, morfologia e sintaxe. São Paulo: Contexto, 2006.
- MADUREIRA FEIJÓ, J. M. **Orthographia, ou arte de escrever, e pronunciar com acerto a língua portuguesa**. Lisboa Occidental: na Officina de Miguel Rodrigues, Impressor do Senhor Patriarca, 1734. Disponível em: <http://purl.pt/13>. Acesso em: 10 jan. 2022.
- MEYERHOFF, M. **Introducing Sociolinguistics**. Routledge, New York, 2011.
- MOREL PINTO, R. **História da língua portuguesa: IV. Século XVIII**. São Paulo: Ática, 1988.
- SILVA, J. A. A. O estado da língua portuguesa num documento do século XVIII: abordagem sócio-histórica de fenômenos linguístico do livro das Monjas. In: CAMBRUSSI, M. F.; ARAGÃO NETO, M. M. (Orgs.). **Léxico e Gramática**. v.1. Curitiba: CVR, 2011. p. 259-283.
- TEIXEIRA, A. **Ensino superior no Brasil: análise e interpretação de sua evolução até 1969**. Rio de Janeiro: Editora da Fundação Getúlio Vargas, 1989.
- TEYSSIER, P. **História da Língua Portuguesa**. Tradução de Celso Cunha. 3.ed. São Paulo: Martins Fontes, 2007.
- ROMAINE, S. **Socio-Historical Linguistics: its status and methodology**. Cambridge University Press, 2009 [1982].
- WEINREICH, U.; LABOV, W.; HERZOG, M. **Fundamentos empíricos para uma teoria da mudança linguística**. São Paulo: Parábola Editorial, 2006 [1968].
- WILLIAMS, E. B. **Do latim ao português**. Fonologia e morfologia históricas da língua portuguesa. 3. ed. Tradução de Antônio Houaiss. Rio de Janeiro: Tempo Brasileiro, 1975.

SOBRE OS ORGANIZADORES

CRISTINA DOS SANTOS CARVALHO

Tem graduação (licenciatura e bacharelado) em Letras Vernáculas (1992-1993) e mestrado em Letras e Linguística (1997) pela Universidade Federal da Bahia (UFBA). Doutora em Linguística (2004) pela Universidade Estadual de Campinas. Realizou estágio pós-doutoral (2016-2018) na Universidade Federal do Rio de Janeiro (UFRJ). Atualmente, é professora plena da Universidade do Estado da Bahia (UNEB) e docente permanente do Programa de Pós-Graduação em Estudo de Linguagens (PPGEL). É líder do *Grupo de Pesquisa Usos e Contextos da Língua Portuguesa* – GConPort (UNEB). Tem experiência na área de letras, com ênfase em linguística e língua portuguesa, atuando principalmente nos seguintes temas: variação e mudança linguísticas, sociolinguística variacionista, funcionalismo, sentenças complexas, gramaticalização e abordagem construcional da gramática.

SILVANA SILVA DE FARIAS ARAÚJO

É graduada em Letras Vernáculas – Licenciatura (1999) e especialista em Língua Portuguesa (2000) pela Universidade Estadual de Feira de Santana (UEFS). Mestre em Letras e Linguística e doutora em Língua e Cultura pela Universidade Federal da Bahia (UFBA). É professora titular de língua portuguesa do Departamento de Letras e Artes da UEFS, atuando em cursos de graduação e de pós-graduação. Realizou dois estágios de pós-doutoramento: um em Linguística (2020), na Universidade Federal de Sergipe (UFS), com bolsa PDJ/CNPq, e outro em Língua Portuguesa (2019), na Universidade Federal do Rio de Janeiro (UFRJ). Tem experiência na área de sociolinguística, investigando principalmente os seguintes temas: formação do português do Brasil, variedades linguísticas minorizadas, contatos linguísticos, fenômenos morfossintáticos e variedades africanas do português.

VALTER DE CARVALHO DIAS

É licenciado em Letras Espanhol (2006) e Letras Português/Inglês (2005) pelo Centro Universitário Jorge Amado. Especialista em Estudos Linguísticos e Literários (2005) pela Universidade Federal da Bahia (UFBA), mestre em Estudo de Linguagens (2010) pela Universidade do Estado da Bahia (UNEB) e doutor em Língua e Cultura (2017) pela Universidade Federal da Bahia (UFBA). Atualmente, é docente de língua portuguesa com dedicação exclusiva do Instituto Federal da Bahia e editor-chefe da *Enlaces: revista de estudos linguísticos e literários*. Dedicar-se a estudos na área de Linguística, com ênfase em sociolinguística variacionista e sua relação com o ensino de língua portuguesa. É líder dos grupos de pesquisa *Estudos da Língua Portuguesa: descrição e ensino* – ELiPor e *Usos e Contextos da Língua Portuguesa* – GConPort, e membro do *Grupo de Pesquisa em Dialetoлогия e Sociogeolingüística* – GPDS-UFOP.

SOBRE OS AUTORES

CAMILO ROSA SILVA

Tem graduação em Letras pela Universidade Federal do Rio Grande do Norte (1989) e mestrado em Letras pela Universidade Federal da Paraíba (1999). Tem doutorado em Letras pela Universidade Federal da Paraíba (2005), com doutorado-sanduíche na Universidade Federal de Santa Catarina (2003). Tem pós-doutorado em Letras pela Universidade Federal de São Paulo (2018). É professor titular da Universidade Federal da Paraíba, onde atua na graduação em Letras e na pós-graduação em Linguística. Desenvolve pesquisas na área de linguística funcional, com ênfase em teoria e análise linguística, situando-se na interface gramática e ensino. Publicou diversos artigos em periódicos qualificados, organizou livros e publicou capítulos de livro com temáticas relacionadas à sua área de atuação.

CLAUDIA ROBERTA TAVARES SILVA

É graduada em Letras (2000) pela Universidade Federal de Alagoas (UFAL). É doutora em Linguística (2004) pela UFAL, tendo realizado o doutorado-sanduíche (2003-2004) na Universidade Nova de Lisboa (UNL). Realizou pós-doutorado (estágio Sênior no Exterior/CAPES) em Linguística na Universidade de Lisboa (2016). É pesquisadora nos grupos de pesquisa *Programa de Estudos Linguísticos* e *Grupo de Estudos em Teoria da Gramática* (GETEGRA), vinculados, respectivamente, à UFAL e à Universidade Federal do Agreste de Pernambuco (UFAPE). Atualmente, é professora associada 3 do Departamento de Letras da Universidade Federal Rural de Pernambuco (UFRPE) e docente do Programa de Pós-Graduação em Estudos da Linguagem (PROGEL) da UFRPE e do Programa de Pós-Graduação em Letras (PPGL) da Universidade Federal de Pernambuco (UFPE). Atua nas áreas de letras e linguística, com ênfase em teoria e análise linguística, centrando a atenção principalmente nos campos da sintaxe, da morfologia, da variação linguística e do contato linguístico.

CLEBER TOURINHO DE SANTANA

É graduado em Letras Vernáculas – licenciatura (2007) pela Universidade Federal da Bahia (UFBA) e especialista em Língua Portuguesa e Matemática numa Perspectiva Transdisciplinar (2016) pelo Instituto Federal do Rio Grande do Norte (IFRN). Tem mestrado profissional em Letras – Proletras (2016-2018) pela Universidade Federal da Paraíba (UFPB). É doutorando em Linguística na UFPB. É professor de língua Portuguesa, vinculado à Secretaria de Educação do Estado da Paraíba e à Secretaria Municipal de Educação e Cultura de João Pessoa (PB). Tem experiência na área de linguística, com ênfase em linguística histórica e sociolinguística.

ELYNE GISELLE DE SANTANA LIMA AGUIAR VITÓRIO

Tem graduação em Letras (2005) e especialização em Linguística e Ensino do Português (2006) pela Universidade Federal do Ceará (UFC). Tem mestrado em Linguística (2008) e doutorado em Linguística (2012) pelo Programa de Pós-Graduação em Letras e Linguística da Universidade Federal de Alagoas (UFAL). Realizou estágio de pós-doutorado na Pós-Graduação em Letras Vernáculas da Universidade Federal do Rio de Janeiro (2013) – PDJ/CNPq e no Programa de Pós-Graduação em Letras e Linguística da Universidade Federal de Alagoas (2014) – PNPd/Capes. É professora de linguística da UFAL (*Campus Arapiraca*), atuando na graduação e no Programa de Pós-Graduação em Linguística e Literatura. Coordena o grupo de pesquisa *Sociolinguística, variação, significados sociais e ensino* e tem experiência na área de linguística, com ênfase em sociolinguística variacionista.

HEBE MACEDO DE CARVALHO

Tem licenciatura em Letras (1993) e mestrado em Letras (1997) pela Universidade Federal da Paraíba (UFPA) e doutorado em Linguística (2007) pela Universidade Federal do Ceará (UFC). Realizou estágio de pós-doutorado em Linguística (2017-2018) na Universidade Federal do Espírito Santo (UFES). É professora do Departamento de Letras Vernáculas – nível associado IV, professora permanente do Programa de Pós-Graduação em Linguística da Universidade Federal do Ceará (UFC), na linha de pesquisa *Descrição e análise linguística*. Desenvolve estudos sobre a variação e mudança de categorias verbais e pronominais na fala do português brasileiro, sob a perspectiva da sociolinguística variacionista.

HUDA DA SILVA SANTIAGO

Tem graduação em Letras Vernáculas (2002), especialização em Estudos Linguísticos (2007) e mestrado em Estudos Linguísticos (2012) pela Universidade Estadual de Feira de Santana (UEFS). Tem doutorado (2019) em Língua e Cultura pelo Programa de Pós-Graduação em Língua e Cultura (PPGLINC), da Universidade Federal da Bahia (UFBA). Integra a equipe do *Projeto Corpus Eletrônico de Documentos Históricos do Sertão* (CE-DOHS), da UEFS, do Programa para a História da Língua Portuguesa (PROHPOR) e do Programa História da Cultura Escrita no Brasil (HISCULTE), da UFBA. É professora assistente da UEFS. Atua como professora permanente no Programa de Pós-Graduação em Estudos Linguísticos (PPGEL/UEFS), na linha *Variação e mudança linguística no português*. Tem experiência na área de linguística, atuando principalmente nos seguintes temas: constituição de *corpus*, sócio-história do português brasileiro, português popular brasileiro, variação linguística, história da cultura escrita, ensino-aprendizagem de língua portuguesa na educação básica.

HUGO LEONARDO GOMES DOS SANTOS

Tem graduação em Letras Português/Literatura – licenciatura plena (2014) pela Universidade Estadual do Ceará (UECE) e especialização em Metodologia de Ensino de Língua Portuguesa e Literatura (2017) pela Faculdade Educacional da Lapa (FAEL). É mestre em Linguística Aplicada (2016) pelo Programa de Pós-Graduação em Linguística Aplicada da UECE e doutorando do Programa de Pós-Graduação em Linguística da Universidade Federal do Ceará (UFC). Membro dos grupos de pesquisa: *Lexicologia, Terminologia e Ensino* (LETENS/UECE), *Pesquisas Sociolinguísticas de Fortaleza* (SOCIOFOR/UECE) e *Pesquisas Sociolinguísticas* (SOCIOLIN-CE/UFC). É professor da rede pública de ensino da Prefeitura Municipal de Fortaleza (CE). Atua principalmente nos seguintes temas: sociolinguística, descrição do português, lexicografia, dicionário escolar e análise de livro didático.

JORGE AUGUSTO ALVES DA SILVA

Tem mestrado (2003) e doutorado em Letras e Linguística (2005) pela Universidade Federal da Bahia (UFBA). É professor pleno do Departamento de Estudos Linguísticos e Literários (DELL) da Universidade Estadual do Sudoeste da Bahia (UESB). É professor do quadro efetivo do Programa de Pós-Graduação *Stricto Sensu* em Linguística e do mestrado acadêmico em Letras (ProfLetras). Atualmente, desenvolve estudos sobre a sócio-história do português do Brasil, com ênfase no português popular (urbano, rural, rurano, afro-brasileiro, variedades de minorias, incluindo línguas de santo e de manifestações linguísticas de religiões de matriz africana). Conjuga a teoria da variação aos estudos filológicos. Dedicar-se, ainda, a estudos voltados às línguas em uso.

LÉCIO BARBOSA DE ASSIS

Tem graduação em Letras - Habilitação em Português, Inglês e respectivas Literaturas (2004) e bacharelado em Administração (2014) pela Universidade do Estado da Bahia (UNEB). Tem pós-graduação *lato sensu* em Língua Portuguesa (2007), MBA em Gestão de Pessoas (2013), pós-graduação *lato sensu* em Metodologia de Ensino e Pesquisa em Língua Inglesa (2014) e pós-graduação *lato sensu* em Educação Inclusiva com ênfase em Libras (2015). É mestre (2019) e doutorando em Linguística pelo Programa de Pós-graduação em Linguística da Universidade Estadual do Sudoeste da Bahia (UESB).

NORMA DA SILVA LOPES

É licenciada em Letras (1973) e mestre em Letras (1981) pela Universidade Federal da Bahia (UFBA). Tem doutorado em Letras (2001) pela UFBA, com bolsa-sanduíche na Universidade de Macau (RAEM), na China. Realizou pós-doutorado (2014-2015) na Universidade Estadual de Feira de Santana (UEFS). Atualmente, é professora plena da Universidade do Estado da Bahia (UNEB) e professora permanente do Programa de Pós-Graduação de Estudo de Linguagens (PPGEL), da UNEB. Tem experiência na área de linguística, com ênfase em variação e mudança linguística, atuando principalmente nos seguintes temas: sócio-história do português brasileiro; contribuição africana na formação do português brasileiro; variação linguística no estado da Bahia.

ROSANA CARVALHO BRITO

É graduada em Letras Vernáculas, especialista em Linguística e Ensino-Aprendizagem de Língua Portuguesa e Mestre em Estudos Linguísticos pela Universidade Estadual de Feira de Santana (UEFS). É doutoranda no Programa de Pós-Graduação em Estudos Linguísticos (PPGEL) da UEFS. Atuou como pesquisadora de Iniciação Científica na modalidade PROBIC/UEFS e na modalidade PIBIC/CNPq. Em 2015, obteve o primeiro lugar pelo melhor trabalho de iniciação científica como bolsista do PIBIC/CNPq – Área de Ciências Humanas, Letras e Sociais Aplicadas. É integrante do Projeto de Pesquisa CE-DOHS - *Corpus* Eletrônico de Documentos Históricos do Sertão (www.uefs.br/dohs) (FAPESB 5566/2010). É professora da educação básica (SEC/BA).

ROSANE BISPO DOS SANTOS

Tem graduação em Letras Português / Inglês (licenciatura) pela Faculdade de Tecnologia e Ciências (FTC) e em Bacharelado Interdisciplinar de Artes pela Universidade Federal da Bahia (UFBA). É especialista em Psicopedagogia Institucional pela Faculdade Afonso Claudio (FAAC) em Gramática e Texto pela Universidade Salvador (UNIFACS). Mestre em Estudo de Linguagens pela Universidade do Estado da Bahia (UNEB). Atualmente, é professora de língua portuguesa pertencente ao quadro efetivo da Prefeitura Municipal de Camaçari. Atua nas áreas de linguística e língua portuguesa, socio-funcionalismo e gramaticalização.

VERA PACHECO

Tem graduação em Linguística (2001) e em Letras – bacharelado e licenciatura (2002) pela Universidade Estadual de Campinas (UNICAMP), onde também cursou o mestrado (2003) e o doutorado (2006) em Linguística. Realizou pós-doutorado na Universidade Paulista “Júlio de Mesquita Filho” / Araraquara. Atualmente, é professora titular da Universidade Estadual do Sudoeste da Bahia (UESB). Tem experiência na área de linguística, com ênfase em teoria e análise linguística, atuando principalmente nos seguintes temas: análise acústica, percepção da fala e prosódia.

WENDEL SILVA DOS SANTOS

É licenciado em Letras (2011) pela Universidade Federal do Maranhão, mestre (2015) e doutor (2020) em Linguística pela Universidade de São Paulo. Atualmente, é professor adjunto da graduação e do mestrado em Letras da Universidade Federal do Maranhão, Centro de Ciências, Educação e Linguagem. É coordenador do *Grupo de Estudos e Pesquisa em Sociolinguística do Maranhão* (GEPeS/UFMA). Tem experiência na área de Linguística, com ênfase em sociolinguística variacionista, atuando em pesquisas sobre variação e mudança linguísticas. Mais recentemente, tem se interessado por pesquisas acerca da percepção sociolinguística de variantes linguísticas e os significados sociais associados a elas.

