

UNIVERSIDADE ESTADUAL DE GOIÁS – UEG
PROGRAMA DE PÓS-GRADUAÇÃO STRICTO SENSU EM LÍNGUA, LITERATURA
E INTERCULTURALIDADE – POSLLI

VANDER SIMÃO MENEZES

A NEGAÇÃO COMO VARIÁVEL SOCIOLINGUÍSTICA NA CIDADE DE GOIÁS-GO

Goiás-GO
2022

VANDER SIMÃO MENEZES

A NEGAÇÃO COMO VARIÁVEL SOCIOLINGUÍSTICA NA CIDADE DE GOIÁS-GO

Dissertação apresentada ao Programa de Pós-Graduação *Stricto Sensu* em Língua, Literatura e Interculturalidade para exame de defesa.

Linha de pesquisa: Estudos de Língua e Interculturalidade

Goiás-GO
2022

TERMO DE AUTORIZAÇÃO PARA PUBLICAÇÃO DE TESES E DISSERTAÇÕES NA BIBLIOTECA DIGITAL (BDTD)

Na qualidade de titular dos direitos de autor, autorizo a Universidade Estadual de Goiás a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UEG), regulamentada pela Resolução, CsA nº 1.087/2019 sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9.610/1998, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira, a partir desta data¹. Estando ciente que o conteúdo disponibilizado é de inteira responsabilidade do(a) autor(a).

Dados do autor (a)

Nome completo Vander Simão Menezes

E-mail vander.simao@gmail.com

Dados do trabalho

Título A negação como variável sociolinguística na cidade de Goiás-GO

Tipo:

Tese Dissertação

Curso/Programa Programa de Pós-graduação em Língua, Literatura e Interculturalidade.

Concorda com a liberação documento

SIM NÃO

¹ Período de embargo é de até **um ano** a partir da data de defesa.

Goiás, 23 de março de 2022

Vander Simão Menezes
Assinatura autor(a)

Marilia Silva Vieira
Assinatura do orientador(a)

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA FONTE

Biblioteca Frei Simão Dorvi – UEG Câmpus Cora Coralina

M543n Menezes, Vander Simão.

A negação como variável sociolinguística na Cidade de Goiás-GO [manuscrito] / Vander Simão Menezes. – Goiás, GO, 2022.
143f.

Orientadora: Profa. Dra. Marília Silva Vieira.

Dissertação (Mestrado em Língua, Literatura e Interculturalidade) – Câmpus Cora Coralina, Universidade Estadual de Goiás, 2022.

1. Sociolinguística - português brasileiro.
1.1. Variação linguística. 1.2. Línguas bantas.
1.3. Fala vilaboense - Goiás, GO. I. Título.
II. Universidade Estadual de Goiás, Câmpus Cora Coralina.

CDU: 81'27(817.3)

Bibliotecária responsável: Marília Linhares Dias – CRB 1/2971

UNIVERSIDADE ESTADUAL DE GOIÁS

(Criada pela lei nº 13.456 de Abril de 1999, publicada no DOE-GO de 20 de Abril de 1999)

Pró-Reitoria de Pesquisa e Pós-Graduação
Coordenação de Pós-Graduação Stricto Sensu
UEG CÂMPUS CORA CORALINA

Av. Dr. Deusdeth Ferreira de Moura Centro - GOIÁS CEP: 76600000

Telefones: (62)3936-2161 / 3371-4971 Fax: (62) 3936-2160 CNPJ: 01.112.580/0001-71

ATA DE EXAME DE DEFESA 02/2022

Aos vinte e seis dias do mês de janeiro de dois mil e dois às catorze horas, realizou-se, por webconferência, o Exame de Defesa da dissertação do(a) mestrando(a) Vander Simão Menezes, intitulado “**A NEGAÇÃO COMO VARIÁVEL SOCIOLINGUÍSTICA NA CIDADE DE GOIÁS-GO**”. A banca examinadora foi composta pelos seguintes professores: Dra. Marília Silva Vieira – Presidente – (POSLLI/UEG), Dr. Scott Schwenter (The Ohio State University), Dr. Antônio Carlos Santana de Souza (UEMS), Dr. Eleone Ferraz de Assis (POSLLI/UEG). Os membros da banca fizeram suas observações e sugestões, as quais deverão ser consideradas pelo(a) mestrando(a) e seu/sua orientador(a). Em seguida, a banca examinadora reuniu-se para proceder à avaliação do exame de defesa. Reaberta a sessão, o(a) presidente da banca examinadora, proclamou o resultado, segundo o qual a dissertação foi (x) aprovada, com as seguintes exigências: realizar as alterações solicitadas pela banca, sobretudo na análise estatística da pesquisa, no que tange à apresentação dos resultados. Cumpridas as formalidades de pauta, às 15h44min a presidência da mesa encerrou esta sessão do Exame de Defesa e lavrou a presente ata que, após lida e aprovada, será assinada pelos membros da banca examinadora.

Goiás-GO, 26 de janeiro de 2022.



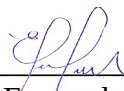
Profa. Dra. Marília Silva Vieira (POSLLI/UEG)



Prof. Dr. Scott Schwenter (The Ohio State University)



Prof. Dr. Antônio Carlos Santana de Souza (UEMS)



Prof. Dr. Eleone Ferraz de Assis (POSLLI/UEG)

À minha esposa, Alaine – meu girassol.

AGRADECIMENTOS

À minha família. Meus pais, Valdeir e Romi, que sempre me guiaram e me apoiaram. À minha querida irmã, Karine, pelo apoio e companheirismo que só irmãos podem ter. À minha lindíssima sobrinha, Ana Cecília, por ser um raio de sol que me iluminou nos momentos mais difíceis.

À minha esposa, Alaine, pela imensa paciência, carinho, compreensão, apoio, incentivo em todos os momentos deste longo trajeto. Sem ela, nada seria possível, nada. Amo-te.

Ao meu nobre amigo Gregory pela camaradagem, boas conversas e pelo clubismo. Agradeço aos colegas Luciana, Cleiton e Patrícia por compartilharem comigo um pouco da experiência que obtiveram ao desenvolver suas respectivas pesquisas, em especial à Luciana por ceder o banco de dados que levantou em 2019.

Ao Grupo de Estudos Funcionalistas, da Universidade Federal de Goiás (GEF/UFG) e ao Grupo de Estudos e Pesquisas em Sociolinguística, da Universidade Estadual de Goiás (Sociolingo/UEG/Cora Coralina) por disponibilizarem os bancos de dados que possibilitaram a realização desta pesquisa.

Agradeço à 4ª turma do POSLLI que mesmo distante, por conta das medidas restritivas impostas pela pandemia da covid-19, não deixou de confraternizar e de dar apoio mútuo.

À Universidade Estadual de Goiás, em especial ao programa de Pós-graduação *Stricto Sensu* em Língua, Literatura e Interculturalidade, por me proporcionar a oportunidade de concretizar o sonho de ser mestre. Igualmente, agradeço aos professores do POSLLI que me ensinaram tanto e propiciaram meu crescimento ao longo desses dois anos.

Ao Prof. Dr. Almir Almeida de Oliveira, da Universidade Estadual de Alagoas (Uneal), pela ajuda com o “R core team”, nossas conversas foram muito esclarecedoras para minha análise de dados.

À banca de qualificação, presidida por minha orientadora, Professora Marília Vieira, e composta pelo Prof. Dr. Antonio Carlos Santana de Souza (UEMS) e pelo Professor Dr. Eleone Ferraz de Assis (UEG/POSLLI), agradeço por aceitarem o convite e pela contribuição.

À banca composta para a defesa, presidida por minha orientadora, Professora Marília Vieira, e composta pelos professores PhD. Scott A. Schwenter, Dr. Antonio Carlos Santana de Souza e Dr. Eleone Ferraz de Assis, agradeço por aceitarem o convite e pelas contribuições.

À Profa. Dra. Marília Silva Vieira por me apresentar a Sociolinguística e me ajudar a compreender as nuances desta corrente de estudos linguísticos que tão belamente permite estudar e compreender os fenômenos da língua(gem). Agradeço pela paciência, pelos conselhos, pelo exemplo, pelas aulas, pela sabedoria, enfim, pela ótima orientação. Obrigado, professora.

Quando os eruditos descobriram a língua, ela já estava completamente pronta pelo povo. Os eruditos tiveram apenas que proibir o povo de falar errado.

Millôr Fernandes

RESUMO

O objeto deste estudo é o fenômeno de variação da negação sentencial. Observamos, no Português Brasileiro (doravante, PB), a ocorrência da negação em três diferentes posições em relação ao verbo: pré-verbal (NEG+V) – NEG1; duplamente marcada, antes e depois do verbo (NEG+V+NEG) – NEG2; e pós-verbal (V+NEG) – NEG3. O objetivo da presente pesquisa é descrever a variação da negação sentencial na fala vilaboense. O fenômeno em questão é objeto de diversos trabalhos, tais como: Avelar; Silva; Almeida (2013); Cavalcante (2007); Cunha (2001); Nascimento (2014); Rocha (2013) e Sousa (2012). Orientamo-nos, para este estudo, pela Sociolinguística Variacionista (LABOV, 1978; LABOV, 2008; WEINREICH; LABOV; HERZOG, 2006) e pela concepção de que o Português Brasileiro se consolidou como uma língua ímpar devido a uma série de fatores, como o contato entre línguas ocorrido ao longo da história nacional (AMARAL, 1976; BORTONI-RICARDO, 2021; LIPSKI, 2008; LUCCHESI, 2017). Há indícios significativos de que a variação da forma de negação utilizada pelo falante tenha influência das línguas bantas e de que seja motivado por fatores pragmáticos, tais como a economia e a informatividade. O estudo da língua, que aqui propomos, parte de evidências empíricas que permitem a observação da variação que ocorre na fala. Esta abordagem fez com que seja necessária a constituição de um *corpus* para a análise. Para isso, lançou-se mão do banco de dados de dois grupos de estudo: do Grupo de Estudos Funcionalistas, da Universidade Federal de Goiás (GEF/UFG); e do Grupo de Estudos e Pesquisas em Sociolinguística, da Universidade Estadual de Goiás (Sociolinco/UEG/Cora Coralina). No *corpus* analisado, encontramos 1990 ocorrências para NEG1 e 553 ocorrências de NEG2, ao passo que NEG3 ocorreu apenas 91 vezes. Verificamos que as chances de NEG2 ser empregada são maiores na amostra coletada em 2003, com informantes do sexo masculino e com menos de 35 anos de idades em 2003. No que se refere às variáveis linguísticas, constatamos que, quando a informação é diretamente ativada no discurso, as sequências discursivas avaliativas e dialogais, as orações absolutas, a ausência de marcadores discursivos aumenta as chances de variação.

Palavras-chave: Variação linguística. Português Brasileiro. Línguas bantas. Cidade de Goiás-GO. Negação sentencial.

ABSTRACT

This study object is the phenomenon of negative sentence variation. Brazilian Portuguese (henceforth BP) is seen to have three different negative positions regarding the verb: pre-verbal (NEG+V) – NEG1; double marked, before and after the verb (NEG+V+NEG) – NEG2; and post-verbal (V+NEG) – NEG3. The objective of the present research is to describe the variation of the sentence denial in Vilaboense speech. Such phenomenon has been the focus of previous research, namely Avelar, Silva and Almeida (2013); Cavalcante (2007); Cunha (2001); Nascimento (2014); Rocha (2013) and Sousa (2012). Our study is oriented by Variationist Sociolinguistics (LABOV, 1978; LABOV, 2008; WEINREICH; LABOV; HERZOG, 2006) and by the assumption that BP has become a unique language due to a series of factors, such as language contact throughout national history (AMARAL, 1976; BORTONI-RICARDO, 2021; LIPSKI, 2008; LUCCHESI, 2017). There are significant clues that the variation in the form of negation chosen by the speaker has been influenced by Bantu languages and motivated by pragmatic factors, such as economy and informational content, which pressure communicative contexts. We propose a language approach that relies on empirical evidence, which enables the observation of speech variation. Thus, we had to constitute a corpus for analysis. We used databanks from two study groups: Functional Study Group, at Goiás Federal University (GEF/UFG); and Study Group and Research on Sociolinguistics, at Goiás State University (Sociolinco/UEG/Cora Coralina). We used R (R CORE TEAM, 2013) so as to carry out the statistical data analysis (OUSHIRO, 2017). In the surveyed corpus, we found 1990 hits for NEG1 and 553 hits for NEG2, whereas NEG3 occurred only 91 times. We found out that NEG2 has an increased chance to occur when compared to the 2003 sample, with male informants and under 35 years of age in 2003. Regarding linguistic variables, we found out that when information is directly activated in discourse, NEG2 has an increased chance to occur in conjunction with evaluative and dialogical discourse sequences, absolute sentences and absence of discourse markers increases the chances of variation.

Keywords: Negative Sentence. Linguistic Variation. Sociolinguistics. Bantu Languages. Town of Goiás-GO.

LISTA DE FIGURAS

Figura 1: Localização da Cidade de Goiás-GO	62
Figura 2: Praça do Coreto (19--).....	67
Figura 3: Centro da Cidade de Goiás – GO (1957)	67
Figura 4: Casarões da Cidade de Goiás (1957).	68
Figura 5: Igreja Nossa Senhora do Rosário e Centro de Cultura e Memória do Poder Judiciário do Estado de Goiás, ao fundo (2021).....	68
Figura 6: Cidade de Goiás: vista panorâmica (2021).	69
Figura 7: Planilha de ocorrência .CSV.....	76
Figura 8: Interface do “RStudio”	78
Figura 9: Resumo do modelo geral	113
Figura 10: Teste “Forward”	116
Figura 11: Teste “Backward”	116
Figura 12: Teste “Both”	117
Figura 13: Resumo do modelo final.....	117

LISTA DE GRÁFICOS

Gráfico 1 - Total de ocorrências com NEG3.....	92
Gráfico 2 - Total de ocorrências	93
Gráfico 3 - Total de ocorrências pelo Status da informação negada - GEF	95
Gráfico 4 - Total de ocorrências pelo Status da informação negada - Sociolinco	96
Gráfico 5 - Total de ocorrências pelo tipo de sequência discursiva - GEF	98
Gráfico 6 - Total de ocorrências pelo tipo de sequência discursiva - Sociolinco	98
Gráfico 7 - Total de ocorrências pelo tipo de oração - GEF	100
Gráfico 8 - Total de ocorrências pelo tipo de oração - Sociolinco	100
Gráfico 9 - Total de ocorrências pela presença ou ausência de marcadores discursivos - GEF	102
Gráfico 10 - Total de ocorrências pela presença ou ausência de marcadores discursivos - Sociolinco	103
Gráfico 11 - Total de ocorrências pelo tipo de sujeito - GEF	104
Gráfico 12 - Total de ocorrências pelo tipo de sujeito - Sociolinco	105
Gráfico 13 – Total de ocorrências por amostra	106
Gráfico 14 – Total de ocorrências pelo sexo do informante - GEF	108
Gráfico 15 – Total de ocorrências pelo sexo do informante - Sociolinco	109
Gráfico 16 – Total de ocorrências faixa etária - GEF	110
Gráfico 17 – Total de ocorrências faixa etária - Sociolinco.....	111
Gráfico 18 – Tipo de oração	118
Gráfico 19 – Tipo de sequência.....	119
Gráfico 20 – Presença ou ausência de marcadores discursivos	119
Gráfico 21 – Status discursivo da informação	119
Gráfico 22 – Tipo de sujeito.....	120
Gráfico 23 – Faixa etária	120

LISTA DE ESQUEMAS

Esquema 1 – Estágio de reanálise da negação	58
--	----

LISTA DE TABELAS

Tabela 1 – Ocorrências das estratégias negativas em Santa Luzia - MG	57
Tabela 2 - Informações por Cidades e Estados - Instituto Brasileiro de Geografia e Estatística.....	63
Tabela 3 – Distribuição geral dos dados com NEG3	92
Tabela 4 – Distribuição geral dos dados sem NEG3	93
Tabela 5 – Distribuição da negação sentencial em diferentes variedades linguísticas	94
Tabela 6 – A negação nas mesorregiões baianas – análise geral de estratégias	94
Tabela 7 – Distribuição dos dados pelo Status discursivo da informação negada - GEF	95
Tabela 8 – Distribuição dos dados pelo Status discursivo da informação negada - Sociolingo.....	96
Tabela 9 – Distribuição dos dados pelo Tipo de sequência discursiva por amostra .	97
Tabela 10 – Distribuição dos dados pelo Tipo de oração	99
Tabela 11 – Distribuição dos dados pela presença ou ausência de marcadores discursivos - GEF	101
Tabela 12 – Distribuição dos dados pela presença ou ausência de marcadores discursivos - Sociolingo	102
Tabela 13 – Distribuição dos dados por tipo de sujeito.....	104
Tabela 14 – Distribuição dos dados pela amostra	106
Tabela 15 – Distribuição dos dados por sexo - GEF.....	107
Tabela 16 – Distribuição dos dados por sexo - Sociolingo.....	108
Tabela 17 – Distribuição dos dados por faixa etária - GEF.....	110
Tabela 18 – Distribuição dos dados por faixa etária - Sociolingo.....	110
Tabela 19 – Distribuição dos dados por informante.....	112
Tabela 20 – Tabela do modelo de regressão logística final.....	118
Tabela 21 – Tabela do modelo de regressão logística com efeitos mistos.....	121
Tabela 22 – NEG2 por <i>status</i> da informação.....	122
Tabela 23 – Negação no PB, por <i>status</i> da proposição negada.....	123
Tabela 24 – NEG2 por <i>status</i> da informação – Amostra GEF	123
Tabela 25 – NEG2 por Tipo de Sequência discursiva	124
Tabela 26 – NEG2 por tipo de Sequência discursiva – Amostra GEF	125
Tabela 27 – NEG2 por tipo de Sequência discursiva – Amostra Sociolingo	125
Tabela 28 – NEG2 por tipo de oração	126
Tabela 29 – NEG2 por tipo de oração – Amostra GEF.....	127
Tabela 30 – NEG2 por tipo de oração – Amostra Sociolingo.....	127
Tabela 31 – NEG2 por presença ou ausência de marcadores discursivos.....	128
Tabela 32 – NEG2 por presença ou ausência de marcadores discursivos – Amostra GEF	128
Tabela 33 – NEG2 por presença ou ausência de marcadores discursivos – Amostra Sociolingo.....	129
Tabela 34 – NEG2 por tipo de sujeito	129

Tabela 35 – NEG2 por tipo de sujeito – Amostra GEF	130
Tabela 36 – NEG2 por tipo de sujeito – Amostra Sociolinco	130
Tabela 37 – NEG2 por Amostra.....	132
Tabela 38 – NEG2 por Sexo.....	133
Tabela 39 – NEG2 por Faixa etária	134

LISTA DE QUADROS

Quadro 1 – Relação entre informatividade e economia.....	30
Quadro 2 – Os tipos de negação no Português Brasileiro.....	32
Quadro 3 – Informantes da amostra GEF	73
Quadro 4 – Informantes da amostra Sociolingo.....	74

LISTA DE ABREVIATURAS E SIGLAS

GEF	Grupo de Estudos Funcionalistas
GO	Goiás
IBGE	Instituto Brasileiro de Geografia e Estatística
NEG1	Negação 1 – pré-verbal
NEG2	Negação 2 – dupla negação
NEG3	Negação – pós-verbal
PA	Português africano
PB	Português brasileiro
PE	Português europeu
UEG	Universidade Estadual de Goiás
UFG	Universidade Federal de Goiás

SUMÁRIO

INTRODUÇÃO	20
1 A SOCIOLINGUÍSTICA VARIACIONISTA	23
1.1 Estruturalismo e Sociolinguística: rupturas	24
1.2 Questões relativas à língua(gem)	26
1.3 A questão da variação	32
1.4 O valor de verdade da variável gramatical.....	36
1.5 A periodização da história sociolinguística do PB	40
1.6 O português brasileiro e as línguas africanas.....	41
2 A NEGAÇÃO COMO VARIÁVEL SOCIOLINGUÍSTICA	47
2.1 A estrutura da negação	47
2.2 O envelope de variação	50
2.2.1 A tripla negação	54
2.3 A coexistência das três formas de negação no PB.....	55
2.3.1 A negação nas diversas comunidades de fala brasileiras.....	56
2.3.2 Possíveis motivações para a variação entre as formas de negação: a abordagem funcional.....	58
2.3.3 Uma abordagem discursivo-pragmática do fenômeno	60
3. MATERIAIS E MÉTODO	62
3.1 A cidade de Goiás – aspectos históricos e socioeconômicos	62
3.2 Constituição do corpus.....	69
3.2.1 Amostras de fala	69
3.2.2 Banco de dados GEF	72
3.2.3 Banco de dados Sociolingo	73
3.3 A análise estatística na sociolinguística: <i>R</i>	76
4 VARIÁVEIS SOCIAIS E LINGUÍSTICAS	79
4.1 Variáveis linguísticas.....	79
4.1.1 Status da informação	79
4.1.2 Tipo de sequência discursiva	81
4.1.2.1 Sequência argumentativa	81
4.1.2.2 Sequência avaliativa.....	82
4.1.2.3 Sequência descritiva.....	82

4.1.2.4 Sequência dialógica.....	83
4.1.2.5 Sequência narrativa.....	83
4.1.2.6 Sequência procedural.....	84
4.1.3 Tipo de oração	85
4.1.4 Marcador discursivo	86
4.1.5 Tipo de sujeito.....	87
4.2 Variáveis sociais	88
4.2.1 Amostra.....	88
4.2.2 Sexo	89
4.2.3 Faixa etária	90
4.2.4 Informante	91
5. ANÁLISE DOS DADOS	92
5.1 Análise quantitativa	92
5.1.1 Variáveis linguísticas.....	94
5.1.1.1 Status da informação.....	95
5.1.1.2 Tipo de sequência discursiva.....	97
5.1.1.3 Tipo de oração.....	99
5.1.1.4 Marcador discursivo.....	101
5.1.1.5 Tipo de sujeito	103
5.1.2 Variáveis sociais.....	105
5.1.2.1 Amostra	105
5.1.2.2 Sexo	107
5.1.2.3 Faixa etária	109
5.1.2.4 Informante	111
5.1.3 Regressão Logística.....	112
5.2 Análise qualitativa	120
5.2.1 Variáveis linguísticas.....	122
5.2.1.1 <i>Status</i> da informação.....	122
5.2.1.2 Tipo de sequência discursiva.....	124
5.2.1.3 Tipo de oração.....	126
5.2.1.4 Marcador discursivo.....	127
5.2.1.5 Tipo de sujeito	129
5.2.2 Variáveis sociais.....	131

5.2.2.1 Amostra	132
5.2.2.2 Sexo	132
5.2.2.3 Faixa etária	133
CONSIDERAÇÕES FINAIS	136
REFERÊNCIAS.....	139
ANEXO A – ROTEIRO DE ENTREVISTA SEMIESTRUTURADA	144

INTRODUÇÃO

Em 1972, era lançado o estudo *Sociolinguistic Patterns* (Padrões Sociolinguísticos), de William Labov (2008 [1972]), e, com a sentença que introduz o texto, Labov expressa o problema de aceitar o termo “Sociolinguística” para os estudos conduzidos por ele. O problema, segundo o autor (LABOV, 2008 [1972], p. 13), está na possibilidade de se conceber algum tipo de abordagem dos fenômenos linguísticos que não incluísse os fatores sociais envolvidos.

Diante disso, inspiramo-nos no posicionamento de Labov (2008 [1972]) para a realização do presente estudo. Não poderíamos conduzir uma pesquisa em linguística que não considerasse aspectos sociais inerentes ao fenômeno de variação linguística. Em outras palavras, por termos tomado a negação como objeto de estudo, não podemos negar o peso dos fatores sociais na ocorrência das formas alternantes.

Esta pesquisa se insere no arcabouço de estudos que descrevem fenômenos linguísticos no português brasileiro (doravante, PB). Assim, para além da prescrição, pretendemos explicar a variação das estratégias de negação no PB. Como toda pesquisa é limitada a um espaço e a um tempo, essa pesquisa pretende averiguar este fenômeno na cidade de Goiás-GO, a partir de amostras coletadas no século XXI.

O fenômeno linguístico aqui descrito é resultado de pressões socioculturais sofridas, históricas, pragmáticas etc. Esforçamo-nos para descrever, da melhor maneira possível, os mecanismos que subjazem a escolha do falante no momento da interação verbal. Esta fotografia pode apontar a tendência na cidade mencionada, além de, contrastada ao atual estado da arte, podemos extrair generalizações acerca do fenômeno no PB como um todo.

O objeto de estudo desta pesquisa, como mencionado acima, é a variação entre as formas de negação na variedade do PB falada na Cidade de Goiás-GO:

- NEG1: pré-verbal (NEG+V): “eu **NÃO** ando de carro sem colocá cinto...” (SOCIOLINCO38FU-Ana¹);
- NEG2: duplamente marcada, antes e depois do verbo (NEG+V+NEG): “Inf. **NÃO** tinha uma matéria pra cada professor **NÃO**...” (SOCIOLINCO38FU-Ana);

¹ Ocorrências extraídas do corpus gravado pelo Grupo de Estudos e pesquisas em Sociolinguística da Universidade Estadual de Goiás (Sociolinco/UEG/Cora Coralina). A sigla refere-se ao código atribuído a cada informante. Desse modo, “SOCIOLINCO” se refere à amostra de onde o excerto foi retirado, “38” é a idade da informante (na época em que foram realizadas as entrevistas), “F” corresponde a sexo/gênero feminino (e M, a masculino); e U indica Ensino Fundamental. Por fim, segue-se o pseudônimo do falante.

- NEG3: pós-verbal (V+NEG): “aí no outro dia falô vamo eu falo vamo **NÃO**...” (SOCIOLINCO38FU-Ana).

Rocha (2013) que elegeu São Paulo como *locus* de pesquisa, este estudo aponta a predominância da negação anteposta ao verbo na capital paulista. Nascimento (2014) investiga a respeito o mesmo fenômeno na fala da cidade de Vitória-ES. Nesta localidade, observou-se que a ativação da informação no discurso é fator determinante para a variação das formas negativas.

Cavalcante (2009) estuda a negação sentencial em comunidades rurais isoladas no interior da Bahia. Este estudo aponta a mesma tendência do anterior, ou seja, o status da informação no discurso é determinante.

Avelar, Silva e Almeida (2013) tratam da variação das formas de negação no português falado na região metropolitana de Belo Horizonte, na cidade de Santa Luzia. Os resultados obtidos apontam para um favorecimento da dupla negação em contexto em que ocorre o *NUM* pré-verbal.

Diante disso, é notável a necessidade de se compreender melhor esta variação no interior do estado de Goiás, visto que na região centro-oeste não há – ou, até a conclusão desta pesquisa não se soube – estudos que tratem do tema. Cientes dessa necessidade, procuramos explorar o fenômeno de modo a descrevê-lo e contribuir para um melhor entendimento das especificidades do PB.

Diante disso, buscamos constituir um *corpus* para a análise, mas, em 2019, o mundo parou por conta da pandemia causada pelo novo coronavírus, o *Sars-Cov-2*, que chegou no Brasil em meados do primeiro semestre de 2020. Para evitar o contágio, o isolamento e o distanciamento sociais eram fundamentais, o que inviabilizou a realização de entrevistas sociolinguística. Assim, valemo-nos de amostras já constituídas na região para a análise aqui realizada.

Desta forma, tomamos como principal objetivo a descrição das estratégias de negação em amostras de língua oral da cidade de Goiás-GO. Mais especificamente, este trabalho objetiva elucidar se este fenômeno de variação aponta para uma mudança linguística ou se se trata de um fenômeno de variação estável e compreender quais são as variáveis independentes que condicionam a variação de formas.

Diante dos objetivos anunciados, confirmamos que se trata de um fenômeno de variação estável. Além disso, os dados apontam que fatores pragmáticos

discursivos, como a ativação da informação negada – a ausência de marcadores discursivos e o tipo de sequência discursiva – são muito importantes para alternância.

Na seção 1 deste trabalho, apresentamos a perspectiva teórica do estudo aqui realizado, contrapondo a abordagem estruturalista à abordagem variacionista da língua – base teórica deste trabalho. Por conta desta escolha, tratamos nesta seção, também, de questões relativas à concepção de língua adotada, argumentando que, conforme Neves (2018), a língua sofre pressões exercidas pelo contexto de comunicação e essas pressões resultam na variação de formas. Ainda na primeira seção, discutimos sobre o valor de verdade para variáveis sintáticas, questão problematizada por Lavandera (1984). Na sequência, realizamos uma periodização sociolinguística do PB, proposta por Lucchesi (2017). Encerramos a seção tratando da relação entre o PB e as línguas bantas.

Na segunda seção, abordamos o fenômeno de variação das estratégias de negação no PB. Discutimos a constituição do envelope de variação, ou seja, em que tipos de construções a negação pode apresentar o mesmo valor de verdade. Além disso, expomos a coexistência das estruturas negativa no PB. Além de dedicarmos as duas últimas subseções à análise de fatores pragmáticos que influem nesse tipo de construção.

Nesta dissertação, tomamos como objeto a língua falada por uma comunidade de fala do interior do estado de Goiás. Lançamos mão, para isso, de entrevistas coletadas na Cidade de Goiás-GO. Esta cidade, outrora capital, conserva bens culturais valiosos, o que garantiu o tombamento da cidade como Patrimônio Mundial da Humanidade. Logo, abordamos em detalhes aspectos históricos e sociais do município na terceira seção. Além disso, na seção 3, são apresentados as amostras e a metodologia utilizados na análise.

Na seção 4, são apresentadas as variáveis linguísticas e sociais controladas para o tratamento estatístico dos dados. Na seção 5, apresentamos a análise dos dados obtidos. Esta última seção é dividida em duas subseções, uma para a análise quantitativa, ou descritiva, e outra para a análise qualitativa.

1 A SOCIOLINGUÍSTICA VARIACIONISTA

O objeto de estudo que abordamos neste trabalho é o fenômeno de variação da negação sentencial. Observamos, no PB, a ocorrência da negação em três diferentes posições em relação ao verbo: pré-verbal (NEG+V) – NEG1; duplamente marcada, antes e depois do verbo (NEG+V+NEG) – NEG2; e pós-verbal (V+NEG) – NEG3 – a estrutura da negação que tomamos aqui como objeto de estudos é detalhada em 2.1.

Segundo Da Costa Souza (2017, p. 17):

Expressões linguísticas à margem das prescrições gramaticais são passíveis de desenvolver uma outra concepção do funcionamento de elementos linguísticos, levando, dessa forma, a formularem leis ou regras próprias. Ao alcançar, por exemplo, outros estratos sociais da língua, perdem-se, ganham-se e se reformulam muito desses elementos. No que tange à concepção dos elementos negativos, podem-se observar dois sistemas que convivem paralelamente.

Depreendemos daí a ideia de que as formas de negação coocorrem no PB, não caracterizando um processo de mudança propriamente dito, afinal, NEG1 não está sendo substituída por NEG2 ou NEG3. Ocorre a variação entre as formas de negação em contextos específicos, ou seja, a escolha do falante sofrerá influências de variáveis sociais, discursivo-pragmáticas e sintático-semânticas. Na seção 4, apresentamos as variáveis linguísticas e sociais.

Diante do fenômeno de variação entre as formas de negação na variedade vilaboense do PB, trataremos, neste capítulo, do referencial teórico que norteará esta pesquisa. Assim, na primeira subseção, abordamos a perspectiva estruturalista da língua e as rupturas propostas por sociolinguistas. Na segunda subseção, discorreremos a respeito da concepção de língua(gem) aqui adotada, visitando as principais noções que guiarão a dissertação.

Na subseção seguinte, tratamos da questão da variação. A quarta subseção trata do problema relativo ao valor de verdade para as variáveis morfossintáticas. Nas duas últimas subseções, discorreremos sobre o PB: na quinta, tratamos da periodização do PB, proposta por Lucchesi (2017); na sexta, da relação entre o PB e as línguas africanas, trazidas ao Brasil pelos africanos escravizados, que influenciaram a língua falada hoje no país e, provavelmente, as formas de negação de que os falantes do PB dispõem atualmente.

1.1 Estruturalismo e Sociolinguística: rupturas

Ferdinand de Saussure apresenta conceitos-chave para o desenvolvimento do estruturalismo linguístico. O autor dispõe de um método e esclarece muito do que havia sido feito e pensado a respeito da língua anteriormente. Talvez o ponto mais relevante de seu pensamento (ou, pelo menos, das ideias atribuídas a ele) seja o desenvolvimento das famosas dicotomias.

Ao se estudar o “Curso de linguística geral” (SAUSSURE, 2006), deparamo-nos, logo de início, com uma concepção de língua:

Mas o que é língua? Para nós, ela não se confunde com a linguagem; é somente uma parte determinada, essencial dela, indubitavelmente. É, ao mesmo tempo, produção social da faculdade da linguagem e um conjunto de convenções necessárias, adotadas pelo corpo social para permitir o exercício dessa faculdade nos indivíduos. Tomada em seu todo, a linguagem é multiforme e heteróclita; o cavaleiro de diferentes domínios, ao mesmo tempo física, fisiológica e psíquica, ela pertence, além disso, ao domínio individual e ao domínio social; não se deixa classificar em nenhuma categoria de fatos humanos, pois não se sabe como inferir sua unidade. (SAUSSURE, 2006, 17).

Destarte, temos aí alguns indícios do que guiaria o estudo da linguagem, pelo menos, no início do século XX, ou seja, a língua entendida como parte constituinte da linguagem; dotada de caráter social e individual; heteróclita e multiforme; e, além disso, sofre pressões dos domínios físico, fisiológico e psíquico. Depreendemos, então, que a língua é parte constituinte da linguagem humana, esta natural e, aquela, convencional. Em outras palavras, segundo Saussure (2006, p. 92): “[...] A língua é para nós a linguagem menos a fala”. Em seguida o autor aponta para a relação existente entre a língua e o *fato social* e o efeito do tempo.

Ressaltamos, ainda, que o autor seleciona a investigação sincrônica como meio para elucidar “[...] os fatores constitutivos de todo estado de língua”, e continua, “[...] pois é somente pelos estados de língua que se estabelecem as diferentes relações que incubem à gramática” (SAUSSURE, 2006, p. 117).

Vale pontuar que, a este respeito, o autor reconhece o caráter heterogêneo das línguas, ao afirmar que:

Um fato de evolução é sempre precedido de um fato, ou melhor, de uma multidão de fatos similares na esfera da fala; [...] pois na história de toda inovação encontram-se sempre dois momentos distintos: 1.º aquele em que ela surge entre os indivíduos; 2.º aquele em que se tornou um fato da língua, exteriormente idêntico, mas adotado pela comunidade. (SAUSSURE, 2006, p. 115).

Entendemos, diante do trecho acima, que Saussure reconhecia que a mudança linguística era um fato. Nesse sentido, o autor delimita as diferenças entre a sincronia e a diacronia, esta última deve estudar os: “termos sucessivos que se substituem uns aos outros no tempo” (SAUSSURE, 2006, p. 163). Por conseguinte, o autor ainda elege a fonética e a fonologia como ponto de partida para o estudo diacrônico das línguas.

A partir de agora, trataremos das rupturas propostas pela Sociolinguística. Acerca da regularidade absoluta da mudança, podemos, aqui, abordar o que se encontra em “Fundamentos empíricos para uma teoria da mudança linguística” (WEINREICH; LABOV; HERZOG, 2006 [1975]), obra de Uriel Weinreich, William Labov e Marvin I. Herzog: “[...] Saussure, claramente, vê a heterogeneidade dentro do uso linguístico de uma comunidade não como sujeita à uma descrição sistemática, mas como um tipo tolerável de imprecisão de desempenho.” (WEINREICH; LABOV; HERZOG, 2006 [1975], p. 56).

Assim, de acordo com a citação, o que observamos, na proposta saussuriana, é que a variação observável seria subproduto da fala. Seguindo o pensamento de Weinreich, Labov e Herzog (2006 [1975], p. 57), podemos compreender que escapa à proposta de Saussure (2006) a noção de que formas variantes coexistam em dada época, até que as inovações individuais se consolidem na língua.

Em 1.4 discutiremos a questão do valor de verdade a partir das reflexões feitas por Lavandera (1984) e Labov (1978). Interessa-nos, aqui, a ideia que se tem a respeito do fenômeno de variação, principalmente quando se trata de um fenômeno que extrapola o nível fonológico. Assim, o ponto de partida é a existência de mais de uma forma de se dizer a mesma coisa, diante dessa possibilidade o falante optará por uma ou outra forma a partir do estado de coisas, de sua autoidentificação e sua acomodação ao ouvinte (LABOV, 1978, p. 8).

Nesse sentido, já que o objeto aqui analisado se constitui da variação de formas, consideramos a variável “Amostra” (discutida em 4.2.1) para que possamos fazer um estudo de tendência desse fenômeno. Em outras palavras, a variável amostra permite a análise em tempo real do fenômeno – ou seja, podemos averiguar a tendência de usos das formas alternantes em diferentes épocas –, pois, as amostras analisadas (os bancos de dados GEF e Sociolinco²) foram constituídas com um

² Grupo de Pesquisas e Estudos em Sociolinguística da Universidade Estadual de Goiás, Câmpus Cora Coralina, coordenado pela Professora Marília Silva Vieira.

intervalo de 16 anos, a primeira em 2003 (GEF) e a segunda em 2019 (Sociolinco), permitindo observar se há alguma diferença significativa nos períodos. Dessa forma, controlamos essa variável para observar como a variável se comporta de acordo com as variáveis previsoras controladas neste estudo – como o falante decide como vai expressar a negação sentencial de acordo com as restrições observáveis.

Em consonância a isso, acrescentamos a afirmação a seguir: “Ela [a língua] é a parte social da linguagem, exterior ao indivíduo, que, por si só, não pode nem criá-la nem modificá-la [...]” (SAUSSURE, 2006, p. 22). Assim, a fala não poderia afetar a língua, visto que esta é social e exterior ao indivíduo e aquela é individual e intrínseca a cada indivíduo.

Por conseguinte, Weinreich, Labov e Herzog (2006 [1975], p. 57) tratam da questão da “antinomia entre o estrutural e o histórico”, ou seja, não se pode “equiparar a justaposição entre estágios remotos de uma língua com a justaposição de estágios em geral”. Dessa forma, compreendemos que a visão estruturalista da língua – em detrimento da linguagem, geral e heterogênea – caracteriza-se por um caráter homogêneo, visto que não se poderiam comparar estágios da língua – estes resultados da fala.

As limitações da abordagem homogênea da língua influenciaram as ideias relativas à variação linguística. Em Weinreich, Labov e Herzog (2006 [1975], p. 85-86), encontramos:

[...] Argumentamos que, enquanto uma mudança linguística está em progresso, uma forma arcaica e uma inovadora coexistem dentro da gramática: esta gramática difere de uma gramática anterior pela adição de uma regra, ou talvez pela conversão de uma regra invariante numa regra variável. [...] teremos de contar com o fato de que a significação social será, mais cedo ou mais tarde, atribuída à oposição das duas formas. Em algum ponto, as questões sociais e linguísticas se resolvem conjuntamente; quando a oposição já não se mantém, a variante conservadora desaparece.

Entendemos, a partir daí, que há a inserção da noção de heterogeneidade. Existem, na língua, formas que coexistem e exercem a mesma função ou função similar, sendo a criatividade do falante, acrescida de fatores sociais, que determinará a variante a ser escolhida no momento da interação.

1.2 Questões relativas à língua(gem)

Para estabelecermos as bases de análise, é necessário que se estabeleça a concepção de língua(gem). Tomamos como pressupostos teóricos os princípios apontados por Neves (2018, p. 30):

- a. a língua (e a gramática) não pode ser descrita nem explicada como sistema autônomo [...];
- b. as formas da língua são meios para um fim, não um fim em si mesmas [...];
- c. na gramática estão integrados os componentes sintático, semântico e pragmático [...];
- d. existe uma relação não arbitrária entre a instrumentalidade do uso da língua (o funcional) e a sistematicidade da estrutura da língua (o gramatical) [...];
- e. a gramática organiza, em alguns conjuntos, opções de uso, e o falante, procede às suas escolhas fazendo seleções simultâneas [...];
- f. a gramática é suscetível as pressões do uso [...];
- g. a gramática resolve-se no equilíbrio das forças internas e forças externas ao sistema [...].

Considerando estes pressupostos, entendemos que a língua não pode ser compreendida como um sistema fechado em si mesmo, que está pronta e acabada. O falante, em uma situação comunicativa real, lançará mão de seu conhecimento linguístico fazendo escolhas que lhe permitam significar o mundo a sua volta, sempre tendo em vista o interlocutor que se insere em seu contexto imediato.

A este respeito, Labov (1978, p. 8-9, tradução nossa³) afirma:

O sociolinguista vê a linguagem construída sobre uma herança biológica comum. Segue-se que as representações lógicas são construídas com faculdades que foram originadas para comunicar sinais de territorialidade e acomodação. Portanto, o componente variável e suas funções características ainda são bastante proeminentes.

Diferenciando-se, portanto, de uma abordagem formalista, a Sociolinguística considera a língua empregada na comunicação diária. Por isso, o estudo busca entender os fatores que motivam os fenômenos de mudança ou variação. É nesse sentido que Weinreich, Labov e Herzog (2006 [1975]) definem a língua como um “Sistema diferenciado”. Isso é importante para que se possa conceber a sistematicidade da língua e, ao mesmo tempo, sua variabilidade. Sendo assim, para que se possa estudar a gramática, a partir desta perspectiva, tanto *corpora* de língua falada quanto de língua escrita são fontes de evidências.

³ “The sociolinguist sees language built on a common biological inheritance. It follows that logical representations are constructed with faculties that were originated to communicate signals of territoriality and accommodation. Therefore the variable component and its characteristic functions are still quite prominent” (LABOV, 1978, p. 8-9).

Para a compreensão dos fenômenos é necessário ter em mente que a língua apresenta dois subsistemas: gramatical e lexical. O primeiro determina a estrutura da representação cognitiva, ao passo que o segundo determina o conteúdo. Isso é possível, pois as pessoas, segundo Neves (2018), possuem conhecimento conceptual, o que lhes permite entender o mundo a sua volta (real e social/psicológico) e dominar a linguagem. Desse modo, em uma interação real, o falante poderia escolher entre as três estratégias de negação disponíveis no PB:

- (1) NEG1 – “eu NÃO sabia que a origem era bíblica...” (SOCIOLINCO31FF-Camila);
- (2) NEG2 – “NUM tem muito tempo NÃO...” (SOCIOLINCO31FF-Camila);
- (3) NEG3 – “tem vontade de ir embora daqui NÃO...” (SOCIOLINCO31FF-Camila).

Segundo Neves (2011, p. 285) a negação sentencial é um modificador que nega a informação contida em seu “âmbito de incidência”, seu escopo. Ainda segundo a autora, o operador de negação ocorre anteposto ao escopo da negação (NEG1), o que chamamos de forma canônica. Todavia, ao observarmos os exemplos do parágrafo acima, podemos afirmar que a negação sentencial no PB pode ocorrer por meio de duas outras estratégias (NEG2 e NEG3) – em 2.2 aprofundamos esta discussão.

Paralelamente às formas de colocação do operador negativo – que não precisam ser necessariamente o advérbio *NÃO* – no PB, Da Costa Souza (2017) apresenta um levantamento a respeito das estratégias de negação disponíveis no PB e em outras línguas, cujo objeto é a dupla negação pré-verbal: “Um exemplo possível do PB seria: ‘Não, nunca dei e se’u pudesse nem num dava nunca’ [...]” (DA COSTA SOUZA, 2017, p. 15, grifo do autor, nota de rodapé).

A autora faz esse levantamento desde o latim, língua em que a dupla negação poderia significar uma afirmação – “*Nemo non uidet.*” (MARTINS, 1996, p. 179 *apud* DA COSTA SOUZA, 2017, p. 15), no exemplo transcrito, observamos o termo *NEMO* (ninguém) se torna uma afirmação (alguém) por conta da relação sintática estabelecida com o negativo *NON* – em outras palavras, uma dupla negação, no latim, gera uma afirmação. O mesmo não se observa no PB, por exemplo, em uma construção como “Ninguém não viu”, se proferida por um falante nativo do PB, não haveria a intenção de afirmar que ninguém, com o sentido de “nenhuma pessoa”

(Dicionário Houaiss da Língua Portuguesa, 2009, p. 1356), “não viu”. Ou seja, afirmar que “ninguém não viu”, seguindo a lógica que se aplica ao latim, implicaria que “alguém viu”, visto que a dupla negativa significaria uma afirmação. Essa relação semântica não parece encontrar materialidade no PB, visto que quando se diz a sentença acima, a intenção do falante é afirmar que *nenhuma pessoa viu*.

Da Costa Souza (2017, p. 17), ao citar Horn (2010, p. 111-112 *apud* DA COSTA SOUZA, 2017, p. 16), argumenta que:

[...] há um sistema que interpreta que duas ou mais negações se anulam (DNA), enquanto no outro, (DNN) os itens negativos da oração, conjuntamente, validam a negação, seja por meio de NC (negative concord) ou PN (pleonastic negation).

Neste sentido, podemos afirmar que, no PB, há o sistema DNN. A respeito do fenômeno observado em Da Costa Souza (2017), a autora ainda afirma que se podem observar ocorrências da “concordância negativa” (DA COSTA SOUZA, 2017, p. 18) desde o português arcaico. A investigação da autora não trata exclusivamente da negação que se caracteriza pela presença exclusiva do *NÃO*, ela aborda outros operadores negativos.

Ponto relevante é o fato de a autora afirmar que o português europeu (doravante, PE) é inovador no uso da dupla negativa pré-verbal, distanciando-se de outras línguas românicas. Ainda assim, mesmo sendo inovador neste sentido o PE, segundo Schwenter (2005, p. 1429), apresenta apenas a estratégia de negação pré-verbal, ao passo que o PB, como dissemos acima, apresenta três estratégias. Todavia, Amadeu Amaral (1976 [1920], s/p) aponta que Gil Vicente fazia uso da dupla negação sentencial, isso será discutido em 1.6. Amadeu Amaral também apresenta a dupla negativa pré-verbal estudada por Da Costa Souza (2017):

22. O emprego de duas negativas - **ninguém não**, **nem não**, etc., assim contíguas, - vulgar na sintaxe portuguesa quinhentista, mas hoje desusado na língua popular de Portugal, e na língua culta tanto lá como cá, - é obrigatório no falar caipira: *Nem eu num disse - Ninguém num viu - Ninhum num fica*. (AMARAL, 1976 [1920], s/p).

Além disso, a ideia de *continuum* (NEVES, 2018) permite compreender que os fatos da língua se dão de forma gradual. Isso porque a estrutura conceptual e estrutura semântica estão intimamente ligadas, tornando as categorias não estanques. Assim, sabe-se que a estrutura gramatical surge no momento do uso linguístico, fixando padrões apenas por meio da repetição daquilo que é mais

comumente realizado pelos falantes. Sendo assim, a linguagem seria uma forma de o ser humano mediar seu conhecimento a respeito do mundo.

A negação é, notadamente, um domínio marcado. Segundo Bagno (2017, p. 268), a forma *não marcada* “é de uso mais amplo e frequente” do que a forma marcada. A partir disso, segundo Furtado da Cunha (2001, p. 4), entendemos que uma categoria marcada “[...] tende a ser mais complexa (maior) [...]; [...] tende a ser menos frequente [...]; tende a ser cognitivamente mais complexa [...]”. Então, podemos observar no PB as formas de afirmação (não marcada) e negação (marcada).

Ainda nesse sentido, quando Neves (2018) aborda a ativação da gramática explica que, pelo fato de a língua codificar expressões comumente usadas pelos falantes, pressões externas acabam por influenciar o uso. A autora define essas pressões como as categorias “informatividade” e “economia” e apresenta a relação entre elas:

Quadro 1 – Relação entre informatividade e economia

Informatividade	Economia	Fatores
Aumento na forma fônica	Redução da forma fônica	Frequência de uso
Aumento da complexidade do enunciado	Perda de complexidade	Existência de marcas
Maior dispêndio de tempo na enunciação	Redução de tempo na enunciação	Velocidade de processamento
[...] uma relação (mais) direta entre forma linguística e estrutura da experiência	[...] relação mais frouxa entre forma linguística e estrutura da experiência	Grau de iconicidade

Fonte: Adaptado de Neves (2018, p. 147-148).

Então, de acordo com a autora, o falante adaptará seu discurso de acordo com a situação comunicativa em que se encontra, considerando as pressões que ele sofre nesse contexto. Se o falante está em um contexto em que uma forma inovadora tem maior frequência de uso, em que todos a compreendam facilmente, facilitando o processamento cognitivo, se não existirem marcas e não houver a necessidade de uma relação mais direta com o evento, o falante optará por esta forma, a fim de obter economia.

Assim, nas construções *NÃO+SV*, *NÃO+SV+NÃO*, e *SV+NÃO* a diferença observada não se limita à estrutura linguística, havendo, aí, motivações pragmáticas para a escolha do falante. Isso será abordado em 2.3.2.

A variação de formas para a negação pode ser entendida como o redobramento da negação, segundo Castilho (2005, p. 32):

A auto-regulagem (sic) da língua pode ser assim explicada: numa língua existem constituintes que se posicionam no centro da gramática, sendo portanto essenciais, e outros que ficam à margem e são circunstanciais. Esses constituintes circunstanciais poderão representar mudanças muito iniciais, ou seja, quando um constituinte do núcleo da sentença começa a perder traços essenciais que o qualificam, a língua resgata, automaticamente, um outro constituinte que contenha traços bem próximos do primeiro e que está na periferia da língua. Assim, esse elemento periférico é levado para o núcleo da língua, se aproximando, lenta ou rapidamente, daquele constituinte que está se descaracterizando e juntos, numa espécie de reforço, constituem um redobramento, que posteriormente poderá ser desfeito ou não.

Castilho (2005) apresenta alguns exemplos relativos à dupla negação pleonástica no castelhano medieval:

- a) "**nengún non** deve tomar el regno"(sec. XIII-XIV),
- b) "**nengún non** demande"/"**nengún non** osme"/"**nengún non** favle"/"**nengún non** conselle" (sec. XIII-XIV),
- c) "**nunca** omne **non** vio tan fiero abramante"(sec. XIII-XIV), "**nada non** pudo adobar" (sec. XIII); (ALONSO, 1964 *apud* CASTILHO, 2005, p. 38).

Além desses exemplos do Castelhana, a autora ainda trata da dupla negação no português medieval: “[XIII CSM1 193:12] Enton cuidei logo como me partisse/daquesta terra que **neun non** me visse,/[...] [duplicação da negação]” (CASTILHO, 2005, p. 33).

De início, observamos que se trata da dupla negação pré-verbal, tal qual abordado por Da Costa Souza (2017). Esta última apresenta algumas ocorrências da negação simples e dupla no catalão antigo: “entre as simples se encontram \emptyset *V pas* [...], *No V* \emptyset [...]; entre as duplas são apontadas *No V pas* [...] e *No V cap* [...]” (DA COSTA SOUZA, 2017, p. 151).

Da Costa Souza (2017, p. 231) apresenta um quadro em que procura sintetizar as possibilidades de negação no PB de acordo com a “língua estandar, mais próxima à variedade europeia, e a língua coloquial/informal/regional” (DA COSTA SOUZA, 2017, p. 231):

Quadro 2 – Os tipos de negação no Português Brasileiro

	Exemplo	Português Brasileiro estândar (padrão)	Variedades do Português Brasileiro
Negação simples [NEG V]	Ninguém chegou.	Sim	Sim
Dupla negação pós-verbal [NÃO V NEG]	Não chegou ninguém.	Sim	Sim
Negação metalinguística	Chegou uma ova!	Sim	Sim
Dupla negação pré-verbal [NEG NÃO V]	Ninguém não chegou.	Não	Sim
Dupla Negação [NEG V NÃO]	Ninguém chegou, não.	Sim	Sim
Negação enfática [V NADA]	Chegou nada!	Não	Sim
Negação à direita [V NEG]	Chegou não.	Não	Sim
Negação pré-verbal [NÃO V].	Não chegou.	Sim	Sim
Dupla negação [NÃO V NÃO]	Não chegou não.	Sim	Sim

Fonte: Adaptado de Da Costa Souza (2017, p. 231).

Ao observarmos o quadro proposto por Da Costa Souza (2017, p. 231) notamos que NEG3 está presente (“Chegou não.”), mas NEG1 e NEG2 não encontram correlatos – essas formas são possibilidades de negação do PB. Por isso, nessa adaptação, acrescentamos NEG1 e NEG2. Isso se faz necessário, pois, a dupla negação é abordada, mas descartando-se a possibilidade da estrutura [NÃO V NÃO] – com a dupla ocorrência do NÃO – e com o uso da vírgula, como se o negação pós-verbal atuasse como um elemento externo à sentença. Da mesma forma, NEG1 não aparecia representada pela estrutura [NÃO V].

1.3 A questão da variação

O estudo da língua, que propomos aqui, parte de evidências empíricas que permitem a observação da variação que ocorre na fala, por meio de um *corpus* de análise. Trataremos dos meios utilizados para a constituição do *corpus* em 3.2

Afirmamos isso, a partir da solução apontada por Weinreich, Labov e Herzog (2006 [1975], p. 65) para a questão da origem e propagação da mudança:

[...] existe variação contínua dentro do sistema de cada dialeto como um elemento estrutural, correlacionado com algum outro fator linguístico ou extralinguístico, e que o movimento constante de itens de uma classe categorial para outra é parte da estrutura subjacente. (LABOV, 1966 *apud* WEINREICH; LABOV; HERZOG, 2006 [1975], p. 65).

A variação linguística, portanto, se dá por conta da heterogeneidade estruturada, ou seja: “a língua comporta [...], ao lado de regras categóricas, regras variáveis” (COELHO *et al.*, 2015, p. 59). Sendo assim, adotamos aqui o entendimento de que a mudança ocorre no interior das comunidades de fala, por conta de condicionadores linguísticos e extralinguísticos, e que pode ser observada a partir de dados empíricos.

Nesta esteira, Weinreich, Labov e Herzog argumenta que a ideia de que o comportamento linguístico seria determinado pelo aparato cognitivo era bastante comum nas teorias que abordavam a mudança linguística. Assim, a variação seria perceptível aos falantes a partir de traços distintivos. Todavia, evidências apontadas por Weinreich, Labov e Herzog (2006 [1975], p. 67) indicam que a estrutura linguística afeta a percepção, de fato, mas fatores sociais e estilísticos também são influem na percepção do falante.

Segundo Bortoni-Ricardo (2021, p. 57-58): “o alinhamento que assumimos em relações a nós mesmos e às outras pessoas presentes na comunicação” é determinado pela situação sócio-histórica em que o falante se insere no momento da comunicação – como o fato de comunicação ser ou não focalizada. Esse alinhamento é chamado de *footing*.

Assim, segundo Weinreich, Labov e Herzog (2006 [1975], p. 70):

[...] temos evidências empíricas para mostrar que, numa comunidade de fala, a fonologia mais altamente sistemática, a que exhibe mais claramente os processos de evolução linguística, é aquela usada na fala casual com o mínimo de distinções e o máximo de suporte contextual.

Objetivamente, essa afirmação traz à tona a perspectiva de que, na interação social, a capacidade cognitiva dos indivíduos de significar o mundo a sua volta manipula a língua de forma a construir significados, todavia, existem forças que exercem pressão para além do nível da consciência.

Assim sendo, entendemos que “a mudança linguística é um processo contínuo e o subproduto inevitável da interação linguística.” (WEINREICH; LABOV; HERZOG, 2006 [1975], p. 87). A mudança ocorre por conta da variação intrínseca que ocorre no sistema linguístico, isto é entendido a partir da concepção de heterogeneidade ordenada. Assim, segundo Weinreich, Labov e Herzog (2006 [1975], p. 105), um conceito essencial para este modelo é a de variável linguística: “um elemento variável dentro do sistema controlado por uma única regra”.

Alguns fatores podem contribuir para a variação linguística, e consequentes mudanças no sistema decorrentes dela, tal como a urbanização. Bortoni-Ricardo (2021, p. 67-68) demonstra que a urbanização – migração do brasileiro, que vivia na zona rural, para os centros urbanos; processo que, no Brasil, não foi precedido pela industrialização da sociedade – causou a integralização do “caipira” com o restante do país. Por conta disso, passa-se de uma rede social mais resistente aos valores dominantes – seja pela distância, seja pela identidade regional – para relações mais esparsas – com maior integralização dos membros e, em alguns casos, busca pelo padrão de linguagem do grupo de referência a fim de adquirir mobilidade social.

A partir daí, entende-se que a variação sofrerá a influência de variáveis independentes, sejam elas linguísticas ou extralinguísticas. Essa variação pode ocorrer quando os falantes que buscam ascensão social procuram corrigir a sua fala a fim de alcançar o padrão do grupo de referência, com a hipercorreção ou em casos em que os falantes mantêm os traços característicos de uma fala caipira, por exemplo, a fim de reforçar a identidade de pertencimento a um determinado grupo.

Assim, concebe-se a língua como um conjunto de subsistemas caracterizados pela variação, determinados pelo “conjunto de regras co-ocorrentes” (WEINREICH; LABOV; HERZOG, 2006 [1975], p. 108). Neste sentido, Weinreich, Labov e Herzog propõem princípios gerais acerca da mudança linguística:

1. A mudança linguística não deve ser identificada com deriva aleatória procedente da variação inerente na fala. A mudança linguística começa quando a generalização de uma alternância particular num dado subgrupo da comunidade de fala toma uma direção e assume o caráter de uma diferenciação ordenada.
2. A associação entre estrutura e homogeneidade é uma ilusão. A estrutura linguística inclui a diferenciação ordenada dos falantes e dos estilos através de regras que governam a variação na comunidade de fala; o domínio do falante nativo sobre a língua inclui o controle destas estruturas heterogêneas.
3. Nem toda variabilidade e heterogeneidade na estrutura linguística implica mudança; mas toda mudança implica variabilidade e heterogeneidade.
4. A generalização da mudança linguística através da estrutura linguística não é uniforme nem instantânea; ela envolve a covariação de mudanças associadas durante substanciais períodos de tempo, e está refletida na difusão de isoglossas por áreas do espaço geográfico.
5. As gramáticas em que ocorre a mudança linguística são gramáticas da comunidade de fala. Como as estruturas variáveis contidas na língua são determinadas por funções sociais, os idioletos não oferecem a base para gramáticas autônomas ou internamente consistentes.
6. A mudança linguística é transmitida dentro da comunidade como um todo; não está confinada a etapas discretas dentro da família. Quaisquer descontinuidades encontradas na mudança linguística são os produtos de descontinuidades específicas dentro da comunidade, mais do que os produtos inevitáveis do lapso geracional entre pais e filhos.

7. Fatores linguísticos e sociais estão intimamente inter-relacionados no desenvolvimento da mudança linguística. Explicações confinadas a um ou outro A variação fonológica e a variação sintática: o valor de verdade aspecto, não importa quão bem construídas, falharão em explicar o rico volume de regularidades que pode ser observado nos estudos empíricos do comportamento linguístico. (WEINREICH; LABOV; HERZOG, 2006 [1975], p. 125-126).

Esses princípios nortearão a noção de variação e mudança aqui adotada. Compreendemos, então, que a mudança linguística não se dá de forma aleatória e, sim, como resultado de variações que têm origens diversas, forçando uma diferenciação ordenada (1). O conjunto de subsistemas é intrínseco à língua e o falante conhece e controla no momento da interação aqueles que lhe estão à disposição (2).

A variação poderá ocorrer sem que se concretize uma mudança de fato (3) – o caso do fenômeno aqui estudado. A mudança linguística é gradual e pode se estender por um longo período (4). O ponto de partida para o exame da mudança linguística é a gramática da comunidade de fala, nela é que ocorrerão as variações – que podem, ou não, resultar em mudança (5). Assim sendo, a variação ocorre na comunidade de fala, a descontinuidade dela, se ocorrer, também (6). Por conta de (1), (2), (3), (4), (5) e (6), entendemos, à luz de Weinreich, Labov e Herzog (2006 [1975]), que é necessário considerar os fatores linguísticos e os extralinguísticos no exame da variação e mudança linguística.

Segundo Reimann e Yacovenco (2014, p. 34) “No âmbito de uma comunidade de fala, uma variante emerge competindo com as formas vigentes, podendo se estabilizar como uma variação ou iniciar um processo de mudança linguística.”. Dessa forma, é possível que o fenômeno aqui estudado não implique uma mudança linguística, a comparação entre os dados das amostras de 2003 e de 2019 aponta para uma estabilização na variação entre as formas negação – discutimos isso na seção 5.

Consideramos ainda, a esse respeito, três questões que Labov (2008 [1972], p. 313-316) aponta como fundamentais para o estudo da variação de um ponto de vista social: “a. O lugar da variação social”; “[...] b. O nível de abstração”; e “[...] c. A função da diversidade”. Os grupos sociais possuem modos de falar e manipular a língua que seriam característicos e, além disso, o falante apresentaria recursos estilísticos para se adequar ao contexto de interação. Isso significa que, em um determinado contexto social, o falante pode optar por formas diferentes de dizer a

mesma coisa, ambas com o mesmo valor de verdade. Em níveis em que a regra gramatical tem um nível de abstração mais alto, mais distante do nível da consciência, a força dos fatores sociais, que ocasionam a variação e mudança, é relativamente menor. E, terceiro, ao seguirmos o raciocínio de Labov (2008 [1972]), admitimos que a mudança linguística não é uma evolução, como se observa nas teorias a respeito da evolução em biologia, sendo, até mesmo, “disfuncional” (LABOV, 2008 [1972], p. 316).

Bortoni-Ricardo (2021, p. 75-76) afirma que a “Teoria da acomodação”, oriunda da área da Psicologia Social, é um guia importante para a pesquisa em Sociolinguística. Assim sendo, entende-se que um falante ajustará sua fala de acordo com a situação comunicativa à qual está inserido – motivado por “convergência, divergência e complementaridade” (BORTONI-RICARDO, 2021, p. 76). Além disso, há ainda que se “[...] considerar [...] as hipóteses de conflito entre orientação para o prestígio e orientação para identidade [...]” (BORTONI-RICARDO, 2021, p. 77), visto que, de acordo com o movimento de mobilidade social – se ascendente, descendente ou estável –, o falante se aproximará ou se distanciará da norma de prestígio social. Todavia, as amostras aqui analisadas restringiram a coleta a informantes com perfil socioeconômico baixo, o que limita a investigação neste sentido.

1.4 O valor de verdade da variável gramatical

No que se refere à análise de variação de itens gramaticais, Lavandera (1984), em um texto intitulado “*Los limites de la variable sociolingüística*”, questiona a metodologia de Labov. Desta forma, é relevante observar a crítica realizada pela autora:

[...] as unidades de níveis além do fonológico, digamos um morfema, um item lexical ou uma construção sintática, cada uma tem um significado por definição. Não são como fonemas, que por definição não possuem 'referência constante' [...]. (LAVANDERA, 1984, p. 42, tradução nossa).⁴

⁴ “[...] las unidades de niveles más allá del fonológico, digamos un morfema, un ítem lexical, o una construcción sitáctica, tienen cada una por definición un significado. No son como los fonemas, los cuales, por definición no tienen ninguna ‘referencia constante’[...]”.(LAVANDERA, 1984, p. 42).

A isto, Labov (1978, p. 7, tradução nossa⁵, grifos nossos) afirma que: “[...] dois enunciados que se referem ao mesmo **estado de coisas** têm o mesmo **valor de verdade** e seguem Weinreich ao limitar o uso de ‘significado’ a esse sentido”.

Assim sendo, no que refere a esta investigação, não basta analisar simplesmente as ocorrências das formas do **NÃO**, é necessário analisar sua equivalência semântica. O excerto a seguir contém uma ocorrência de NEG1:

(1)foi com meu irmão mais novo... tamém foi coisa passageira que hoje eu intendo que isso aí é bobera... que né? porque antes eu num sabia nada... **NUM** sabia perdoá... hoje eu aprendi a perdoá... hoje eu gosto muito deles né? (GEF33FU-Helena)

Observamos, em (1), uma ocorrência de NEG1: “[...] num sabia perdoá...”. Neste trecho, a informação negada incide em toda a sentença, o argumento externo do verbo está implícito, sendo recuperável por meio da desinência⁶, ao passo que o argumento interno é uma oração completiva verbal de infinitivo.

Para que se possa afirmar, então, que as três formas de negação são variantes da mesma variável é necessário que se possa fazer a substituição delas sem prejuízo semântico. Para NEG2, a ocorrência seria: “[...] num sabia perdoá NÃO...”. Observamos, desta forma, que a negação continua a incidir e “*saber perdoar*” continua sendo negado. O mesmo observado em NEG2 é perfeitamente aplicável a NEG3: “[...] sabia perdoá NÃO...”, visto que a informação continua a ser negada. Esses testes indicam que as formas são, de fato, intercambiáveis sem prejuízos semânticos.

No excerto a seguir há uma ocorrência de NEG2:

(2)[...] tive ele... ele tamém é um minino muito bom... ele já foi muito doente mais agora graças a Deus ele tá bem né? ele mora com minha sogra... **NUM** tá comigo **NÃO** mais... mora com ela... (GEF33FU-Helena - 2003)

Observamos que o conteúdo negado é “*estar comigo*”. Novamente o argumento externo é recuperável pela desinência (\emptyset) e o argumento interno é “*comigo*” – neste caso, assumimos que este pronome pessoal tem o sentido de “12

⁵ “[...] two utterances that refer to the same state of affairs have the same truthvalue, and follow Weinreich in limiting the use of ‘meaning’ to this sense” (LABOV, 1978, p. 7).

⁶ Nesse caso, entende-se, conforme Bechara (2009, p. 235), que a desinência de 1ª pessoa do singular no pretérito imperfeito do modo indicativo é \emptyset , havendo apenas a desinência modo-temporal tal qual P3, que apresentam a mesma forma. Assim, a ausência de uma desinência aponta para P1 e P3. A distinção é feita pelo contexto em que a sentença está inserida.

coabitando com o locutor <ela está c.>.” (HOUAISS, VILLAR, 2009, p. 500). Caso alternássemos o tipo de negação usada, para NEG1, teríamos: “**NUM** tá comigo mais...”. É notável que a negação de “*estar comigo*” se mantém. NEG3 também mantém esse sentido, ou seja, caso o informante optasse por esta forma, seria possível que seu interlocutor compreendesse claramente que a pessoa em questão não reside com ela: “tá comigo **NÃO** mais...”.

Agora, observemos o mesmo teste com uma ocorrência de NEG3:

(3)[...] converso com elas na rua... mais... tem muita intimidade também **NÃO**... (GEF28FF-Renata)

Há, nesta ocorrência, a coordenação de sentenças estabelecida pela conjunção “*mais*”. A declaração “converso com elas na rua...” se opõe a “mais... tem muita intimidade também **NÃO**...”. É notável que a conjunção já dá uma pista de que “ter intimidade” será negado, mesmo que apenas no fim, com NEG3. Outro ponto que chama a atenção é a forma “tem”, visto que sabemos que as formas para verbos das três conjugações P1, no presente do indicativo, é marcada pela desinência -o (e a alomorfa da raiz). Todavia, a forma que está transcrita indica que: ou o transcritor se equivocou no momento da transcrição; ou o informante se valeu do mesmo sistema de outros tempos do indicativo e marcou P1 da mesma forma como marca P3, com Ø.

Para além dessas discussões, observamos que se aplicarmos NEG1 neste excerto o resultado será: “converso com elas na rua... mais... **NÃO** tem[tenho] muita intimidade também...”. Com isso, percebemos que “*ter intimidade*” se mantém como informação negada. O mesmo ocorre com NEG2: “converso com elas na rua... mais... **NÃO** tem[tenho] muita intimidade também **NÃO**...”. Com isso, é perceptível o fato de que as três formas podem ser intercambiáveis em alguns contextos. Ao longo desta dissertação, discutiremos em que contextos essa troca é possível e em quais não, além de observar quais são as variáveis que favorecem uma ou outra forma.

No contexto dos estudos de Língua e Interculturalidade, buscamos averiguar a variação na marcação da negação sentencial na fala vilaboense (relativa aos nativos da Cidade de Goiás-GO) e quais fatores influenciam essas escolhas. Nesse sentido, é salutar recorrer aos questionamentos feitos por Lavandera (1984), aos quais Labov (1978) responde com seu artigo: “*Where does the linguistic variable stop? A response to Beatriz Lavandera*”. O que se destaca desta publicação é o seguinte:

É verdade que nossos primeiros trabalhos se desenvolveram como uma forma de estudar a estratificação sociolinguística. Não percebemos quão grande era o efeito da variação nas regras linguísticas e só gradualmente passamos a perceber o valor total da análise da variação. É óbvio que Lavandera está correta ao dizer que o resultado de uma análise da variação sintática não é em si um achado interpretável. É a explicação das restrições variáveis que nos leva a conclusões sobre a forma da gramática. Quando chegarmos a essas conclusões, não hesitaremos em colocar pesos probabilísticos em nossas regras gramaticais, não importa onde elas ocorram. Há ampla evidência de que a competência linguística humana inclui restrições quantitativas, bem como restrições discretas, e que o reconhecimento de tais restrições nos permitirá construir nossa teoria gramatical sobre a evidência de produção e percepção na vida cotidiana. (LABOV, 1978, p. 13, tradução nossa⁷).

Depreendemos disso que, apesar da crítica da Lavandera acerca dos limites da variável sociolinguística, referindo-se à sintaxe, Labov (1978) indica a possibilidade de lidar com este tipo de variação a partir de um ponto de vista que englobe mais do que fatores linguísticos. Assim, não se avalia a variação somente tendo em vista a estratificação social, é necessário considerar toda a comunidade de fala e as regras variáveis. Como dissemos acima, o falante fará escolhas estilística, depender do contexto interacional em que se insere, e isso pode resultar no aparecimento da variação em certo estilo e em outro não.

Além disso, por mais que se explore a variável em questão na fala vilaboense⁸ é necessário considerar que:

A transformação de falares rurais em variedades urbanas de pouco prestígio [...] está no âmago dos processos de mudança linguística e padronização da língua no Brasil [...] e deve ser pesquisada levando-se em conta certas características sócio-históricas brasileiras a partir do século XX, como a migração de massas populacionais do campo para a cidade, a introdução da tecnologia e dos valores urbanos em áreas rurais e um alto nível de movimentação populacional inter-regional que tem marcado o Brasil como um país com sérios desequilíbrios regionais e uma perversa concentração de renda. (BORTONI-RICARDO, 2021, p. 63).

Ademais, acrescenta-se que fatores pragmáticos influenciam nas escolhas sintáticas dos falantes, isso é discutido mais detalhadamente em 2.3.2.

⁷ It is true enough that our early studies developed as a way of studying sociolinguistic stratification. We didn't realize how massive was the effect of variation in linguistic rules and we have only gradually come to realize the full value of the analysis of variation. It is obvious that Lavandera is correct in saying that the result of an analysis of syntactic variation is not in itself an interpretable finding. It is the explanation of the variable constraints that lead us to conclusions about the form of the grammar. When we reach these conclusions we will not hesitate to place probabilistic weights upon our grammatical rules, no matter where they occur. There is ample evidence that human linguistic competence includes quantitative constraints as well as discrete ones, and that the recognition of such constraints will allow us to build our grammatical theory on the evidence of production and perception in every-day life.

⁸ Gentílico da Cidade de Goiás-GO.

Para além de fatores pragmáticos, é necessário resgatar na história do Português Brasileiro possíveis influências que tenham motivado as três formas de negação aqui investigadas, estabelecendo um paralelo com outras línguas neolatinas, como o francês.

1.5 A periodização da história sociolinguística do PB

Trataremos, agora, da proposta Lucchesi (2017) para uma periodização da história da Língua Portuguesa do Brasil. A divisão da história das línguas tem caráter didático, visto que não se podem estabelecer pontualmente as divisões entre os períodos, pois, segundo Faraco (2005, p. 49):

[...] Do mesmo modo que nunca há um momento definido em que uma língua (metaforicamente chamada de mãe) deixa de ser falada e é substituída por outra (metaforicamente chamada de filha), também não há possibilidade de se dizer que o português arcaico é substituído globalmente pelo português moderno num determinado ponto do tempo.

Todavia, sabemos que uma língua, parte constituinte da cultura de um povo, sofre alterações ao longo do tempo que podem ser constatadas como características de um determinado período. Assim, Lucchesi (2017, p. 350) trata de características pertinentes à periodização do PB e trata, logo de início, de situar sua abordagem fora antinomia “história interna *versus* história externa”, pois a língua não existe fora do contexto histórico e social em que está inserida.

Por conta disso, o autor opta por salientar fatores que foram importantes para a mudança linguística e a caracterização do PB como ele é hoje. É, portanto, uma “Periodização da história sociolinguística do Brasil” (LUCCHESI, 2017). A mudança ocorrida no português, trazido ao Brasil pelos portugueses, deve-se, principalmente, ao contato entre esta língua e, primeiramente, as línguas indígenas, em seguida, as línguas africanas e, por fim, as línguas dos imigrantes – além dos pidgins e/ou as línguas crioulas que surgiram. Dessa forma, Lucchesi (2017, p. 375-376) propõe uma periodização dividida em 4 fases: “Tupinização da costa”; “Multilinguismo generalizado”; “Homogeneização linguística”; e “Nivelamento linguístico”.

A primeira fase da periodização ocorre entre os anos de 1000 e 1532, com a migração dos índios Tupi para o litoral e a região sul do país, sendo o tupi e o tupinambá as línguas predominantes no território brasileiro daquele período.

A segunda fase tem seus limites fixados entre os anos de 1532 e 1695, ocorrendo antes da imposição do português como língua hegemônica e comportando o período em que se inicia e se consolida a escravização dos africanos. Neste período o português sofre pidginização/crioulização no interior e muitas línguas indígenas ainda coexistem.

A terceira fase se dá no período de 1695 a 1930 e é marcada pela imposição do português como língua oficial. Nesse período, há grande contato entre as línguas aqui existentes – português, línguas indígenas e, mais à frente, línguas de imigrantes europeus –, este contato gera, por um lado, algumas variações – por conta da transmissão linguística irregular e o contato dialetal (LUCCHESI, 2017, p. 359). Destarte, a lusitanização da cultura, imposta pela chegada da família real – com a fundação de escolas e faculdades e maior urbanização, não consegue frear as mudanças que estão em curso.

A última fase começa no ano de 1930 e se estende até os dias atuais, caracterizando-se pela sobreposição da norma de prestígio (urbana) às variantes populares e o distanciamento do PE. Esses fatores geram o preconceito linguístico, tema amplamente debatido na academia, principalmente em Bagno (2002).

1.6 O português brasileiro e as línguas africanas

Para que possamos refletir a respeito do PB é, antes de qualquer coisa, refletir acerca da história do Brasil e, mais importante do que isso, a história dos povos que formaram e formam o Brasil. O PB é a língua falada por 98% da população brasileira atualmente (LUCCHESI, 2017, p. 349), faz sentido, então, que se estudem os fenômenos que ocorrem nesta língua.

Todavia, nem sempre foi assim, antes da chegada dos portugueses, como supracitado, o tupi e o tupinambá eram as línguas mais faladas. Depois disso, falavam-se no Brasil as línguas africanas, indígenas e o português, mais a frente, as línguas gerais, depois a língua franca, etc. Esse quadro só muda, de fato, com a imposição do português – imposição de fato, visto que isso não ocorreu de forma pacífica.

Há aí, elementos da identidade cultural que se tentavam construir no Brasil, como uma identidade derivada da identidade nacional portuguesa. Nesse sentido, fundamenta-se tal argumento na ideia de Hall (2006, p. 52-56) acerca da narrativa nacional. Assim, uma narrativa acerca da história de uma determinada nação é

construída, e ela encontra sua continuidade na tradição – esta, inventada – e na “intemporalidade”. A tradição se funda em mito fundacional advindo de um povo puro (original) – neste caso, os portugueses.

Todavia, em que pese essa narrativa nacional, segundo Hall (2006, p. 61-62):

Em vez de pensar as culturas nacionais como unificadas, deveríamos pensá-las como constituindo um dispositivo discursivo que representa a diferença como unidade ou identidade. Elas são atravessadas por profundas divisões e diferenças internas, sendo "unificadas" apenas através do exercício de diferentes formas de poder cultural. Entretanto....: como nas fantasias do eu "inteiro" de que fala a psicanálise lacaniana - as identidades nacionais continuam a ser representadas como unificadas.

Assim, apesar da tentativa de apagar a história dos africanos que vieram ao Brasil e dos indígenas que aqui habitavam, por meio da imposição cultural, as marcas dessa cultura híbrida manifestam-se em fenômenos linguísticos do PB, para além das estratégias de negação verbal⁹, nas religiões, na culinária, entre outros

Ponto importante, a este respeito, é discutido por Lucchesi (2012, p. 50):

Historicamente, essas características mais notáveis da fala popular resultam de processos de mudança induzidos pelo contato do português com as línguas indígenas e africanas, ocorridos nos primeiros séculos de formação da sociedade brasileira. Portanto, essas marcas nada mais são do que o reflexo linguístico do caráter pluri-étnico do Brasil, de modo que o preconceito que sobre elas se manifesta pode ser visto como a expressão mais clara do racismo no plano da língua. Hoje, no Brasil, o racismo é tipificado como crime inafiançável pelo Código Penal, porém ele continua impune quando seu alvo é fala, que pode ser definida como “a segunda pele do indivíduo”.

A partir do que discutido por Lucchesi, podemos entender que o estigma recebido pelas variedades populares do PB é carregado de racismo.

Por conta disso, discutiremos como algumas línguas influenciaram na constituição do PB como é hoje. Segundo Ribeiro (1995, p. 113-114):

Arthur Ramos (1940, 1942, 1946), prosseguindo os estudos de Nina Rodrigues (1939, 1945), distingue, quanto aos tipos culturais, três grandes grupos. O primeiro, das culturas sudanesas, é representado, principalmente, pelos grupos Yoruba - chamados nagô -, pelos Dahomey – designados geralmente como gegê - e pelos Fanti-Ashanti – conhecidos como mircas -,

⁹ Outros fenômenos linguísticos também sofreram a influências dessas línguas, é o caso da concordância verbal e nominal que, segundo Lucchesi (2012, p. 59): “[...] o amplo e profundo quadro de variação no uso das regras de concordância nominal e verbal que se observa hoje nas variedades rurais e populares do português brasileiro constitui o reflexo mais notável do processo de transmissão lingüística irregular de tipo leve que afetou o português no Brasil em função de sua aquisição imperfeita por milhões de escravos africanos e índios aculturados e da socialização e nativização desse modelo defectivo de segunda língua nesses segmentos. E o grau de variação no emprego de tais mecanismos gramaticais podem ser tomados como um dos parâmetros para analisar a norma popular brasileira, distinguindo cada uma de suas variedades, em termos sócio-históricos. As variedades que hoje exibem um grau maior de variação são aquelas que historicamente foram afetadas mais diretamente pelo contato entre línguas”.

além de muitos representantes de grupos menores da Gâmbia, Serra Leoa, Costa da Malagueta e Costa do Marfim. O segundo grupo trouxe ao Brasil culturas africanas islamizadas, principalmente os Peuhl, os Mandinga e os Haussa, do norte da Nigéria, identificados na Bahia como negros malé e no Rio de Janeiro como negros alufá. O terceiro grupo cultural africano era integrado por tribos Bantu, do grupo congo-angolês, provenientes da área hoje compreendida pela Angola e a "Contra Costa", que corresponde ao atual território de Moçambique.

Com isso, entendemos que a chegada dos africanos escravizados não representava a chegada de uma cultura homogênea, os africanos, trazidos à força para o Brasil, vieram de diferentes territórios, cada qual com sua cultura ímpar. Diante da periodização proposta por Lucchesi (2017), é possível dizer que o PB se distinguiu e muito da língua originalmente trazida da Europa há cerca de 500 anos.

São vários os fatores apontados por Lucchesi (2017): contato entre língua; pidginização/crioulização; transmissão linguística irregular; e contato dialetal. Diante disso, os esforços de preservação do português, tal como se encontrava na terra lusitana, não surtiram efeito.

Assim sendo, pode-se afirmar que as línguas indígenas também deixaram sua marca, tanto na gramática do PB quanto na toponímia:

Os bandeirantes, exploradores da vastidão do Brasil, foram outros propagandistas insuperáveis do tupi, porque lhe marcaram um lugar invejável na toponímia brasileira. Seria estulto, portanto, menosprezar uma influência secular. A tendência, porém, a exagerar tal influência em detrimento de outras, não se compreende. (MENDONÇA, 2012, p. 79).

Este, *per si*, já é um fator determinante para a diferenciação observada entre PB e PE. Mas precisamos observar mais profundamente a influência das línguas africanas. Segundo Mendonça (2012, p. 80):

O negro influenciou sensivelmente a nossa língua popular. Um contato prolongado de duas línguas sempre produz em ambos fenômenos de osmose.

Ao lado da contribuição genérica e imprecisa que deu o africano para o alongamento das pretônicas e a elocução clara e arrastada, deixou sinais bem seus nos dialetos do interior, principalmente.

Em contraponto a essa ideia, podemos argumentar que Amaral (1976 [1920], s/p) – contradizendo o que afirma Da Costa Souza (2017), como mencionado em 1.2 – constata a existência de NEG2 na obra de Gil Vicente:

23. Mas há fato mais interessante. A negativa não repetida depois do verbo: não quero não, não vou não, parece puro brasileiro. Encontra-se, porém, repetidas vezes em Gil V., como neste passo:

Este serão glorioso

Não he de justiça, **não**.

(Auto da Barca do Purg.) (AMARAL, 1976 [1920], s/p).

Esse tipo de construção se deve à necessidade comunicativa do falante, conforme Cunha (2001, p. 21):

Uma evidência do caráter originalmente enfático do não pós-verbal é o fato de, na língua escrita, esse elemento geralmente vir separado da construção negativa por meio de vírgula. A vírgula é uma tentativa de representar a pausa da língua falada, que deixa o não fora da fronteira do SV[...]. A pausa, nesse caso, tem uma função icônica: indica que o marcador negativo se aplica a toda a sentença, e não apenas ao V. Logo, a pausa pode ser vista como um recurso prosódico de marcar iconicamente o isolamento relativo do segundo não, isto é, a distância conceitual relativa entre o SV e o reforço negativo.

Esses indícios podem apontar que a origem desta variação não possa ser facilmente recuperada. Todavia, as línguas bantas têm forte influência sobre o PB e, nelas, constata-se a dupla negação. Todavia, conforme Bagno (2016), os estudos a respeito do impacto das línguas africanas no PB se limitaram a aspectos superficiais.

No Brasil não houve uma política estatal de segregação¹⁰, tal qual houve nos Estados Unidos da América (BAGNO, 2016, p. 22-23), mas é preciso reconhecer que:

Embora, no plano social, o Brasil seja um país impregnado de racismo, no plano linguístico as diferenças que separam as variedades urbanas privilegiadas das demais, estigmatizadas, são de ordem socioeconômica: a gramática dos negros pobres e dos brancos pobres é a mesma.

Logo, as variedades mais estigmatizadas são aquelas usadas por grupos que, historicamente, tiveram um contato mais próximo dos escravizados. Sob uma perspectiva gramatical, alguns fenômenos do PB possuem fortes indícios de influência das línguas bantas:

- simplificação da flexão verbal;
- simplificação da flexão nominal (as línguas bantas não classificam as palavras segundo os gêneros);
- omissão (embora não-sistemática) do artigo;
- omissão de preposições (inexistentes nas línguas bantas);
- extensão funcional da preposição em para indicar localização mas também direção (Fui na venda) – as línguas bantas não ‘distinguem’ ‘lugar onde’ e ‘lugar aonde’: ‘para casa’ e ‘em casa’ se diz, indistintamente, ‘ku nzo’ (quicongo), ‘konjo’ (umbundo) e ‘ku bata’ (quimbundo);
- simplificação da flexão pronominal (como vimos acima);

¹⁰ Segundo Bagno (2016, p. 22): “[...] Nos Estados Unidos, a profunda e duradoura segregação racial, que permaneceu amparada pela legislação até a década de 1960, e a virtual inexistência de miscigenação – decorrente, entre outras coisas, da criminalização das relações interracialis por leis que vigoraram também até 1967 –, fez surgir uma língua característica dos guetos urbanos das grandes cidades americanas, língua que já foi denominada BEV (Black English Vernacular), atualmente designada na literatura sociolinguística como AAVE (Afro-American Vernacular English), e que foi batizada de Ebonics (derivado de ebony, ‘ébano’) por intelectuais negros na década de 1970”.

- omissão do pronome pessoal objeto (sintaxe típica do português brasileiro, que nisso difere das demais línguas românicas: Aluguei o filme, ainda não vi [], mas já emprestei [] à Denise, que ficou de me devolver [] no sábado.);
- pronome sujeito em função de objeto (Eu conheço ela; leva eu etc.). (GÄTNER, 2002 *apud* BAGNO, 2016, p. 28)

Acrescentamos a esses fenômenos, com fortes indícios da influência das línguas bantas, (LIPSKI,2008), a dupla negação sentencial.

Por conseguinte, encontramos em Lipski (2008) uma reflexão acerca das aproximações entre o PB e o português africano (doravante, PA). Em se tratando das formas de negação sentencial observáveis, o autor afirma que:

Outra característica do português vernacular brasileiro que não ocorre nas variedades europeias é a negação dupla do tipo *não sei não* e ainda a negação posposta *sei não*. Entre os dialetos hispano-americanos a negação dupla só ocorre em algumas variedades de forte influência africana: na República Dominicana e na região de Choco, Colômbia (SCHWEGLER, 1996). No século XIX também ocorria no dialeto afro-cubano falado pelos bozais (negros nascidos em África). Hoje em dia a negação dupla é frequente no português vernacular angolano [...]. (LIPSKI, 2008, p. 89).

A partir disso, pode-se perceber que há uma correlação entre a existência da variação das formas de negação sentencial e a influência de africanos, visto que este tipo de variação não ocorre no PE e apresenta forte tendência no PB e, segundo o autor, ocorre no português de Angola. Isso pode ser observado em:

ke be sumba ko
NEG Cl. Comprar NEG = 'eles nao compram' (LIPSKI, 2008, p. 91).

De acordo com a notação do autor, *ke* é uma partícula de negação, seguida de *be*, um pronome clítico, o verbo *sumba* (comprar) e outra negação *ko*. Pedro Dias descreve a gramática de uma língua falada em um quilombo do Brasil. A obra trata de uma língua semelhante ao *quimbundo* – língua do tronco *banto*. Nessa gramática, o autor destaca:

Do verbo negativo
Tem o verbo negativo a mesma conjugação que o verbo, Cuzóla, de que fallamos, ao qual acrescentando esta palavra Caná, antes, ou depois do verbo, fica negativo. v.g. Canangazóla, não amo, Canággiba, não matei. Porém posta antes, & depois do verbo, nega com efficacia. v.g. Canángagiba caná, não matei não. (PEDRO DIAS, 1697 *apud* BONVINI, 2008, p. 38).

Bonvini (2008, p. 39), ao citar este trecho da gramática de Pedro Dias (1697), acrescenta:

“Concordância de Gênero”, “passiva”, “dupla negação”... eis o que parecia insólito para um falante de português do século XVII diante de uma língua

africana. No século XX, esse espanto recai sobre o próprio português, ao menos em sua variante popular. (BONVINI, 2008, p. 39).

Dessa forma, verifica-se que há a possibilidade de que alguns fenômenos do PB possam ter tido origem o contato com as línguas africanas. Todavia, não documentação histórica suficiente para que se comprove esta tese. Seguindo esta esteira, mais à frente, Lipski (2008) argumenta que o mesmo fenômeno pode ser observado em São Tomé. Dentre as línguas bantas, além da língua Kikongo, o Kimbundu – língua falada em Luanda, capital da Angola – ganha destaque por apresentar como marcação da negação um prefixo *ki-*. O autor salienta o fato de este tipo de marcação permite a dupla negação:

[...] pode existir a negação dupla: “o prefixo *ki-* é facultativo e depois do verbo coloca-se o pronome disjuntivo das construções de posse: esta estratégia fornece outro modelo para a negação nos dialetos portugueses formados em contato com as línguas da família banto.” (LIPSKI, 2008, p. 91).

Este fato, apontado por Lipski (2008), vai de encontro ao que foi exposto por Ribeiro (1995, p. 114), ao afirmar que: “O terceiro grupo cultural africano era integrado por tribos Bantu, do grupo congo-angolês”. Assim, sabemos que, de fato, houve o contato entre os africanos, que falavam essa língua e foram trazidos ao Brasil, e os falantes de português do país.

Diante disso, percebemos que, além do supramencionado desbotamento semântico do *NÃO* pré-verbal e suas consequências, o contato entre o português e as línguas africanas pode ter influenciado essa variação. Lipski (2008, p. 91) argumenta que as línguas bantas, que ocorrem principalmente na região da Angola e do Congo, apresentam estruturas diferentes para a negação se comparadas ao português.

Pode-se, então, concluir que a relação entre os africanos escravizados no período e o português existente no Brasil, naquele período, somados aos fatores apontados por Lucchesi (2017) – contato entre língua; pidginização/crioulização; e transmissão linguística irregular – foram, aparentemente, determinantes para o surgimento da dupla negação (NEG2) e da negação posposta ao verbo (NEG3).

2 A NEGAÇÃO COMO VARIÁVEL SOCIOLINGUÍSTICA

Nesta seção, apresentamos a estrutura da negação no PB, com base no funcionamento dos advérbios até o nosso envelope de variação. Além disso, apresentamos ocorrências de tripla negação, em períodos complexos. Apresentamos alguns trabalhos que tratam da variação das formas de negação em outras línguas e aspectos pragmáticos que estão envolvidos neste fenômeno.

2.1 A estrutura da negação

Para um melhor entendimento acerca da variação negação no PB, é preciso definir a estrutura gramatical das variantes. A função de negação é exercida por um advérbio ou locução adverbial (ex.: “de forma alguma”). De acordo com Neves (2000, p. 234), o advérbio, de um ponto de vista sintático, é um item periférico, ou seja, “funciona como satélite de um núcleo”. Essa classe de palavras pode ter como núcleo um verbo, um adjetivo, outro advérbio, um numeral, um substantivo, um pronome. Além disso, o advérbio pode atuar sobre uma oração e advérbios modalizados podem atuar sobre todo um enunciado.

Nessa esteira, passamos aos detalhes relativos à negação. Castilho (2012, p. 542) afirma que o advérbio *NÃO* – e o *SIM* – não transferem propriedades ao verbo, sendo operadores de verificação. Neste sentido, Neves (2000, p. 285) afirma que a negação atua nos níveis sintático-semântico e pragmático, constituindo um recurso argumentativo bastante usual. A negação, então, atua sobre um escopo, ou seja, haverá um componente sobre o qual incidirá seu efeito. Este componente pode ser uma oração ou um constituinte – sintagma –, além de aparecer sozinho como resposta direta a interrogações. A negação aqui estudada será a que incide sobre o sintagma verbal.

Realizaremos esta pesquisa de acordo com a metodologia da pesquisa em tempo real – a comparação “dos usos ao longo de uma faixa de tempo” (BAGNO, 2017, p. 455). Para a realização de uma pesquisa que considera a variação em tempo real, lançaremos mão de duas amostras de análise levantadas com o intervalo de 16 anos: a primeira é parte do banco de dados o Grupo de Estudos Funcionalistas, da Universidade Federal de Goiás (GEF/UFG); e a segunda é parte do banco de dados do Grupo de Estudos e Pesquisas em Sociolinguística, da Universidade Estadual de Goiás (Sociolinco/UEG/Cora Coralina).

Portanto, sabemos que a negação pode ocorrer com o próprio NÃO ou com outros advérbios – isso será discutido em 2.2. A negação canônica no PB é a pré-verbal (NÃO+SV), todavia observam-se as ocorrências de dupla negação (NÃO+SV+NÃO) e negação pós-verbal (SV+NÃO). Essas construções emergem na situação comunicativa em que há a negociação de sentidos entre os interlocutores. Nos exemplos abaixo, extraídos da amostra GEF (Grupo de Estudos Funcionalistas, da Universidade Federal de Goiás), podemos verificar essas ocorrências na fala da Cidade de Goiás (GO):

(4) você conheci tamém o ladu da pessoa porque a gente imagina que as pessoas são tudu bonzinhas i **não** é nada dissu... (GEF33FU-Helena) - (Não+SV)

(5) Ele num... **num** deu a vorta por cima **não**... acho que faltou um pôco de Deus na cabeça dele ((risos)). (GEF30MF-Carlos) - (Não+SV+Não)

(6) olha explico... tanto que eu falo pra ela assim... você tem que aprendé a costurá pra você consegui pagá a faculdade... mais não pra ficá como costureira... a não se qui vai costurá pra vendé né... mais pra costurá pro outros **não**... (GEF30MF-Sérgio) - (SV+Não)

No excerto (01), ocorre a marcação da negação com o advérbio anteposto ao verbo; já no excerto (02) há a negação anteposta ao verbo e posterior ao mesmo, marcando duplamente a negação. O excerto (03), por sua vez, apresenta a marcação da negação posposta ao verbo.

A partir disso, considera-se a existência de um envelope de variação com base nessas três variantes, o que resvala no valor de verdade dessas sentenças: é preciso considerar se seria possível alterar o tipo de negação presente nas sentenças sem alterar seu sentido. O excerto (01), por exemplo, poderia alternar entre as formas:

(7)a. você conheci tamém o ladu da pessoa porque a gente imagina que as pessoas são tudu bonzinhas i **não** é nada dissu... (GEF33FU-Helena) – (Não+SV).

b. você conheci tamém o ladu da pessoa porque a gente imagina que as pessoas são tudu bonzinhas i **não** é nada dissu **não**... (GEF33FU-Helena) – (Não+SV+Não).

c. você conheci tamém o ladu da pessoa porque a gente imagina que as pessoas são tudu bonzinhas i é nada dissu **não**... (GEF33FU-Helena) – (SV+Não).

Assim sendo, pode-se entender que NEG1, nesse contexto, pode ser substituída por NEG2 e NEG3.

No excerto (02), há uma ocorrência de NEG2, sendo possível a substituição por NEG1 ou NEG3:

- (8)a.** Ele num... **num** deu a vorta por cima **não**... acho que faltou um pôco de Deus na cabeça dele ((risos)). (GEF30MF-Carlos) - (Não+SV+Não)
- b.** Ele **num**... deu a vorta por cima... acho que faltou um pôco de Deus na cabeça dele ((risos)). (GEF30MF-Carlos) - (Não+SV)
- c.** Ele deu a vorta por cima **não**... acho que faltou um pôco de Deus na cabeça dele ((risos)). (GEF30MF-Carlos) - (SV+Não)

Por fim, no terceiro excerto, observa-se a ocorrência de NEG3 que também permite a substituição pelas outras duas variantes:

- (9)**olha explico... tanto que eu falo pra ela assim... você tem que aprendé a costurá pra você consegui pagá a faculdade... mais não pra ficá como costureira... a não se qui vai costurá pra vendé né... mais pra costurá pro outros **não**... (GEF30MF-Sérgio) - (SV+Não)
- b.** olha explico... tanto que eu falo pra ela assim... você tem que aprendé a costurá pra você consegui pagá a faculdade... mais não pra ficá como costureira... a não se qui vai costurá pra vendé né... mais **não** pra costurá pro outros... (GEF30MF-Sérgio) - (Não+SV)
- c.** olha explico... tanto que eu falo pra ela assim... você tem que aprendé a costurá pra você consegui pagá a faculdade... mais não pra ficá como costureira... a não se qui vai costurá pra vendé né... mais **não** pra costurá pro outros **não**... (GEF30MF-Sérgio) - (Não+SV+Não)

Assim sendo, percebe-se que o advérbio de negação que orbita o sintagma verbal não o antecederá. O fenômeno em questão é abordado em Vitória (ES), por Yacovenco e Nascimento (2016), que apontam para a tendência do uso da dupla negação nessa comunidade linguística. O percentual de ocorrências de dupla negação em tal localidade é superior aos dados das amostras que analisamos aqui: 21,1%. Além disso, as autoras afirmam que a dupla negação é fenômeno comum nas diversas regiões do país, porém, perde força nas regiões mais ao Sul:

Os resultados revelaram que a dupla negação é uma tendência na fala do capixaba, atingindo um percentual de 21,1% dos 2263 dados analisados. Nossos resultados se aproximam dos resultados encontrados nas regiões nordeste - em Natal (20,6%) e Fortaleza (18%) -, sudeste - em Mariana/MG (19,5%), e se afastam dos obtidos na cidade de São Paulo (5,8%) e da região sul - Curitiba (2,6%), Florianópolis (4,4%) e Porto Alegre (0,6%), onde predomina a negação pré-verbal. (YACOVENCO; NASCIMENTO, 2016, p. 136).

Tendo como base o estudo realizado pelas autoras, não se têm informações a respeito das ocorrências de NEG2 – e de NEG3 – na região Centro-Oeste. Assim, evidencia-se a importância do estudo que propomos aqui. A partir da análise das ocorrências, temos um panorama das variantes independentes que favorecem ou desfavorecem as formas não canônicas de negação. Além disso, o fato de estudarmos

este fenômeno em tempo real, por meio de um estudo de tendência, possibilita a avaliação, também, das mudanças ocorridas nesses 16 anos, podemos verificar se algum dos fatores de favorecimento ganhou ou perdeu força no período, configurando um estudo em tempo real.

2.2O envelope de variação

Diante do objeto descrito acima, há a necessidade de se discutir a composição do envelope de variação. Em outras palavras, quais ocorrências encontradas no *corpus* revelam o mesmo valor de verdade e quais não. Diante disso, escolheram-se as variáveis linguísticas para apurar esta variação e, também, as sociais para analisarmos o peso que exercem sobre a variável – isso é apresentado na seção 3.

A princípio, discutiremos a presença de negações com outros advérbios, visto que estes podem ocorrer antepostos ao verbo, como em NEG1, e antepostos ao verbo com um *NÃO* pós-verbal, como em NEG2. Não foram encontradas ocorrências de advérbios diferentes *NÃO* pospostos ao verbo. Os advérbios encontrados são: *NEM* e *NUNCA*.

No excerto a seguir podemos observar a ocorrência do advérbio *NEM*:

(10) INF. não... o meu irmão era muito custoso...
 DOC. Hum...
 INF. Nossa Senhora... já teve uã vez d'eu tê que pulá a janela pá i lá no serviço dela pá chamá ela que ele cortô o pé com caco de vidro...
 DOC. Hum...
 INF. Aí eu tive que pulá a janela pá mim i lá chamá ela pra ela vim aí socorrê ele...
 DOC. Nossa mas o quê que ele fez pra cortá o pé?
 INF. quebrô copo e eu não vi...
 DOC. Ahn...
 INF. aí ele foi e pisô... eu **NEM** lembro quantos anos tinha mas era... pequeno...
 DOC. Pequeno... ahn... aí sua mãe vei correno?
 INF. vei... aí ela veio e levô ele po hospital... (SOCIOLINCO38FU-ANA)

Nesta ocorrência, “eu **NEM** lembro quantos anos tinha mas era... pequeno...”, verificamos que o *NEM* é empregado com sentido similar ao *NÃO* que ocorre em NEG1, todavia não apresenta o mesmo valor de verdade. Caso ocorresse “eu **NÃO** lembro quantos anos tinha mas era... pequeno...” a informação seria negada da mesma forma, mas o *NEM* dá uma ideia restritiva. O falante deixa claro que se lembra de alguns aspectos concernentes a este evento com exclusão da idade, mas se

lembra que era pequeno, ou seja, o *NEM* não nega toda a informação, apenas parte dela (NEVES, 2000, p. 301).

Abaixo há outro excerto, desta vez com o advérbio *NUNCA* em uma ocorrência em que há um reforço negativo – o *NÃO* posposto como em NEG2:

(11) DOC. cê era tranquilo assim que...
 INF. NÃO eu **NUNCA** dava trabai pra es NÃO...
 DOC. hum...
 INF. NÃO...
 DOC. era comportadim né?
 INF. era...

Nessa ocorrência, verificamos a seguinte estrutura: ADV+V+ADV. O segundo advérbio reforça a negação. Se fizermos a substituição por *NÃO* teremos: “eu **NÃO** dava trabai pra es NÃO...”. Fica claro que a informação “dar trabalho” continua sendo negada, mas o valor de verdade não é o mesmo. Segundo Neves (2000, p. 287) o advérbio *NUNCA*, assim como *JAMAIS* – este último não ocorreu nos dados –, “mesclam ideia aspectual e temporal”. Assim, observamos o *NUNCA* tem sentido próximo de “em tempo algum”, sendo mais restritivo do que o *NÃO*.

Por essas razões, delimitamos a análise às ocorrências que contenham apenas o *NÃO* atuando como advérbio de negação. Diante disso, consideramos a variação entre as formas de negação na fala vilaboense e o peso das variáveis selecionadas.

As variáveis independentes serão discutidas na seção 3 deste trabalho, porém, vale ressaltar que uma delas tem maior peso sobre as ocorrências da negação. Trata-se do *status* informacional do discurso:

Passando pelo paradigma de negação, NEG1 a NEG2 para NEG3, então, descobre-se que essas formas mostram um grau crescente de confiança no contexto do discurso. NEG1 é livre para negar expectativas que são estritamente antigas do falante e/ou ouvinte, mas discursivas novas, ou seja, proposições que não foram “acionadas” de qualquer forma pelo conteúdo do discurso em andamento. NEG2, em contraste, requer uma proposição que é ativada pelo conteúdo do discurso, mas essa proposição pode ser inferencialmente derivada de outra proposição do discurso antigo ou mesmo, em alguns casos [...], aquela que é apresentada como se fosse um discurso antigo pelos próprios falantes. NEG3, por fim, é o negador de sentença mais restrito do paradigma, [...] na medida em que se restringe a negações de proposições que são ativadas diretamente no discurso em curso.¹¹ (SCHWENTER, 2005, p. 1452, tradução nossa).

¹¹ Moving through the paradigm of negatives from NEG1 through NEG2 to NEG3, then, one finds that these forms show an increasing degree of reliance on the discourse context. NEG1 is free to deny expectations that are strictly speaker- and/or hearer-old but discoursenew, i.e., propositions that have not been “triggered” in any way by the content of the ongoing discourse.²⁰ NEG2, in contrast, requires

Observamos que, na pesquisa desenvolvida por Schwenter (2005), os dados revelaram que NEG1 pode ocorrer em todos os contextos discursivos, seja com uma informação nova no discurso, inferível ou ativada diretamente, ao passo que NEG2 exige que a informação seja inferível ou ativa e NEG3 exige uma informação diretamente ativada discursivamente.

No excerto a seguir, podemos observar NEG1 em uma ocorrência em que a informação negada é nova no contexto de fala:

(12) DOC. Ele tá morano lá por quê?

INF. Uai porque ela dá... ela é uma pessoa sozinha né? ela mora sozinha... assim com Deus e ela né? e ela resolveu í pra lá i tem sete ano qu/ela mora lá

DOC. Lá onde?

INF. I toda vida... Uruaçu... aí toda vida ele... ela tem... assim tentando que a gente dexa ele morá com ela né? aí a gente nunca dexô porque é ruim demais ficá longe dum fii... né? aí eu peguei... de tanto conversá com meu esposo ele tamém foi intendeno a gente entrô em acordo e dexô ele morá lá com ela porque ele já tem quinze anos aqui em Goiás **NÃO** tem serviço que... qui dá pra ele assim... que dá pra ele trabaia é difici dimais... (GEF33FF-Joana)

Observa-se que o estado de coisa “ter serviço em Goiás” não aparece no excerto, não sendo ativado nem pelo documentador nem pela informante, também não há menção que pudesse tornar essa informação disponível por inferência, assim, podemos afirmar que se trata de uma informação nova.

Nesse caso, de acordo com a proposição de Schwenter (2005), NEG2 e NEG3 não poderiam ocorrer – caso ocorressem, poderiam até ter o mesmo valor de verdade, mas não ocorreram nos dados observados pelo autor. A seguir, apresentamos um excerto em que a informação negada é inferível:

(13) DOC. Seu nome é M. I. né?

INF. É...

DOC. É M. é um pouco mais comum... mais I. é diferente né? quem que escolheu?

INF. Esse nome quem... escolheu pra mim foi minha vó

DOC. Foi?

INF. Minha vó D.

DOC. Por quê?

a proposition that is activated by discourse-old content, but this proposition may be one that is inferentially derived from another discourse-old proposition or even, in some cases [...], one that is presented as if it were discourse-old by speakers themselves. NEG3, lastly, is the most restricted sentence negator in the paradigm, [...] insofar as it is restricted to denials of propositions that are directly activated in the ongoing discourse. (SCHWENTER, 2004, p. 1452).

INF. Ah... NUM sei o porque NÃO qu/ela escolheu NÃO...foi ela que escolheu meu nome...
 DOC. As veis porque já tinha alguém na família né?
 INF. É... NÃO... **NUM** mim lembro **NÃO** que nossa família é muito grande tem uns parente qu/eu nem conheço... (GEF33FF-Joana)

No excerto acima, o documentador demonstra interesse a respeito do nome da informante, ressaltando o fato de que não se trata de um nome comum. A informação duplamente negada (NEG2) é “lembrar”, visto que a informante afirma não se lembrar se há alguém na família que possua o mesmo nome que ela. A informação é inferível, visto que é possível, a partir do contexto, dizer qual é o objeto de “lembrar”.

Por fim, apresentamos um exemplo de ocorrência em que a informação fora diretamente ativada no discurso:

(14) DOC. ((risos)) Tá certo... e aí foi crescendo... é... começô a namorá:::?
 INF. É foi cresce:::no... sim...namorá nós **NUM** namorô mui:::to rápido **NÃO** né? que::: nós era sim mei vergonhosa quais num saía né (GEF28FF-Renata)

No excerto acima, a informação “namorar” é diretamente ativada pelo documentador em seu turno de fala: “começô a namorá:::?”. A informante, então, retoma a informação e, ao repeti-la, nega-a – lançando mão de NEG2. É importante mencionar que ocorrências como (15) e (16) constituem nosso envelope de variação, visto que, sendo a informação diretamente ativada ou inferível no contexto, é possível a ocorrência de NEG1, NEG2 e NEG3. Por outro lado, conforme supramencionado, ocorrências como (7) não comportam a alternância entre as formas estudada aqui.

Assim, em Schwenter (2005), NEG1, NEG2 e NEG3 serão variantes da mesma variável somente em contexto em que a informação tenha sido diretamente ativada. Da mesma forma, NEG1 e NEG2 constituirão uma variável quando a informação for, pelo menos, inferível.

Dessa forma, Serra (2018, p. 37), citando Schwenter (2006) avalia que, na realidade, os fatores pragmático-discursivos têm maior peso no uso variável das estratégias de negação no PB do que os fatores sociais. Por essa razão, controlamos o *status* informacional do discurso para verificar se, na amostra vilaboense, as ocorrências se distribuirão da mesma forma.

2.2.1 A tripla negação

Durante a classificação dos dados, encontramos uma ocorrência que, à primeira vista, fora classificada como sendo duas ocorrências: uma de NEG2 e uma de NEG3. Todavia, ao se observar mais atentamente, verificou-se que se tratava de um mesmo período:

(15) [...] era uma pessoa assim hora das... das encrenca dela ela sabia encrencá fazê as encrenca mais na hora dela conversá com a gente ela era muito de conversá... ele era uma pessoa boa... a gente tem esses momento mesmo di... di bobera né? di... di briga di encrenca... mais ela era uma pessoa boa... minha mãe meu pai... eu **NUM** tenho nada de regramá **NÃO** da minha infância **NÃO**... foi uma infância boa... (GEF33FF-Joana)

Antes de se analisar a estrutura da ocorrência acima, devemos fazer algumas observações relativas ao *status* da informação negada acima. Primeiramente, este excerto foi extraído de um longo trecho narrativo da entrevista, são quatro páginas de transcrição para esta sequência e o excerto (17) está no fim da terceira página. A informação negada *não ter nada de reclamar* foi ativada na primeira página deste turno de fala da informante – e, no momento da ativação, negada por meio de NEG1. Isto é suficiente para que possamos afirmar que se trata de uma informação velha no discurso.

Além disso, é inferível a partir do contraste estabelecido entre a “hora das... das encrenca” e a “hora dela conversá” que a conclusão seria essa. Entendemos tendo por base o que afirma Schwenter, valendo-se das proposições de Prince (1992 *apud* SCHWENTER, 2005, p. 1435):

A concepção de "discurso" pretendida no modelo de Prince é ampla, abrangendo não apenas o "co-texto" ou o material linguístico real produzido no discurso, mas também o "contexto", incluindo características situacionais tais como o cenário e os participantes. (Tradução nossa¹²).

Com isso, verificamos que a análise não deve se basear apenas no material linguístico, mas considerar aspectos situacionais envolvidos no momento da comunicação. Para os participantes da comunicação a informação pode ser nova ou velha e, da mesma forma, a informação pode ser nova ou velha no discurso

¹² The conception of “discourse” intended in Prince’s model is a broad one, encompassing not only the “co-text” or the actual linguistic material produced in discourse, but also the “context”, including situational features such as the setting and the participants.

(SCHWENTER, 2005, p. 1434). Assim, quando a informante nega a informação em (10) ela se vale de NEG2, pois ela foi ativada anteriormente no discurso.

A partir daí, concordamos com as proposições de Schwenter (2005, p. 1434), ao afirmar que NEG2 é sensível ao *status* discursivo da informação negada, mas não ao *status* da mesma informação para os interlocutores envolvidos, ou seja, mesmo que para a informante a informação fosse nova, ela só poderia ser duplamente negada se fosse direta ou indiretamente ativada no discurso. Como dissemos acima, essa informação é ativada antes no mesmo turno de fala e negada com NEG1 – este tipo de negação, segundo Schwenter (2005, p. 1452), pode negar tanto informações novas quanto as que foram ativadas direta ou indiretamente.

Observa-se que: “eu **NUM** tenho nada de recramá **NÃO**” já é um período fechado, que dá uma informação completa, mas, considerando-se o contexto, a falante sentiu a necessidade de enfatizar a informação, a fim de deixar claro que não aconteceu nada durante sua infância que a fizesse “reclamar”, por isso, acrescenta: “da minha infância **NÃO**...”. A construção poderia aparecer da seguinte forma: “eu **NUM** tenho nada de recramá da minha infância **NÃO**”, ou ter sido encerrada como no primeiro recorte deste parágrafo, mas há a continuação.

O excerto em questão é constituído por uma oração principal em que o verbo *TER* seleciona os argumentos – A1: *EU*; A2: *nada de recramá NÃO da minha infância NÃO*. Ocorre que, para complementar o A2, que possui como núcleo o pronome indefinido *NADA*, há uma oração subordinada adjetiva reduzida de infinitivo. Essa oração subordinada é constituída pelo verbo *RECLAMAR* que, segundo o Dicionário Houaiss da Língua Portuguesa (2009), pode ser transitivo indireto e intransitivo quando assume o sentido de “opor-se por meio de palavras; fazer reclamação; queixar-se” [...].

Diante dessas duas possibilidades de transitividade, a informante iniciou o período com a negação com a forma *NUM*, selecionou o verbo e seu complemento e, depois dele, reforçou a negação com outro *NÃO*, todavia, sentiu a necessidade de acrescentar informações a respeito da relação que tinha com os pais durante a infância e mudou a transitividade do verbo da oração subordinada, dando-lhe um complemento e, por consequência, um reforço negativo.

2.3 A coexistência das três formas de negação no PB

A negação sentencial é objeto de estudos em todas as regiões do Brasil, de Norte a Sul. Assim, trataremos, nesta seção, da revisão da literatura a respeito do fenômeno em diversas localidades e sob diversas orientações teóricas. Vale ressaltar que, para nosso trabalho, a corrente teórica que norteará o trabalho será a Sociolinguística Variacionista.

2.3.1 A negação nas diversas comunidades de fala brasileiras

De início, trataremos, aqui, do trabalho de Rocha (2013) que elegeu São Paulo como *locus* de pesquisa. A princípio, ganha destaque o fato de NEG3 não ter ocorrido o suficiente para ser considerada na análise – apenas 4 ocorrências de um total de 5607. Este estudo abrange diversas regiões da cidade de São Paulo – vale ressaltar que se trata da mais populosa da América Latina.

Diante disso, o autor mostra que os fatores sociais que propiciam a ocorrência de NEG2 são: “região periférica da cidade”; “nível de escolaridade: não superior”; “informantes mais velhos”; “Paulistanos de 1ª geração (filhos de migrantes)” (ROCHA, 2013, p. 69).

Os fatores linguísticos controlados pelo pesquisador foram a presença ou ausência de marcadores conversacionais e outro termo negativo na sentença e a ativação da informação negada, se direta ou indireta. Os resultados por ele obtidos indicam que a presença de marcadores conversacionais e de outros termos negativos na sentença desfavorece a ocorrência de NEG2 e a ativação direta da informação negada favorece esta ocorrência – ao passo que informações inferíveis a partir do contexto favorecem a ocorrência de NEG1.

Ainda na região sudeste do país, encontramos no trabalho de Nascimento (2014) uma investigação a respeito do mesmo fenômeno na fala da cidade de Vitória-ES. Assim, observamos, a partir dos resultados obtidos, um favorecimento para NEG2 e NEG3 em contexto em que a informação tenha sido diretamente ativada e em ocorrência onde a informação negada é nova no discurso sempre ocorreu NEG1.

Um fator que desfavorece NEG2, apontado por Nascimento (2014), é a presença de reforço negativo na sentença ou de marcadores conversacionais, quando isto ocorre a negação é marcada apenas uma vez. E, de acordo com a autora, o fator social que se demonstrou relevante para o fenômeno foi a variante sexo/gênero, visto que os homens favorecem a ocorrência de NEG1.

Além disso, vale ressaltar a relação existente entre o PB e as línguas de origem africana. O fato de o Brasil ser uma antiga colônia de Portugal gerou o contato entre diversas línguas, como apontado em 1.6. Além das línguas dos indígenas que aqui viviam houve, também, o contato com as línguas dos africanos que foram tirados de suas nações e escravizados no país. Esse contato interferiu diretamente na constituição do PB.

Cavalcante (2009) estuda a negação sentencial em comunidades rurais isoladas que são constituídas, principalmente, por afrodescendentes. O dialeto observado no estudo apresenta ocorrências das três formas de negação que são objeto deste estudo.

O autor aponta que a frequência de ocorrência de NEG1 é de 66%, NEG2 possui frequência de 28% e 6% para NEG3 para um *corpus* de 2026 ocorrências. Além disso, as variantes NEG2 e NEG3 são favorecidas quando a informação negada foi ativada discursivamente, ao passo que, em contexto de perguntas introduzidas por um pronome interrogativo, são agramaticais. Não obstante, NEG3 é agramatical em orações encaixadas ou orações que tenham seu complemento topicalizado. Por fim, as variantes sociais favorecem NEG1, em detrimento a NEG2, já NEG3 não muda de acordo com a alternância das variáveis sociais controladas.

O estudo realizado por Avelar, Silva e Almeida (2013) trata da variação das formas de negação no português falado na região metropolitana de Belo Horizonte, na cidade de Santa Luzia, conforme a tabela a seguir:

Tabela 1 – Ocorrências das estratégias negativas em Santa Luzia - MG

Tipos de Estratégia	Ocorrências	Porcentagem
[Não V]	163	91,6%
[V Não]	1	0,6%
[Não V Não]	14	7,8%
Total	178	100%

Fonte: Adaptado de Avelar, Silva e Almeida (2013, p. 32).

O objeto investigado mais profundamente, no trabalho que aborda a variedade mineira do PB, foi a alternância entre as formas plena [NÃO] e reduzida [NUM] do advérbio *NÃO*. Assim, verificou-se que, para a NEG1, não há uma diferença significativa entre as formas – 79 ocorrências para a forma plena e 84 para a forma reduzida. Quando se trata de NEG2, a diferença é mais expressiva, já que, em 12

ocorrências, a forma reduzida foi a escolhida – 85,7% das vezes –, na medida em que apenas em duas 2 ocorrências a forma plena foi selecionada – 14,3%.

2.3.2 Possíveis motivações para a variação entre as formas de negação: a abordagem funcional

Cunha (2001) sustenta, em sua pesquisa, que a variação observada no domínio da negação é motivada por razões pragmáticas, ou seja, a economia e a informatividade exercem pressão nos contextos comunicativos, levando os falantes, muitas vezes, a se valer de uma forma inovadora, preterindo a forma canônica.

A referida pesquisa constatou que, primeiramente, as estratégias não canônicas de negação não apareceram em textos escritos, mas são intercambiáveis em textos falados: para pessoas que concluíram apenas o ensino fundamental NEG1 foi utilizada 293, enquanto que NEG2 foi utilizada 67 e NEG3 foi utilizada 5 vezes; para informantes que concluíram o ensino médio, NEG1 foi a escolhida em 508 vezes, NEG2 apareceu em 52 e NEG3 ocorre apenas 2 vezes; já para os entrevistados com ensino superior completo, NEG1 ocorreu 497 vezes, NEG2 39 e NEG3 2.

Podemos constatar, a partir do quadro de distribuição acima, que os informantes com escolaridade básica (até a 8ª série do ensino fundamental) empregam mais NEG2 e NEG3 se comparados aos estudantes que concluíram o ensino médio e os que cursaram o ensino superior. Apesar disso, esse tipo de estratégia não aparece em textos escritos.

A partir dos resultados de Cunha (2001), é possível afirmar que essa variação ocorre por conta de pressões sofridas nas interações sociais características da modalidade oral de comunicação, sendo bastante raras as ocorrências de NEG2 e NEG3 em textos escritos – o que sustenta a hipótese de que a variação ocorre por pressões pragmáticas. Nesse sentido, a autora aponta um caminho para o surgimento das formas inovadoras de negação:

Esquema 1 – Estágio de reanálise da negação

[[não + SV] não]	>	[não + SV + não]	>	não [SV + não]	>	[SV + não]
1º estágio		2º estágio		3º estágio		4º estágio

Fonte: Adaptado de Cunha (2001, p. 20).

Dessa forma, entendemos que por conta de pressões pragmáticas por economia houve o desgaste fonológico de NEG1: não > num. Segundo Cunha (2001, p. 19):

Por um lado, no discurso falado rápido verifica-se uma tendência para a redução do signo. Por outro lado, quando um item (ou construção) é frequentemente repetido, ele perde sua força expressiva. A repetição frequente resulta, por economia, no enfraquecimento desse item, o que motiva, por sua vez, uma busca por redundância.

Isso pode ser observado em:

(16) Inf. [...] ainda NUM tinha energia durmia a luis de vela () lamparina né? pegava água nus vizim... eu já passei muitas dificuldade... é... aí logo a gente meu sogro né? seu () também já é falecido ele já faleceu... juntô mais minha sogra e comprô as teia... (GEF33FF-Joana)

No excerto acima, fica evidenciado o desgaste fonológico. Esse tipo de desgaste, em situações de fala, acaba por fazer o operador de verificação pareça mais fraco, diminuindo o nível informacional do mesmo. Por isso, o falante busca a redundância de marcação. Essa redundância gera NEG2, NUM+SV+NÃO. Segundo a autora, em um primeiro momento esse segundo NÃO tinha como escopo toda a sentença, sendo separado por vírgula na escrita. Porém, com o tempo, ele se aproximou do sintagma verbal, gerando NEG2.

Nesse segundo estágio, há a prioridade informacional, visto que o falante opta por esta construção a fim de garantir que seu interlocutor entenderá o que está sendo dito. Este tipo de construção pode ser observado em:

(17) Inf. [...] tem meu esposo tem minha casa graças a Deus... tem o que cumê graças a Deus... tem roupa tem calçado... as veis **NUM** pode sê do melhó **NÃO** mais eu tenho graças a Deus e gosto muito dessa vida qu/eu levo e tem fé que um dia meu esposo vai tá junto comigo na igreja com meus filho... (GEF33FF-Joana)

De acordo com Cunha (2001), NEG3 também surge do desgaste fonológico sofrido por NEG1:

A redução fonológica do não pré-verbal tem duas consequências relacionadas. Primeiro, o marcador pós-verbal enfático é reanalisado como um constituinte da sentença negativa. Segundo, dado o seu enfraquecimento fonológico, o próprio marcador pré-verbal é reinterpretado como um elemento opcional, o que leva à emergência da construção SV + não. (CUNHA, 2001, p. 20).

No estágio 3, observa-se o afastamento do *NÃO* pré-verbal do sintagma nominal o que acaba desembocando no 4º estágio que consiste no apagamento total dele, mantendo-se apenas NEG3. Pode-se observar isto em:

(18) Inf. [...]Inf. Fiis uma festinha até boa... festa como podia fazê né? já tava começano ficá brefado fazê quando eu casei aí... ô fui só casano passô::: casamento foi muito bom pro oro *NÃO* que o oro sumiu de mim virô um setenta que fiquei brefado muito tempo (GEF38MF-Sérgio)

No entanto, a autora argumenta que, em seus dados NEG3, ocorre somente como resposta a perguntas diretas, o que não é verificado no trecho exposto acima e em outras ocorrências no *corpus*. Assim sendo, nota-se uma diferenciação do que foi observado pela autora em Natal-RN e o que se observa em Goiás-GO.

De acordo com a reflexão exposta aqui, nota-se que o domínio da negação pode ser ocupado por 3 construções diferentes: pré-verbal (*NÃO*+SV); dupla negação (*NÃO*+SV+*NÃO*); e negação pós-verbal (SV+*NÃO*). A escolha de uma dessas alternativas se dará no momento da interação verbal por conta de pressões externas, como o fato de a informação ter sido diretamente ativada no discurso. Essa alternância de formas gera o que é chamado sincronicamente de variação. Essa variação se dá por razões diacrônicas, visto que o surgimento de novas formas para funções já existentes não ocorre abruptamente.

Apesar de se tratar de uma variável morfossintática, devem ser consideradas, para a análise das formas variantes de negação, alguns fatores de ordem discursivo-pragmática, como descritos a seção a seguir.

2.3.3 Uma abordagem discursivo-pragmática do fenômeno

A pesquisa de Lima e Silva (2016) apresentam uma abordagem da negação a partir de uma perspectiva da pragmática. Assim, para a constituição do *corpus*, o autor deixa claro que ele é composto por amostras da língua espontânea, descartando-se as ocorrências de fala em que o que é dito tenha sido previamente elaborado. Em outras palavras, para a Teoria da Língua em Ato, a unidade o ponto de partida para a análise são os atos de fala (LIMA E SILVA, 2016, p. 79).

Lima e Silva (2016) desenvolveram sua análise com base em um *corpus* de 3114 ocorrências, das quais 2262 (72,63%) eram de NEG1, 704 (22,62%) eram de NEG2 e 148 (4,75%) de NEG3. O autor afirma que: “a negação dupla possa conter

traços funcionais tanto da negação pré-verbal (livre em unidades informacionais textuais) e da pós-verbal (restrita a unidades ilocucionárias).” (LIMA E SILVA, 2016, p. 166). Além disso, as variantes sociais selecionadas para o estudo não apresentaram peso significativo para a variação entre as formas de negação.

Um fato relevante, apontado por Lima e Silva (2016), é que, em seu estudo, foram encontradas ocorrências de negação não canônica que se diferenciam das que foram apontadas anteriormente em outros estudos: para NEG2, foram encontradas ocorrências em orações coordenadas e em orações infinitivas; NEG3, segundo o autor, ocorreu em orações que possuíam o sujeito explícito e em orações encaixadas, além de ocorrer em trechos onde a informação negada não havia sido ativada discursivamente.

Assim, de acordo com os apontamentos do autor, resulta de seu estudo a seguinte hipótese:

A proposta lançada aqui é a de que a negação pré-verbal nega a proposição, a pós-verbal nega uma explicatura/implicatura e a negação dupla pode negar tanto uma explicatura/implicatura, quanto a proposição (no caso de ocorrer em unidades de PAR). (LIMA E SILVA, 2016, p. 166).

A partir disso, verificaremos em nossos dados de essa hipótese se confirma também na fala vilaboense, ou se, nessa variedade de fala, não encontraremos tais ocorrências.

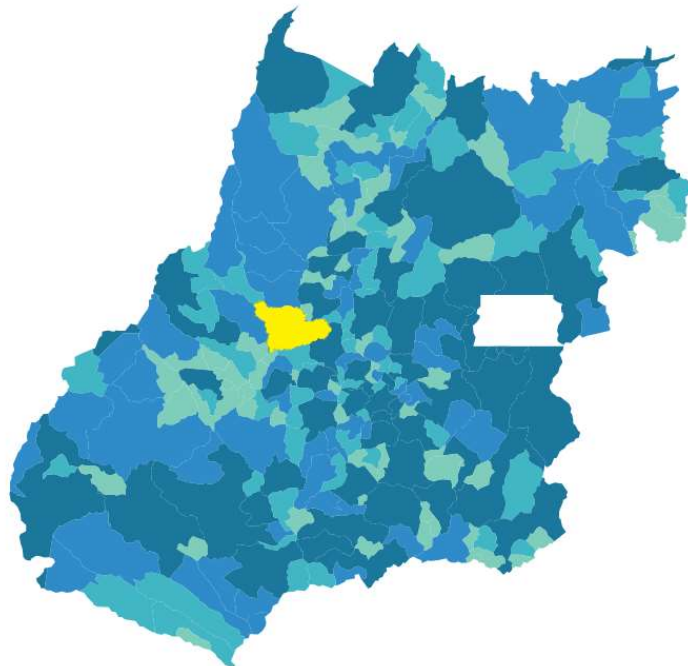
3. MATERIAIS E MÉTODO

Nesta seção, fazemos uma breve apresentação do município de Goiás-GO, *locus* de nossa pesquisa. Em seguida, em 3.2, serão descritas as amostras de fala analisadas neste estudo. A escolha das amostras já constituídas se deu devido a impossibilidade de realização de entrevistas sociolinguísticas durante o período de pandemia da *Covid-19*. Encerramos a seção apresentando o método e os critérios que usaremos para a análise dos dados.

3.1 A cidade de Goiás – aspectos históricos e socioeconômicos

A cidade de Goiás-GO é a antiga capital do estado homônimo. Localiza-se a 148 km¹³ da atual capital do estado, Goiânia. O município está inserido na Mesorregião Noroeste do estado¹⁴. Na imagem abaixo, é possível visualizar a localização do município:

Figura 1: Localização da Cidade de Goiás-GO



Fonte: Site do IBGE (Disponível em: <https://cidades.ibge.gov.br/brasil/go/goias/panorama>. Acesso em 27 de abr. de 2021).

¹³ Disponível em: [https://pt.wikipedia.org/wiki/Goi%C3%A1s_\(munic%C3%ADpio\)](https://pt.wikipedia.org/wiki/Goi%C3%A1s_(munic%C3%ADpio)). Acesso em 25 de abr. de 2021.

¹⁴ Disponível em: https://www.imb.go.gov.br/files/docs/mapas/mesorregioes-ibge/meso_noroeste.pdf. Acesso em 25 de abr. de 2021.

De acordo com o Instituto Brasileiro de Geografia e Estatística, a cidade de Goiás-GO apresenta os indicadores a seguir:

Tabela 2 - Informações por Cidades e Estados - Instituto Brasileiro de Geografia e Estatística

Goiás	
Gentílico	Goiano
Área Territorial	3.108,020 km ² [2020]
População estimada	22.381 pessoas [2020]
Densidade demográfica	7,96 hab/km ² [2010]
Escolarização 6 a 14 anos	98,2 % [2010]
IDHM (Índice de desenvolvimento humano municipal)	0,709 [2010]
Mortalidade infantil	21,43 óbitos por mil nascidos vivos [2017]
Receitas realizadas	77.228,05746 R\$ (×1000) [2017]
Despesas empenhadas	74.785,93658 R\$ (×1000) [2017]
PIB per capita	20.136,77 R\$ [2018]

Fonte: Site do IBGE (Disponível em: <https://www.ibge.gov.br/cidades-e-estados/go/goias.html>. Acesso em 27 de abr. de 2021).

Observa-se, à primeira vista, que o IDH do município é de 0,709, ao passo que a média do mesmo índice para o estado de Goiás é de 0,735. Outro dado que se destaca, ainda um indicador do desenvolvimento humano da cidade, é a escolarização entre 6 e 14 anos, visto que o município alcançou 98,2% no ano de 2010.

O Produto Interno Bruto (PIB) *per capita* é relativamente alto, R\$ 20.136,77. Todavia, de acordo com o IBGE, 88,5% das receitas do município são de fontes externas. Além disso, a de se considerar que esse valor do PIB não reflete uma realidade de distribuição de renda equânime.

Borges (2010, p. 27) pontua que o município possui parte de sua economia advinda do comércio, da agropecuária e do turismo. Este último se destaca pelo valor cultural, visto que o município foi tombado como Patrimônio Histórico pela UNESCO (Organização das Nações Unidas para a educação, a ciência e a cultura). Segundo Borges (2010, p. 28):

A cidade histórica possui uma admirável riqueza arquitetônica do período colonial, restaurado e conservado com o tempo, seus museus, igrejas, coretos, chafarizes nos levam a uma viagem ao passado. Há destaque a casa da grande poetisa Cora Coralina e também a pintora Goiandira do Couto [...]. Como produtos turísticos destacam-se também: O Palácio Conde dos Arcos, sede do governo; o Museu das Bandeiras, antiga Cadeia Municipal; o Chafariz de Cauda; Museu de Arte Sacra da Boa Morte; Casa de Fundação do Ouro, Ministério Público; Catedral de Santana; Coreto do Jardim; Cruz do Anhanguera; Igreja Nossa Senhora do Carmo; Prédio da Real Fazenda e outros de uma imensa lista.

A ocupação da região onde hoje se encontra o estado de Goiás remete ao início da colonização portuguesa em território brasileiro. As primeiras expedições que chegaram à região vinham, principalmente, das regiões onde hoje se encontram a Bahia e São Paulo.

Segundo Borges (2010, p. 29) foi “Bartolomeu Bueno da Silva” o primeiro a chegar à região do Rio Vermelho com a intenção de permanecer. A expedição do “Anhanguera” ocorreu em 1722:

[...] a motivação predominantemente para a exploração das terras goianas foi o índio, devido a sua extinção nos grandes centros colonizadores da costa – Pernambuco, Bahia, Rio – e a dificuldade de importação, em certos períodos, de negros da África, empurravam os paulistas a dedicarem-se, ao negócio lucrativo que era o ciclo comercial do índio. (BORGES, 2010, p. 30).

Apesar disso, sabe-se que este explorador encontrou ouro na região, o que deu origem à corrida do ouro no século XVIII.

(...) de uma atividade febril. (...) Bueno funda solenemente o primeiro arraial, o arraial de Sant’Anna. Localizado entre morros, numa quebrada no sopé da Serra Dourada, muito próximo das nascentes do Rio Vermelho, a nova povoação – que deveria converter-se doze anos depois em vila e tornar-se capital – geograficamente se encontra deslocada, como centro de operações, no território goiano, e climaticamente exposta aos rigores de uma insolação concentrada, sem ventilação. Mas a urgência do momento não admite dilações. Há ouro e água, isto basta. Este será o critério com que irão surgindo os demais arraiais. (PALACIN, 1994, p. 19-25 *apud* BORGES, 2010, p. 31).

A fundação do arraial de Sant’Anna, no ano de 1722, ocorre no que Lucchesi (2017, p. 375) define como 3ª fase da “Periodização da História Sociolinguística do Brasil”. Segundo o autor, este período, denominado por ele como “Homogeneização Linguística”, subdivide-se em três fases: “1695-1808 - Primeira vaga de aportuguesamento; 1808-1850 - [...] Segunda vaga de aportuguesamento; [...] 1850-1930 - Terceira vaga de aportuguesamento”.

Assim, a *primeira vaga de aportuguesamento* é que está em voga no período de ocupação da região do Rio Vermelho, caracterizada por: “O ciclo do ouro promove o afluxo maciço de colonos portugueses e escravos africanos, com o primeiro surto de urbanização do Brasil: avanço da língua portuguesa no sudeste, e em outras regiões, com a pecuária”. (LUCCHESI, 2017, p. 376).

A este respeito, no qual foi dado o início da constituição da comunidade de fala vilaboense, Lucchesi (2017, p. 365-366) observa que:

[...] Essa fase se define por um amplo, profundo e violento processo de homogeneização linguística, no qual o português se impõe como língua hegemônica, tornando-se a língua materna da imensa maioria da população brasileira. Como resultado de todo esse processo, nenhuma das línguas africanas conservou-se no Brasil, e o uso das línguas indígenas foi drasticamente reduzido, ficando confinado às regiões mais recônditas do país, nomeadamente na Amazônia e no Centro-Oeste. O desaparecimento das línguas africanas deveu-se à desmedida violência física, cultural, simbólica e psicológica inerente ao processo de escravidão, ao passo em que a extinção de muitas línguas indígenas deveu-se ao extermínio dos povos que as falavam, não obstante muitas outras tenham desaparecido pela aculturação forçada de seus falantes, inclusive pela catequização dos jesuítas, que lhes impunham a língua geral de base tupi ou tupinambá, de modo que a homogeneização linguística de muitas populações indígenas passou por uma fase intermediária de assimilação da língua geral, até a imposição do português.

Percebemos, com isso, que a principal via de imposição do português como língua hegemônica no Brasil foi a violência – praticada contra indígenas e contra os africanos que eram escravizados no Brasil naquele período.

Depois disso, a capital passou a receber a denominação de *Vila Boa de Goyas* em 1739. A descoberta de ouro na “Serra dos Pirineus” (BORGES, 2010, p. 31), região onde hoje se localiza a cidade de Pirenópolis-GO, fez surgir a ideia de mudança da capital, visto que essas jazidas se encontravam em uma região mais central e de clima mais ameno. Todavia, a mudança da capital postergada, visto que a coroa portuguesa a considerou inviável para o momento.

A ocupação da região se deu principalmente pela exploração do ouro. Segundo Prudente (2006, p. 18) afirma que:

No que tange à formação populacional, em 1750, a Capitania de Goiás possuía pouco menos de 40.000 habitantes, tendo sido dobrada a população com relação a sua fundação cerca de 25 anos antes. A população prosseguiu aumentando, havendo em 1783 quase 80.000 habitantes.

Esse aumento da população urbana se encaixa no que diz Bortoni-Ricardo (2021, p. 63): “A grande distância entre a chamada língua padrão, ensinada nas escolas, e as variedades regionais-rurais pode possivelmente ser explicada pela influência de uma língua *pidginizada* nos últimos séculos de colonização[...]”. Ocorre que, apesar de no período em questão já ter se iniciado o aportuguesamento (LUCCHESI, 2017), havia ainda uma forte influência do período anterior, denominado por Lucchesi (2017, p. 375) de Multilinguismo Generalizado. Assim, apesar da urbanização propiciada pela corrida do ouro, havia uma tendência do uso dos *pidgins* no interior do Brasil.

No entanto, segundo Prudente (2006, p. 18), o ano de 1778 marca o início do declínio da extração de ouro na região. Em decorrência disso há a mudança de modelo econômico, chegando ao ponto de haver apenas agricultura de subsistência, devido ao isolamento da região.

Fato que chama a atenção é o apresentado no dossiê intitulado “Proposição de inscrição da cidade de Goiás na lista do Patrimônio da Humanidade” (GOIÁS, 1999¹⁵ *apud* BORGES, 2010, p. 33):

Quando a Abolição foi decretada, em 1888, quase não havia mais escravos para libertar na capital, pois as três raças – índia, portuguesa e africana – misturam-se desde o início da ocupação da região, fazendo com que o movimento pela abolição da escravatura fosse muito intenso em Goiás (DOSSIÊ, 1999).

O isolamento das cidades e o lento desenvolvimento levou o interventor Pedro Ludovico Teixeira a solicitar a mudança da capital do estado de Goiás. Essa mudança de capital, da Cidade de Goiás para Goiânia, consolidou-se no ano de 1933.

Segundo Prudente (2006, p. 57-58), alguns prédios setecentistas do município foram tombados como Patrimônios Nacionais já na década de 1950 pela Diretoria do Patrimônio Histórico e Artístico Nacional (DPHAN):

- Igreja de São Francisco de Paula;
- Igreja de Nossa Senhora do Carmo;
- Igreja de Nossa Senhora D’Abadia;
- Igreja de Santa Bárbara;
- Igreja da Boa Morte;
- Imagem de Nossa Senhora do Rosário;
- Quartel do XXº Batalhão da Infantaria;
- Museu das Bandeiras;
- Conjunto Arquitetônico e Urbanístico da Praça Brasil Caiado;
- Palácio Conde dos Arcos, armas de Portugal e dois bustos em pedra;
- Conjunto Arquitetônico e Urbanístico da Rua da Fundação; e
- Igreja de São João Batista do antigo Arraial do Ferreiro.

Em 1978 houve a expansão do tombamento, o que encaminhou a delimitação do centro histórico da cidade. Prudente (2006, p. 58) destaca os principais pontos, elencados no Dossiê supramencionado, que justificam o tombamento:

[...] de ser o primeiro centro urbano oficialmente planejado a oeste da linha de Tordesilhas; de possuir uma estrutura urbana informal vernacular, diferente dos centros urbanos de Minas Gerais; de refletir nas estruturas urbanas das cidades goianas congêneres; e por fim, de seu núcleo de arquitetura vernacular expressar a cultura da coletividade, suas relações com o território e a expressão da diversidade cultural planetária

¹⁵ Por se tratar de um arquivo em formato de CR-ROM, não foi possível ter acesso ao original.

Em outras palavras, Goiás preserva características das cidades brasileiras dos séculos XVIII e XIX. Como podemos observar nas imagens abaixo:

Figura 2: Praça do Coreto (19--)



Fonte: Acervo IBGE. Disponível em: <https://servicodados.ibge.gov.br/api/v1/resize/image?maxwidth=600&maxheight=600&caminho=biblioteca.ibge.gov.br/visualizacao/fotografias/GEBIS%20-%20RJ/go39718.jpg>. Acesso em 12 de jun. de 2021.

Figura 3: Centro da Cidade de Goiás – GO (1957)



Fonte: Acervo IBGE. Disponível em: <https://servicodados.ibge.gov.br/api/v1/resize/image?maxwidth=600&maxheight=600&caminho=biblioteca.ibge.gov.br/visualizacao/fotografias/GEBIS%20-%20RJ/GO24699.jpg>. Acesso em 12 de jun. de 2021.

Figura 4: Casarões da Cidade de Goiás (1957).



Fonte: Acervo IBGE. Disponível em: <https://servicodados.ibge.gov.br/api/v1/resize/image?maxwidth=600&maxheight=600&caminho=biblioteca.ibge.gov.br/visualizacao/fotografias/GEBIS%20-%20RJ/GO24701.jpg>. Acesso em 12 de jun. de 2021.

Figura 5: Igreja Nossa Senhora do Rosário e Centro de Cultura e Memória do Poder Judiciário do Estado de Goiás, ao fundo (2021).



Fonte: Instagram da Prefeitura de Goiás. Disponível em: <https://www.instagram.com/p/CNvqhp8hHR8/>. Acesso em 30 de mai. de 2021.

Figura 6: Cidade de Goiás: vista panorâmica (2021).



Fonte: Instagram da Prefeitura de Goiás. Disponível em: <https://www.instagram.com/p/CNvqhp8hHR8/>. Acesso em 30 de mai. de 2021.

Por conta do nível de preservação do patrimônio na Cidade de Goiás – GO, no dia 13 de dezembro de 2001, a cidade foi reconhecida como Patrimônio da Humanidade pela UNESCO.

3.2 Constituição do corpus

O *corpus* do trabalho será constituído por amostras de língua oral do município Goiás-GO. Para isso, lança-se mão do banco de dados de dois grupos de estudo: o Grupo de Estudos Funcionalistas, da Universidade Federal de Goiás (GEF/UFG); e do Grupo de Estudos e Pesquisas em Sociolinguística, da Universidade Estadual de Goiás (Sociolinco/UEG/Cora Coralina).

3.2.1 Amostras de fala

Para nosso trabalho nos valeremos de amostras de língua oral do município Goiás, localizado no Estado Goiás, Brasil. Para isso, lançamos mão dos bancos de dados de dois grupos de estudo: o Grupo de Estudos Funcionalistas, da Universidade Federal de Goiás (GEF/UFG); e do Grupo de Estudos e Pesquisas em

Sociolinguística, da Universidade Estadual de Goiás (Sociolinco/UEG/Cora Coralina). As entrevistas que constituem o banco de dados do GEF foram gravadas e transcritas no ano de 2003, tendo como informantes pessoas que não concluíram o ensino fundamental. Para complementar a análise, usamos as entrevistas oriundas do banco de dados do Sociolinco, gravadas no ano de 2019, com o mesmo perfil de informantes. O fato de as entrevistas terem sido gravadas com os mesmos critérios permitiu um estudo em tempo real, abordando a língua falada na cidade em épocas diferentes.

Mais especificamente, as entrevistas realizadas em 2019 seguiram os moldes utilizados em 2003, abordando informantes com o mesmo perfil e com o mesmo roteiro das anteriores. As entrevistas do banco de dados do dividem-se em duas faixas etárias. Cada uma delas é representada por 3 informantes de cada sexo/gênero. Todos eles têm padrão socioeconômico baixo e falante de variedades populares do PB. Em consonância a isso, as entrevistas do banco de dados do Sociolinco (RIBEIRO, 2020) também são divididas em duas faixas etárias e cada uma delas é também representada por 3 informantes de cada sexo/gênero. Assim, no geral, a amostra é constituída por 2 faixas etárias, com 12 informantes cada uma, compostas por 6 homens e mulheres, totalizando 24 entrevistas.

É importante destacar que, segundo Freitag (2017), a entrevista sociolinguística é “um protocolo que visa fazer emergir o vernáculo de um falante representativo de uma dada comunidade de fala” (LABOV, 1972 *apud* FREITAG, 2017, p. 18) que deve garantir a:

[...] confiabilidade (os mesmos resultados devem ser repetidos na análise de um mesmo fenômeno) e intersubjetividade (dois pesquisadores diferentes devem obter os mesmos resultados seguindo a mesma metodologia) da análise. (FREITAG, 2017, p. 18).

Para as entrevistas dos bancos de dados (GEF/UFG; e Sociolinco/UEG/Cora Coralina) o método seguido para a escolha de informantes foi a “amostragem de conglomerados não aleatórios de adesão voluntária por indicação ou julgamento” (FREITAG, 2017, p. 19), ou seja, a seleção dos informantes não ocorreu modo aleatório, mas, na verdade, buscou-se um primeiro informante e, feita a entrevista com ele, foi solicitada a indicação de outra pessoa que pudesse colaborar, formando, assim, redes de relações sociais (FREITAG, 2017).

Seguindo este princípio, o pesquisador deve se inserir na comunidade e observar, de forma discreta, a fim de escolher o informante inicial. Isso deve ser feito

seguindo alguns critérios: ele deve desempenhar um papel de liderança na comunidade e não apresentar restrições fonológicas que prejudiquem a transcrição. Feita a primeira entrevista, o informante indica outros que possam ser informantes também. Para a realização da entrevista, o documentador informa que o procedimento dura cerca de 1 hora e é registrado em áudio. O informante deve preencher a ficha social e, “para quebrar o gelo”, pode ser feita uma checagem desta ficha no início da entrevista. Ainda a respeito da entrevista, é preciso levar em conta que a entrevista sociolinguística não é uma “conversa entre amigos” (FREITAG, 2017, p. 23). Assim, apelar para assuntos que sejam mais emocionais pode ajudar a emergir o vernáculo.

As entrevistas realizadas seguiram um roteiro que se repetiu com todos os informantes, a fim de guiar as entrevistas e amenizar o efeito do paradoxo do observador. Labov (2008 [1972], p. 244) define o paradoxo do observador da seguinte forma:

[...] o objetivo da pesquisa linguística na comunidade deve ser descobrir como as pessoas falam quando não estão sendo sistematicamente observadas – no entanto, só podemos obter tais dados por meio da observação sistemática.

Desta forma, o roteiro semiestruturado serve para conduzir a entrevista de modo a fazer emergir o vernáculo, como dito acima.

A transcrição das entrevistas seguiu o método semiortográfico, ou seja, explicitam-se graficamente algumas características fonológicas da fala, diferenciando-se da escrita ortográfica em que se faria a adaptação de para atender às convenções da escrita. Diferencia-se, também, da transcrição fonética, que geralmente é feita com a utilização de *softwares* (FREITAG, 2017, p. 14). Assim, em Silva (2005, p. 25) e em Ribeiro (2020, p. 68) é explicitado que a transcrição, em ambos os trabalhos, seguiu as normas propostas para o projeto Norma Linguística Urbana Culta (NURC) (CASTILHO; PRETI, 1987, p. 15), apesar de as chaves de transcrição não serem as mesmas há uma aproximação.

As entrevistas foram colhidas com um intervalo de 16 anos, o que nos permite a análise de tendência em tempo real. A pesquisa em tempo real se caracteriza por uma amostra randômica de uma determinada comunidade de fala que, depois de certo período é repetida (COELHO *et al.*, 2015, p. 118). Assim, coletaram-se entrevistas no ano de 2003 e em 2019 as outras entrevistas seguindo o mesmo roteiro e com informantes de mesmo perfil. A pesquisa de tendência se difere da de painel, pois esta exige que a entrevista seja realizada com o mesmo informante abordado na primeira

e aquela aborda outros informantes que possuam o mesmo perfil (FREITAG, 2005). Embora não permita abordar os mesmos informantes, essa metodologia viabiliza a análise da língua falada pela comunidade em duas épocas distintas, possibilitando uma abordagem panorâmica do fenômeno.

Para Freitag (2017), “A entrevista sociolinguística é resultado de um processo que se inicia com a seleção do informante, a sua adesão ao experimento e a realização da documentação em áudio propriamente dita”. (FREITAG, 2017, p. 22). Este, portanto, é um recurso bastante produtivo, pois possibilita a contemplação de diferentes estilos de fala, visto que abrange “tipos textuais/sequências discursivas e conteúdo específico/tópicos discursivos”.

Junto a isso, levantamos dados históricos e oficiais que refletem as características socioculturais presente no município, a fim de alcançar uma análise fidedigna. Em seguida, fizemos a análise dos dados e traçamos um panorama do que se obteve como resultado.

Para a verificação da influência de cada variante independente no fenômeno em estudo, usamos o programa R (R CORE TEAM 2013), visto que ele permite compilar os dados, realiza análises estatísticas, listas de frequência, dentre outros. Este *software* é distribuído gratuitamente e dispensa a elaboração de códigos, necessários para outros programas, o que facilita o trabalho propriamente dito da análise sociolinguística: descreve um fenômeno de variação. Assim, será possível entender quais as variáveis independentes realmente exercem influenciam na variação que é objeto desta pesquisa.

3.2.2 Banco de dados GEF

O “Grupo de Estudos Funcionalistas (GEF): análise, descrição e ensino”¹⁶, da Faculdade de Letras da Universidade Federal de Goiás, foi fundado no ano de 2003 e tem por objetivo o estudo dos fenômenos que constituem o PB, principalmente fenômenos observáveis na fala goiana.

Dentre os empreendimentos do Grupo de Estudos em questão, há o “Projeto Português Contemporâneo Falado em Goiás – ‘Fala goiana’” (PFG). Este projeto é dedicado à análise da variedade goiana do PB e tem como objetivo geral a

¹⁶ Disponível em: <https://gef.letras.ufg.br/>. Acesso em 30 de abr. de 2021.

“documentação e análise do português falado em Goiás, considerando-se os aspectos da constituição social e cultural das comunidades de fala envolvidas, visível especialmente em fenômenos gerais de mudança linguística”¹⁷.

A amostra que os pesquisadores deste grupo de estudo coletaram para o PFG foi disponibilizado para acesso público, tendo apenas a identidade e a voz dos informantes preservadas. Os informantes são de algumas cidades do interior goiano e de Goiânia. A maioria das entrevistas é de informantes da Cidade de Goiás, sendo que a faixa etária destes se estende entre 25 e 72 anos e há entrevistas com informantes de ambos os sexos – para estas entrevistas não foi considerada a variável gênero, o que inviabilizou o controle da mesma nas entrevistas realizadas posteriormente (apresentadas em 3.2.2). A coleta dessas entrevistas foi realizada nos anos de 2003 e 2004, sendo que as últimas foram gravadas há 18 anos.

Quadro 3 – Informantes da amostra GEF

PROJETO FALA GOIANA			
	IDADE	SEXO	IDENTIFICAÇÃO DO INFORMANTE
FAIXA ETÁRIA 1	28	F	[GEFf1F-001]
	33	F	[GEFf1F-002]
	43	F	[GEFf1F-003]
	25	M	[GEFf1M-007]
	30	M	[GEFf1M-12]
	36	M	[GEFf1M-008]
FAIXA ETÁRIA 2	48	F	[GEFf2F-004]
	65	F	[GEFf2F-005]
	70	F	[GEFf2F-006]
	65	M	[GEFf1M-009]
	72	M	[GEFf2M-010]
	75	M	[GEFf2M-011]

Fonte: adaptado de Ribeiro (2020, p. 65).

Dessa forma, faremos a análise dos dados tendo como ponto de partida os dados coletados em 2003, que reflete a língua falada em um período diferente do atual. Com isso, será possível contrastar os dois períodos para obter maior profundidade na análise.

3.2.3 Banco de dados Sociolingo

¹⁷ Disponível em: <https://gef.letras.ufg.br/p/11947-projetos-tematicos>. Acesso em 30 de abr. de 2021.

Para complementar a análise, serão utilizadas as entrevistas oriundas do banco de dados do Sociolingo¹⁸, gravadas no ano de 2019, com o mesmo perfil de informantes. Essas entrevistas garantirão um estudo em tempo real, abordando a língua falada na cidade em épocas diferentes.

Segundo Ribeiro (2020, p. 66), os roteiros utilizados para a realização das entrevistas do PFG não foram disponibilizados, assim sendo, para as entrevistas realizadas pela autora, o roteiro foi adaptado a partir da observação das entrevistas do primeiro banco de dados.

As entrevistas do banco de dados do GEF dividem-se em duas faixas etárias. Os informantes têm padrão socioeconômico baixo, ou seja, em suas atividades profissionais, não lhes é exigido o domínio da norma culta. Em consonância a isso, as entrevistas do banco de dados do Sociolingo (RIBEIRO, 2020) também são divididas em duas faixas etárias. As entrevistas do Sociolingo se organizam conforme tabela a seguir:

Quadro 4 – Informantes da amostra Sociolingo

SOCIOLINGO			
	IDADE	SEXO	IDENTIFICAÇÃO DO INFORMANTE
FAIXA ETÁRIA 1	31	F	[AMGf1F-001]
	38	F	[AMGf1F-002]
	40	F	[AMGf1F-003]
	33	M	[AMGf1M-007]
	38	M	[AMGf1M-008]
	40	M	[AMGf1M-009]
FAIXA ETÁRIA 2	46	F	[AMGf2F-004]
	46	F	[AMGf2F-005]
	60	F	[AMGf2F-006]
	47	M	[AMGf2M-010]
	48	M	[AMGf2M-011]
	49	M	[AMGf2M-012]

Fonte: adaptado de Ribeiro (2020, p. 65).

É importante frisar que, segundo Freitag (2017), a entrevista sociolinguística é “um protocolo que visa fazer emergir o vernáculo de um falante representativo de uma dada comunidade de fala” (LABOV, 1972 *apud* FREITAG, 2017, p. 18) que deve garantir a “confiabilidade (os mesmos resultados devem ser repetidos na análise de

¹⁸ Grupo de Pesquisas e Estudos em Sociolinguística da Universidade Estadual de Goiás, Câmpus Cora Coralina, coordenado pela Professora Marília Silva Vieira.

um mesmo fenômeno) e intersubjetividade (dois pesquisadores diferentes devem obter os mesmos resultados seguindo a mesma metodologia) da análise” (FREITAG, 2017, p. 18).

As entrevistas foram colhidas com um intervalo de 16 anos, o que permite a análise de tendência em tempo real. A pesquisa em tempo real se caracteriza por uma amostra randômica de uma determinada comunidade de fala que, depois de certo período é repetida. Assim, coletaram-se entrevistas no ano de 2003 e, em 2018, gravaram-se outras, entrevistas seguindo o mesmo roteiro e com informantes de mesmo perfil. A pesquisa de tendência se difere da de painel, pois esta exige que a entrevista seja realizada com o mesmo informante abordado na primeira e aquela aborda outros informantes que possuam o mesmo perfil (FREITAG, 2005). Embora não permita abordar os mesmos informantes, essa metodologia viabiliza a análise da língua falada pela comunidade em duas épocas distintas, possibilitando uma abordagem panorâmica do fenômeno.

Para Freitag (2017) “A entrevista sociolinguística é resultado de um processo que se inicia com a seleção do informante, a sua adesão ao experimento e a realização da documentação em áudio propriamente dita”. (FREITAG, 2017, p. 22). Este, portanto, é um recurso bastante produtivo, pois possibilita a contemplação de diferentes estilos de fala, visto que abrange “tipos textuais/sequências discursivas e conteúdo específico/tópicos discursivos”.

Freitag (2017) define a entrevista sociolinguística como um método por meio do qual é possível registrar o vernáculo do informante, sendo necessário garantir a “confiabilidade (os mesmos resultados devem ser repetidos na análise de um mesmo fenômeno) e intersubjetividade (dois pesquisadores diferentes devem obter os mesmos resultados seguindo a mesma metodologia) da análise” (FREITAG, 2017, p. 18).

Junto a isso, é necessário levantar dados históricos e oficiais que reflitam as características socioculturais do município, a fim de alcançar uma análise fidedigna. Em seguida, será feita a análise dos dados e traçado um panorama do que se obteve como resultado. Em seguida, será realizada uma revisão de literatura, oriunda do levantamento bibliográfico anterior, para comparar os dados obtidos na pesquisa com outros obtidos em localidades distintas. Feito isso, será elaborado o texto para a dissertação.

3.3A análise estatística na sociolinguística: R

A partir das ocorrências encontradas no *corpus*, fizemos a análise das variáveis independentes a fim de verificar o peso destas para a ocorrência da variável dependente. Antes de mais nada, para que se entenda o processo de análise, é preciso enumerar os passos tomados para que, enfim, se chegasse à análise no R.

Em um primeiro momento, analisamos as entrevistas, que estão disponíveis em formato *PDF*. Nessas entrevistas, buscamos as ocorrências do fenômeno e as isolamos em uma planilha de dados. Depois de isoladas, fazemos a classificação das ocorrências de acordo com cada variável independente selecionada.

O R (*R CORE TEAM 2013*) é um software que auxilia a análise de dados. Ele permite selecionar, classificar, quantificar e visualizar a relação entre as variáveis dependentes e as independentes. Como se pode observar na imagem abaixo, para o tratamento dos dados no R, depois de terminada a classificação dos dados, é necessário que o arquivo seja convertido para o formato “comma-separated-values” (valores separados por vírgulas), “.csv”. Desta forma, todas as ocorrências são distribuídas nas linhas e em apenas uma coluna, sendo as variáveis colocadas logo adiante, separadas por vírgulas:

Figura 7: Planilha de ocorrência .CSV

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	INFORMANTE	TIPO	STATUS	TIPOSEQ	TIPOORA	MARCADORES	SUJ	SEXO	FAIXAETARIA	AMOSTRA										
2	A	NEG1	INF	DIA	SUB	N	EXP	FEM	1a	GEF										
3	A	NEG1	ATV	DIA	SUB	N	EXP	FEM	1a	GEF										
4	A	NEG1	NOVA	NFP	ABS	S	EXP	FEM	1a	GEF										
5	A	NEG1	ATV	NFP	SUB	N	EXP	FEM	1a	GEF										
6	A	NEG1	NOVA	NFP	COO	N	IMPL	FEM	1a	GEF										
7	A	NEG1	INF	NFP	COO	S	EXP	FEM	1a	GEF										
8	A	NEG1	INF	NFP	COO	S	EXP	FEM	1a	GEF										
9	A	NEG1	ATV	NFP	COO	N	INEX	FEM	1a	GEF										
10	A	NEG1	ATV	NFP	SUB	N	EXP	FEM	1a	GEF										
11	A	NEG1	INF	NFP	SUB	N	IMPL	FEM	1a	GEF										
12	A	NEG1	INF	NFP	ABS	N	IMPL	FEM	1a	GEF										
13	A	NEG1	INF	NFP	ABS	N	IMPL	FEM	1a	GEF										
14	A	NEG1	INF	NFP	ABS	N	IMPL	FEM	1a	GEF										
15	A	NEG1	ATV	NFP	ABS	N	IMPL	FEM	1a	GEF										
16	A	NEG1	INF	NFP	ABS	S	IMPL	FEM	1a	GEF										
17	A	NEG1	NOVA	NFP	COO	N	EXP	FEM	1a	GEF										
18	A	NEG1	INF	NFP	ABS	N	EXP	FEM	1a	GEF										
19	A	NEG1	ATV	NFP	ABS	S	INEX	FEM	1a	GEF										
20	A	NEG1	ATV	NFP	ABS	N	INEX	FEM	1a	GEF										
21	A	NEG1	INF	NFP	PR	N	IMPL	FEM	1a	GEF										
22	A	NEG1	ATV	NFP	PR	N	IMPL	FEM	1a	GEF										
23	A	NEG1	INF	NFP	COO	N	INEX	FEM	1a	GEF										
24	A	NEG1	ATV	NFP	COO	S	INEX	FEM	1a	GEF										
25	A	NEG1	INF	NFP	ABS	N	EXP	FEM	1a	GEF										
26	A	NEG1	INF	NFP	COO	S	EXP	FEM	1a	GEF										
27	A	NEG1	INF	AVA	SUB	N	INEX	FEM	1a	GEF										
28	A	NEG1	INF	DIA	ABS	N	INEX	FEM	1a	GEF										
29	A	NEG1	INF	NFP	ABS	N	IMPL	FEM	1a	GEF										
30	A	NEG1	NOVA	NFP	SUB	N	INEX	FEM	1a	GEF										
31	A	NEG1	INF	NFP	COO	S	IMPL	FEM	1a	GEF										
32	A	NEG1	ATV	AVA	PR	N	IMPL	FEM	1a	GEF										
33	A	NEG1	INF	NFP	COO	S	IMPL	FEM	1a	GEF										

Fonte: Elaboração nossa.

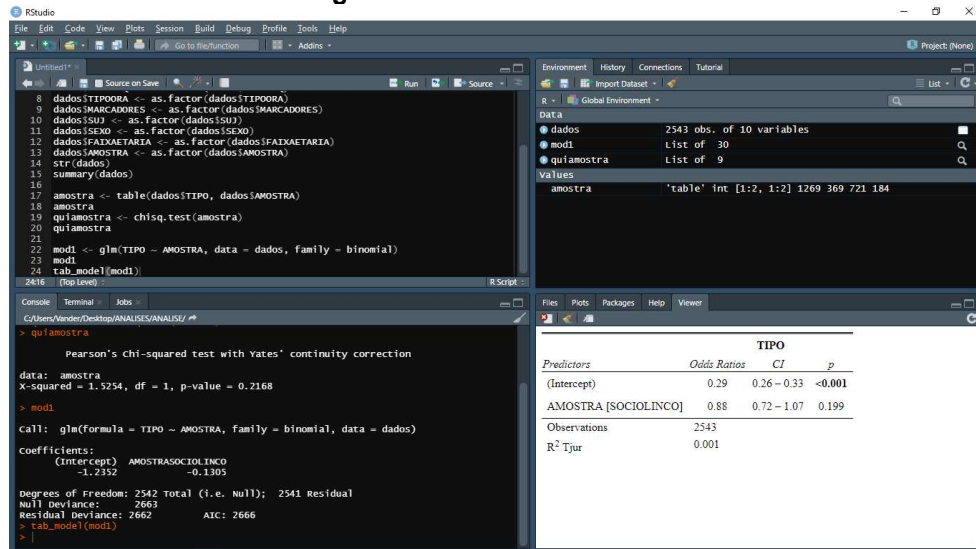
Isso é necessário visto que uma análise univariada não dá conta de diversos fatores contextuais envolvidos na variação linguística. Segundo Guy e Zilles (2007, p. 34):

Uma vez que a distribuição dos dados linguísticos geralmente é, de fato, desigual, uma análise multivariada dará resultados mais precisos, porque ao mesmo tempo em que computa o efeito de uma variável independente, ela controla explicitamente o efeito de todas as outras variáveis independentes conhecidas.

Por ser um *software* que se baseia em scripts, a interface não é muito intuitiva, para isto, foi desenvolvido o “RStudio”, que permite uma interação simplificada. A imagem abaixo é um exemplo da funcionalidade do “RStudio”. Do lado direito superior temos o *console - input*, onde se introduzem as linhas de comando. No lado direito, na parte inferior, há o *console - output*, onde aparecem as linhas de comando digitadas no *Script* e os resultados obtidos a partir dos comandos executados. Do lado esquerdo, na parte superior, temos o ambiente de trabalho (“*environment*”), onde aparecem os dados – *dataframes*, tabelas, modelos, etc. – que estão sendo manipulados. Do lado direito, na parte inferior, aparecem 5 abas que apresentam com as opções relativas ao ambiente de trabalho - *File*, os gráficos gerados – *Plots*, os pacotes instalados e carregados – *Packages*, ajuda a respeito de comandos, pacotes, etc. – *Help*, e a visualização de tabelas – *Viewer*.

Na figura, podemos observar as linhas de comando para a realização do teste “qui-quadrado”: *quiamostra <- chisq.test(amostra)*; e o comando para de um modelo que testa os efeitos da variável previsora “Amostra” (GEF ou SOCIOLINCO) sobre a variável resposta “Tipo” (NEG1 ou NEG2): *mod1 <- glm(TIPO ~ AMOSTRA, data = dados, family = binomial)*. Logo após esses comandos, no *output*, podemos observar o resultado desses comandos. Em *Viewer* temos o resultado do modelo plotado em uma tabela a partir do comando: *tab_model(mod1)*. É importante mencionar que o termo que antecede o “<-” serve para a criação do modelo ou do dataframe resultante, que estarão disponíveis no ambiente de trabalho (“*environment*”).

Figura 8: Interface do “RStudio”



Fonte: Elaboração nossa (print da tela).

Esses são alguns comandos disponíveis no R para a análise. O R é um software livre e de fácil manuseio, utilizado para pesquisas em diversas áreas do conhecimento. Por vezes, os usuários contribuem para o desenvolvimento do *software* criando pacotes para agrupar algumas funções e facilitar seu uso.

No próximo capítulo, serão descritas as variáveis linguísticas e sociais que compõem o envelope de variação da negação na fala vilaboense e que recebeu tratamento estatístico do R.

4 VARIÁVEIS SOCIAIS E LINGUÍSTICAS

De um modo geral, podemos afirmar que “a variação e a mudança linguísticas são inerentes ao próprio sistema, podendo ser controladas por restrições de caráter interno (estrutural) ou externo (social, contextual, discursivo etc.)” (SCHERRE; YACOVENCO, 2011, p. 122).

Por conta disso, dividimos as variáveis aqui controladas em dois grandes grupos: Variáveis Linguísticas, abordadas na subseção 4.1; e Variáveis Sociais, abordadas na subseção 4.2.

4.1 Variáveis linguísticas

Esta seção é dedicada à apresentação das variáveis linguísticas: status da informação; tipo de sequência discursiva; tipo de oração; presença ou ausência de marcadores discursivos e tipo de sujeito.

4.1.1 Status da informação

Essa variável se refere ao *status* discursivo da informação negada. Assim sendo, observamos que a informação pode ser nova ou velha no discurso, em outras palavras, ativada ou não ativada.

Em 2.2, discutimos os desdobramentos da pesquisa de Schwenter (2005) que concluiu que, no PB, NEG1 aparecerá livremente – ou seja, pode ocorrer como negação para informações novas ou velhas no discurso. Em contrapartida, NEG2 e NEG3 não aparecem ocorrem em contextos em que a informação tenha sido ativada, direta ou inferencialmente.

Ainda nesse sentido, Schwenter (2005) propõe os chamados “pontos fora da curva” que indicam ocorrências de NEG2 negando informações não ativadas, mas a informação nova negada por NEG2 “[...] é apresentada como se fosse um discurso antigo pelos próprios falantes¹⁹” (SCHWENTER, 2005, p. 1452, tradução nossa). NEG3, por sua vez, caracteriza-se por negar informações que foram ativadas diretamente no discurso, não ocorrendo em contextos diferentes desses.

Nas *amostras* aqui analisadas, encontramos a variação entre NEG1, NEG2 e NEG3 principalmente quando a informação é discursivamente recuperável, seja

¹⁹ “[...] one that is presented as if it were discourse-old by speakers themselves”.

diretamente ou por meio de inferência. Há raros casos de NEG2 em contexto de informações novas e nenhum para NEG3 – isso será apresentado na seção 5.

NOVA:

(19) DOC. Ah::: cê NUM gostô

INF. NÃO gostei:::

DOC. Porque:::?

INF. Porque as pessoa falava pra mim assim... ah::: cê **NUM** tem sorte... porque cê nasceu no di... no mês de agosto... aí quando eu falava qu/eu nasci quinze... es falava meu::: Deus... cê tem que orá muito... cê vai sê sem sorte... () (GEF48FF-Luana)

ATIVADA:

(20) DOC. Água... energia...

INF. É NUM é fáci NÃO mais eu vô:::... vô viveno né? é o jeito ((risos))

DOC. É... mais Deus num abandona né?

INF. NÃO... isso eu tenho certeza **NUM** abandona **NÃO**...

DOC. Isso... muito bom... e seus filhos qual... qual é o mais velho...o... o rapais... o a moça? (GEF48FF-Luana)

INFERÍVEL:

(21) DOC. Então ele plantava pro outro também?

INF. É... prantava na meia com fazendero... no caso assim... seu () dava a terra e os... os grão pra prantá... aí o que dá divide... aí dava treis saca... uma... aí ele nem pegava aquilo... dexava... nós era muito fii né... aí cabava rápido né... aí NUM tinha jeito

DOC. O outro é que tinha de dexá né? ((risos))

INF. É...

DOC. Porque provavelmente ele tinha né? mais condições

INF. Tinha... aí ele **NUM** pegava a renda **NÃO**... ficava por lá mesmo... nesse ponto também ele foi um bom patrão...

DOC. Ah tá... (GEF48FF-Luana)

Observamos em (21) a negação de uma informação nova, visto que “ter sorte” não está inserido no contexto do diálogo. Em (22), a conversa trata de algumas dificuldades enfrentadas pela informante e há a negação de uma informação diretamente ativa no discurso: “**NUM** abandona **NÃO**...”. Já em (23), observamos a ocorrência em que a informação negada é ativada, porém inferível, visto que não havia “renda”, pois o trabalho era remunerado com a divisão dos “grãos”.

Nascimento (2014, p. 82) indicou uma maior suscetibilidade de NEG3 para esta variável, ao passo que NEG2 e NEG1 mantiveram pesos relativos próximos. Rocha (2013, p. 64) constatou que, em São Paulo, NEG2 é favorecida por informações diretamente ativadas no discurso. Diante do que foi exposto, acreditamos que NEG2 será mais frequente em contextos em que a informação negada tenha sido diretamente ativada no discurso.

4.1.2 Tipo de sequência discursiva

Outra variável linguística que controlamos foi o *tipo de sequência discursiva* em que a informação negada se insere. Assim, valemo-nos da concepção de “discurso como enunciado” (PAREDES SILVA, 1997, p. 81). Assim sendo, entendemos que os gêneros textuais terão, então, funções de ordem sequencial – que diz respeito à organização da sequencial dos termos em um dado enunciado – e de ordem semântica e pragmática – aí se trata da comunicação propriamente dita, na medida em que as informações se interpretam e transmitem.

Em outras palavras, segundo Di Palma Back *et al.* (2004, s/p):

Sequências discursivas (tipos de texto) são as estruturas convencionalizadas de que o falante dispõe na língua para organizar o seu discurso, marcadas por características como tempo, modo e aspecto verbal, pessoa do discurso em referência, unidades sintáticas e semântica predominantes. O uso destas estruturas em situações reais de comunicação caracteriza os *gêneros do discurso*.

Neste sentido, observamos nas ocorrências, extraídas das amostras, as sequências discursivas a seguir: argumentação, avaliação, descrição, diálogo, narração e procedural. Logo a hipótese é de que as sequências dialogais favorecerão as variantes inovadoras.

4.1.2.1 Sequência argumentativa

A sequência discursiva argumentativa/expositiva, segundo Paredes Silva (1997, p. 90), caracteriza-se por sentenças complexas, geralmente contendo subordinação, os verbos apresentam aspectualidade imperfectiva, o que resulta em construções hipotéticas. Apresentamos esse tipo de sequência em (24):

(22) INF. hoje eu falo... ô meu Deus... **se eu tivesse seguido as instruções da minha mãe tivesse feito tudo que ela falava e que a gente NUM passava metade do que a gente passa hoje... porque se eu tivesse seguido tudo que ela falô à risca... minha vida era totalmente diferente... mas eu NUM reclamo da vida que tem também NÃO...** (SOCIOLINCO31FF-Camila - 2019).

Observa-se em (24) uma sequência argumentativa e, neste excerto, há uma ocorrência de NEG1, “a gente NUM passava metade do que a gente passa hoje”, e uma de NEG2 “mas eu NUM reclamo da vida que tem também NÃO...”.

Nascimento (2014, p. 72) aponta que em sequências argumentativas as variantes inovadoras são desfavorecidas, com pesos relativos baixos – 0.36 para NEG2; e 0.39 para NEG3.

4.1.2.2 Sequência avaliativa

Paredes Silva (1997) se refere a este tipo de sequência como *expressiva*, há a predominância de verbos de “[...] cognição, percepção, volição, sentimento, por exemplo [...]” Paredes Silva (1997, p. 90), exprime-se este tipo de sequência com o verbo no tempo presente e, na maioria dos casos, na primeira pessoa. Em (25) apresentamos este tipo de sequência:

(23) DOC. cê faz muitas viagens assim?
 INF. não... eu **NUM** sô quase de viajá só ficá aqui mesmo...
 DOC. só quando cê tem que trabalhá mesmo fora né?
 INF. só...
 DOC. e cê vai muito em Goiânia... sabe andá em Goiânia?
 INF. não... não... **NÃO** sou de ii muito **NÃO** conheço nada lá...
 quando eu vô fico loquim pra vim embora... (SOCIOLINCO31MF-
 Marcelo - 2019).

Neste excerto, à primeira vista, percebemos que existem muitas ocorrências do **NÃO**, mas algumas delas são respostas diretas, não se caracterizando como uma negação sentencial. No primeiro turno de fala do informante em (5), notamos que ele responde à pergunta de forma avaliativa – ou expressiva –, visto que demonstra que não lhe agrada viajar: “[...] eu num sô quase de viaja [...]”, o verbo está no presente, a pessoa é notadamente P1 e *SER*, neste contexto, tem sentido cognitivo. O mesmo ocorre no último turno de fala do informante no excerto acima.

Nascimento (2014, p. 73) aponta o favorecimento de NEG1 e NEG3 em sequências discursivas deste tipo. Assim, acreditamos que este resultado deve se repetir em nossos dados, havendo favorecimento para NEG1.

4.1.2.3 Sequência descritiva

As sequências descritivas se caracterizam pela presença de estruturas nominais, verbos imperfectivos denotando estado (PAREDES SILVA, 1997, p. 90). Em (26) podemos observar esse tipo de sequência:

(24) DOC. É... você que cuidô dela até agora tem que orientá mesmo

INF. Pois é... Nossa Senhora eu morro de medo... ela e minha tia tudo já teve na pele... minha tia entrô até em depressão quando ela descobriu sabe? qu/ela tava... minha mãe não minha mãe foi mais forte...

DOC. E ela é branquinha assim?

INF. Minha mãe? não, **NUM** é branquinha **NÃO**... morena claro... (GEF28FF).

No excerto observamos a ocorrência de uma sequência descritiva, a informante, ao ser indagada a respeito do tom de pele da mãe, nega a informação apresentada pelo documentador e num ato descritivo: verbo denotando estado com um predicativo como complemento.

Acreditamos que haverá, também, um favorecimento de NEG1 nesses contextos.

4.1.2.4 Sequência dialógica

As sequências dialógicas se caracterizam “pela alternância de participantes/pessoas do discurso envolvidas” (PAREDES SILVA, 1997, p. 90). Em (7) esse tipo de sequência fica evidente:

(25) DOC. aí cê num tem uma religião que cê vai?

INF. NÃO... NÃO vô mais... uma que meu tempo tamém NUM dexa...

DOC. uhum... é... trabalha muito né?

INF. e ota que ah NÃO... a gente desgosta... a gente escuta tanta coisa... (SOCIOLINCO 31FF)

Esse tipo de sequência ocorre, segundo Nascimento (2014, p. 58), com perguntas e respostas (diretas) ou conversação.

Segundo Nascimento (2014, p. 72), esse tipo de sequência discursiva favorece significativamente as ocorrências de NEG2 – peso relativo de 0.78. A hipótese é que o mesmo padrão se repita na análise de nossos dados.

4.1.2.5 Sequência narrativa

Este é o tipo de sequência mais comum nas ocorrências encontradas. Segundo Paredes Silva (1997, p. 90): “uma estrutura narrativa será caracterizada pelo verbo no pretérito perfeito em predicados de ação, em torno de eventos referentes à primeira ou à terceira pessoas, sintaticamente organizadas em orações com junção temporal”. No excerto (28) há uma ocorrência de narrativa:

(26) DOC. hum...
 INF. Deus que me livre...
 DOC. aí cê ficô muito tempo casada?
 INF. fiquei cinco anos casa com ele...
 DOC. hum...
 INF. **aí eu larguei dele na marra... porque ele NUM queria dexá... tive que ii morá com a minha irmã... minha irmã morava nuã casinha menor que essa daqui... aí minha irmã tinha quatro filhos que ela cria do meu irmão e mais três dela... e eu com três filhos pequeno... pensa... nuã casinha minúscula... (SOCIOLINCO 31FF)**

Neste excerto, encontramos uma narrativa de eventos pessoais, ou seja, a informante narra acontecimentos em primeira pessoa, o verbo no passado preenche outro dos requisitos.

Nascimento (2014, p. 72) aponta para o favorecimento da negação anteposta ao verbo quando a sequência discursiva é narrativa. Novamente, acreditamos que o mesmo acontecerá em nossos dados, ou seja, esse tipo de sequência não favorecerá NEG2.

4.1.2.6 Sequência procedural

Esse tipo de sequência se caracteriza pelo uso do imperativo e/ou formas no futuro e infinitivo. O foco é o processo, assim o sujeito acaba por ser impessoal ou genérico. Esse tipo de sequência é pouco frequente nas amostras. *São apenas 16 ocorrências, no total*, e, por isso, será controlada apenas qualitativamente. Em (9) apresentamos uma ocorrência que se caracteriza por este tipo de sequência discursiva:

(27) DOC. Sei...
 INF. Ela NUM é vermeinha
 DOC. Mais parece qu/eles tão fazeno agora modelo de panela diferente... tem a vermelhinha né? e uma mais PREta... cê já viu
 INF. Aquela é porque cura ela né?
 DOC. Como assim?
 INF. **Pra í direto pro fogo... que essas vermeinha que a gente... a gente fais... ela NUM pode i direto no fogo... no forno... cê tem que passá um óleo nela por fora... por dentro e pô água... e dexâ fervê algumas horas sabe?**
 DOC. Ham
 INF. **Alguns minutos... que cê já pegá ela e pô direto no forno ela estora... ela NUM... NUM guenta... aí sempre cê tem que fazê fritá um toicim... ô se NÃO colocá matega e dexá ela fervê algum tempo... aqui NÃO... quando eu vô pegá va... é... mexê... quando**

eu quero uma panela... eu ponho ela já direto no fogo... na fornaia NÃO estora... (GEF28FF).

Observa-se aí que a informante está explicando o procedimento necessário para que se fazer as panelas de barro, famosas na região.

Como este tipo de sequência não se encontra facilmente em entrevistas sociolinguísticas, não temos revisão de literatura para fundamentar a hipótese.

4.1.3 Tipo de oração

O tipo de oração é uma variável que pode influir na ocorrência da negação. Nascimento (2014, p. 64) demonstra que, na variedade capixaba do PB as orações principais favorecem NEG1 e orações absolutas favorecem NEG2. Em contraste a isso, encontramos em Goldnaldel *et al.* (2013, p. 53) a tendência de NEG2 e NEG3 serem favorecidas por orações absolutas.

A este respeito, nas amostras analisadas, encontramos ocorrências para as estruturas: absolutas; coordenadas; subordinadas; e principais.

(28) INF. Ambição... é o zoí grande... o sempre eu falo lá im Uruaçú ond/ela mora lá casona boa qu/ela tem lá, tem duas casa queria qu/eu fosse pra lá, falei não vô... vô ficá queto aqui em Goiás... **eu NUM fui nascido aqui em Goiás mais fui criado aqui** tento de Goiás aqui dento de Goiás eu num saio não... (GEF36MF).

No excerto (30) observamos a ocorrência de uma oração coordenada.

(29) Inf. Mais Deus determinô aquilo pra mim. Determinô aquilo **e/u NUM sei fazê otra coisa**

Em (31) observamos uma oração absoluta, sintaticamente independente.

(30) INF. não... mais a gente que invento esse apilido pra ela... ela era uma criança tamém ela é da minha idade... **eu NUM sei [se ainda é viva hoje]**... mais eu acredito que seja... ai inventamo esse apilido pra ela... mai é um apelido cê vê... de mal gosto tamém mais só que a gente brincava né eu peguei e cai nessa ()... (GEF36MF).

Em (32) a construção que observamos é composta por subordinação. O verbo da oração principal *SABER* tem como seu argumento interno uma oração subordinada introduzida pela conjunção *SE*. Consideramos, neste caso exemplar, a oração principal para análise estatística, visto que é sobre ela que incide a negação – claro que a negação se estende ao argumento interno, mas, como veremos, se a negação

incidisse apenas no argumento interno constituído por uma oração subordinada, consideraríamos a oração subordinada no momento da classificação.

(31) Inf. É uma coisa assim qu/eu fico sem jeito de falá... sempre eu falo pra ela... ela pega e fala esque::ce esse trem mais... **eu fico guardando é uma coisa [que NUM tem como eu ficá falano]...** (GEF36MF).

Em (33) observamos uma oração subordinada, introduzida pela conjunção *QUE*, que funciona como adjunto adnominal de *COISA*. Neste caso, o advérbio de negação incide somente no verbo da oração subordinada, assim, consideramos a oração subordinada para a análise. Acreditamos, diante do exposto, que NEG2 será favorecido por orações absolutas.

4.1.4 Marcador discursivo

A variável *marcador discursivo* foi controlada, visto que alguns trabalhos, como o de Nascimento (2014, p. 75) e Rocha (2013, p. 63), demonstram o desfavorecimento desses marcadores para as ocorrências de NEG2 e NEG3.

Segundo Freitag (2007, p. 23) o marcador discursivo:

[...] é um rótulo amplo que recobre construções que atuam tanto no plano textual, estabelecendo elos coesivos entre partes do texto, como no plano interpessoal, mantendo a interação falante/ouvinte e auxiliando no planejamento da fala.

Assim, para a análise qualitativa, controlamos a presença e/ou ausência dos marcadores nas ocorrências extraídas das amostras.

(32) Inf. cunhicia porque a gente era mais nova **NUM tinha ideia né?** (SOCIOLINCO 46-2FF).

(33) Inf. ai antigamente assim tinha o Saiog tinha forró ali eu sô mais assim **NUM gosto desses batidão NÃO sabe?** Nunca gostei... mas aí o Saiog eu ía muito lá... festa na praça também eu ía... (SOCIOLINCO40FF).

Em (34) verificamos uma ocorrência de NEG1 em que há um marcador discursivo e em (35) o mesmo vale para a ocorrência de NEG2.

(34) Inf. aí eu fui embora... fui pra Goiânia... aí eu fiquei... **quando eu NUM tinha meus filhos eu morava mais lá do que aqui....** (SOCIOLINCO 40FF).

(35) Inf. aí até minha enteada a enteada do meu meninim lá a B. que mora com meu fii lá ela falô que era pra eu i pra lá né leva os trem pra

ela i me ensiná mas té hoje **eu NÃO fui lá NÃO...** (SOCIOLINCO46-2FF).

Em (36) há uma ocorrência de NEG1 sem marcador discursivo, o mesmo vale para (37), uma ocorrência de NEG2. A hipótese é que, em nossos dados, a presença de marcadores discursivos desfavoreça a escolha pelas variantes inovadoras.

4.1.5 Tipo de sujeito

O não preenchimento do argumento externo pode, por hipótese, favorecer a ocorrência de NEG2. Nascimento (2014, p. 87), em sua pesquisa, constatou que, na variedade do PB de Vitória (ES), o tipo de sujeito não exerce grande influência para as formas não canônicas de negação.

Para este trabalho, consideramos três tipos de sujeito: explícito; implícito (oculto; ou indeterminado); e oração sem sujeito.

(36) Inf. aí no caso eu precisava né? e minha mãe era doente tinha problema de rins... **minhas irmã era mais nova NUM trabalhava né?** e eu que assumi a despesa todinha de casa (SOCIOLINCO47MF- - 2019)

(37) Inf. é... que se não **NUM ganha nada né?** (SOCIOLINCO47MF- - 2019)

(38) DOC. hum... e plantava lá como é que era?
INF. prantava... tinha roça grande né? milho esses trem... criava bastante capado... né? tinha um manguerão... tem lá até hoje só que porque hoje meu tii mora lá num tem poi... **lá NUM tem nada né?** (SOCIOLINCO42MF)

Em (38) observamos uma oração com o sujeito explicitamente realizado: “minhas irmã [...] **NUM** trabalhava né”. Em (39), por outro lado, o sujeito não está explicitamente posto, mas pode ser recuperado pelo contexto e pela desinência verbal: “é... que se não **NUM** ganha nada né?”. Por fim, em (40) o verbo *TER*, neste caso, ao exprimir o sentido de “existir”, não possui argumento externo, sendo considerado sem sujeito. Dessa forma, a hipótese que levantamos é que, conforme Nascimento (2014), o tipo de sujeito não influirá significativamente nas ocorrências de negação.

No capítulo 5, serão expostos os resultados obtidos, a partir da análise estatística no R, para cada uma das variáveis descritas, a fim de confirmar ou refutar as hipóteses elencadas.

4.2 Variáveis sociais

Considerando, então, os diversos fatores que influenciam a variação de formas linguísticas, nesta subseção são apresentadas as variáveis sociais: Amostra; Sexo; Faixa etária e Informante.

4.2.1 Amostra

A constituição do material a analisado neste trabalho se deu por meio de entrevistas de duas amostras – uma delas pertencente ao Grupo de Estudos Funcionalistas, da Universidade Federal de Goiás (GEF/UFG); e, a outra, ao Grupo de Estudos e Pesquisas em Sociolinguística, da Universidade Estadual de Goiás (Sociolinco/UEG/Cora Coralina). Os detalhes a respeito de cada uma dessas amostras foram discutidos em 3.2. Cabe, aqui, apresentar as possibilidades que surgem ao se abordar dados da mesma comunidade de fala, gravados com certo intervalo de tempo e seguindo os mesmos padrões de entrevistas.

As entrevistas foram coletadas com um intervalo de 16 anos, o que permite a análise de tendência em tempo real. A pesquisa em tempo real se caracteriza por uma amostra randômica de uma determinada comunidade de fala que, depois de certo período é repetida. Assim, coletaram-se entrevistas no ano de 2003 e, em 2019, gravaram-se outras entrevistas seguindo o mesmo roteiro e com informantes de mesmo perfil. A pesquisa de tendência se difere da de painel, pois esta exige que a entrevista seja realizada com o mesmo informante abordado na primeira e aquela aborda outros informantes que possuam o mesmo perfil (FREITAG, 2005). Embora não permita abordar os mesmos informantes, essa metodologia viabiliza a análise da língua falada pela comunidade em duas épocas distintas, possibilitando uma abordagem panorâmica do fenômeno.

Assim, a partir da amostra coletada em 2019, é possível comparar as ocorrências – variável resposta – com as variáveis previsoras a fim de averiguar se há alguma mudança nesse período no peso exercido por cada uma delas nesses dois momentos. Além disso, fizemos a comparação direta entre as ocorrências encontradas na amostra do GEF às ocorrências encontradas na amostra do Sociolinco – o que será apresentado na seção seguinte.

4.2.2 Sexo

Para este estudo, foram abordados informantes de ambos os sexos, sendo 6 homens e 6 mulheres para cada um dos bancos de dados. Assim, no total, as entrevistas foram realizadas com 12 informantes de cada sexo. Precisamos pontuar que não houve o controle do gênero, “abordagem sociocultural” (SCHERRE; YACOVENCO, 2011, p. 122).

Segundo Labov (2001 *apud* SCHERRE; YACOVENCO, 2011, p. 122), ao se analisar a influência dessa variável em alguns fenômenos de variação, constata-se que as mulheres são mais conservadoras quando o fenômeno ocorre acima do nível da consciência social e preferem as formas inovadoras quando o fenômeno em questão está abaixo do nível da consciência.

Scherre e Yacovenco (2011) discutem a questão do peso da variável sexo à luz de uma série de pesquisas que abordam o fenômeno da variação entre *tu* e *você* em várias regiões do Brasil. A questão central, para as autoras, é porque há essa diferença de comportamento, ora progressista, ora inovador.

A partir dessas constatações, e de outras relativas a mais dois fenômenos, as autoras propõem o seguinte no que diz respeito à variável gênero:

- 1) Traços linguísticos menos marcados, no sentido de serem menos dependentes das relações interacionais ou mais frequentes ou mais aceitos socialmente, tendem a ser favorecidos pelas mulheres [...].
Generalização: em configurações menos marcadas - e não necessariamente mais prestigiadas - as mulheres estão à frente na variação ou na mudança.
- 2) Traços linguísticos mais marcados, no sentido de serem mais dependentes das relações interacionais ou menos frequentes ou menos aceitos socialmente, tendem a ser favorecidos pelos homens[...].
Generalização: em configurações mais marcadas - e não necessariamente menos prestigiadas – os homens estão à frente na variação ou na mudança.

Em outras palavras, as mulheres estarão à frente quando o fenômeno variável for menos marcado. O fenômeno aqui é estudado ocorre, notadamente, abaixo do nível da consciência e, por isso, podemos considerá-lo como não marcado.

Nascimento (2014, p. 83) afirma que NEG2 é levemente favorecida em ocorrências com informantes do sexo feminino. Rocha (2013, p. 66), por sua vez, não aponta uma diferença significativa na variação da negação em seus dados. Reimann e Yacovenco (2011) apontam que os homens favorecem as variantes inovadoras, tendo sido feita uma rodada no *GoldVarb X* com NEG2 e NEG3 amalgamadas.

Levando o que foi discutido em consideração, a hipótese que se levanta, a respeito desta variável, é a de que não haverá diferença significativa nas escolhas das

estratégias de negação entre homens e mulheres. Isso porque, conforme argumentam Scherre e Yacovenco (2011), o fenômeno em questão não parece sofrer nenhum tipo de estigma social. Assim sendo, a variação de formas ocorrerá de forma equilibrada entre os falantes de ambos os sexos.

Assim, na seção 5, abordaremos novamente esta questão, a fim de averiguar se essa proposição também se aplica para o fenômeno que tomamos como objeto para este trabalho.

4.2.3 Faixa etária

Abordaremos a variável faixa etária, neste trabalho, dividindo os informantes em dois grupos: 30-45 anos; e 46-60 anos. A faixa etária não se refere apenas à idade do informante, ela pode apontar questões como a inserção ou não do mesmo no mercado de trabalho, o grau de escolaridade e as relações que ele mantém na sociedade (FREITAG, 2005, p. 106).

Além disso, vale comentar que a análise em tempo real, como é o caso das amostras utilizadas aqui, permite observar falantes da mesma faixa etária em momentos distintos. Ou seja, aquele que possuem 46 anos ou mais e foram entrevistados em 2019 – para constituição da amostra Sociolingo – pertenciam à primeira faixa etária (30-45 anos) à época da pesquisa realizada pelo GEF.

Além disso, Freitag (2005, p. 109) afirma que:

Se o comportamento linguístico dos indivíduos é estável durante toda a sua vida e o comportamento linguístico da comunidade também, a intuição do pesquisador falhou e não há variação a analisar: há estabilidade, como reflete o padrão (1). Já se os indivíduos mudam seu comportamento linguístico durante o decorrer da sua vida e a comunidade não mostra a mesma mudança, o padrão é caracterizado como gradação etária, como o padrão (2). Os padrões (3) e (4) não são transparentes, requerem um controle mais refinado para serem identificados. O padrão (3) reflete mudança geracional, no qual alguns membros da comunidade desencadeiam a mudança. E o padrão (4) reflete mudança comunitária, no qual toda a comunidade desencadeia a mudança.

Rocha (2013, p. 66), novamente, não encontrou diferença significativa entre as ocorrências de acordo com a faixa etária. Em Nascimento (2014), a variável aqui discutida não foi selecionada, ou seja, também não apresenta grande disparidade entre as faixas etárias. Já em Reimann e Yacovenco (2011), o peso relativo da 1ª faixa etária (15-25 anos) é maior que na outra faixa etária controlada (>49 anos). Em nossa

pesquisa, a hipótese que aventamos é a de que as formas inovadoras são mais frequentes entre os mais jovens.

Cabe, então, a análise das ocorrências à luz das faixas etárias selecionadas a fim de verificar se há instabilidade entre elas. A hipótese que levantamos, quanto à faixa etária dos informantes, é que não haverá grande diferença entre as gerações. Ou seja, acreditamos que a variação entre as estratégias de negação na variedade vilaboense do PB é um fenômeno estável.

4.2.4 Informante

Além das variáveis sociais previsoras apresentadas acima, decidimos analisar se há variação de uso significativa entre os informantes. Logo, como são 24 informantes, 12 para cada uma das amostras, analisamos o uso variável da negação e da dupla negação entre os indivíduos.

Faremos isso a fim de verificar os possíveis *outliers* e entender como cada um deles influenciou no resultado que obtivemos a partir da análise estatística. Essa avaliação será realizada a partir da análise do perfil dos informantes e das ocorrências das variantes inovadoras. Conforme discutimos em 1.2, tendo como referência Neves (2018), a língua não é autônoma, o falante, na interação social, é parte determinante da língua. Segundo Machado (2010, p. s/p):

Uma teoria social da linguagem deve perceber a langue como um processo de convencionalização, e a competência do falante individual não pode ser vista apenas como internalização de uma convenção. A prática linguística não é apenas o uso consensual de um sistema, pois uma convenção é um processo em que o indivíduo age sobre o outro e faz hipóteses sobre seu comportamento; ou seja, o sujeito falante é um agente criador de convenções.

Tendo essa concepção, de que o sujeito é o agente que cria, aceitando ou não as convenções linguísticas, é que analisaremos esta variável. Isso porque, segundo Eckert (2000 *apud* MACHADO, 2010, p. s/p): “[...] o estudo da variação em comunidade [...] explicita o uso da linguagem vernacular e/ou padrão de acordo com uma estratificação socioeconômica que, depois, passa a ser ampliado para um esquema de investigação de modos de vida”. O objetivo é compreender se há alguma diferença significativa entre os usos da negação entre os informantes.

5. ANÁLISE DOS DADOS

Nesta seção apresentaremos a análise dos dados feitas com o auxílio do programa *R*. A seção está dividida em duas subseções: na primeira, expomos a distribuição das ocorrências da variável resposta e sua proporção, de acordo com cada variável previsor; na segunda, analisamos os efeitos de cada variável independente, aqui chamada de variável previsor, nas ocorrências da variável dependente, aqui chamada de variável resposta.

5.1 Análise quantitativa

Dividimos essa subseção em duas partes: na primeira, abordamos as variáveis linguísticas; na segunda, as variáveis sociais. Encontramos, nas duas amostras analisadas, um total de 3029 ocorrências de negação sentencial, dessas 1909 são de NEG1, 516 são de NEG2 e 88 de NEG3. Dessa forma, as ocorrências estão distribuídas da seguinte forma:

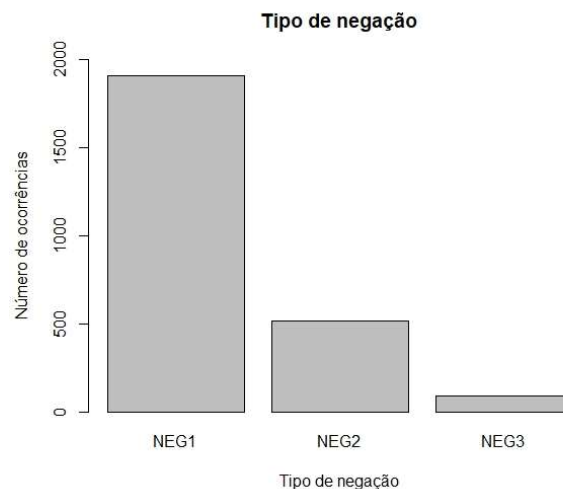
Tabela 3 – Distribuição geral dos dados com NEG3

	Frequência	Porcentagem
NEG1	1909	75,9%
NEG2	516	20,5%
NEG 3	88	3,6%

Fonte: Elaboração nossa.

A baixa ocorrência de NEG3 na amostra se evidencia. O contraste das três formas alternantes fica bastante claro ao observarmos o gráfico a seguir:

Gráfico 1 - Total de ocorrências com NEG3



Fonte: Elaboração nossa.

Nota-se que NEG3 não ocorre o suficiente para ser estatisticamente relevante, assim sendo, consideraremos um universo de 2941 ocorrências, das quais 1909 são de negação pré-verbal e 516 de dupla negação. A distribuição dos dados pode ser observada na tabela a seguir:

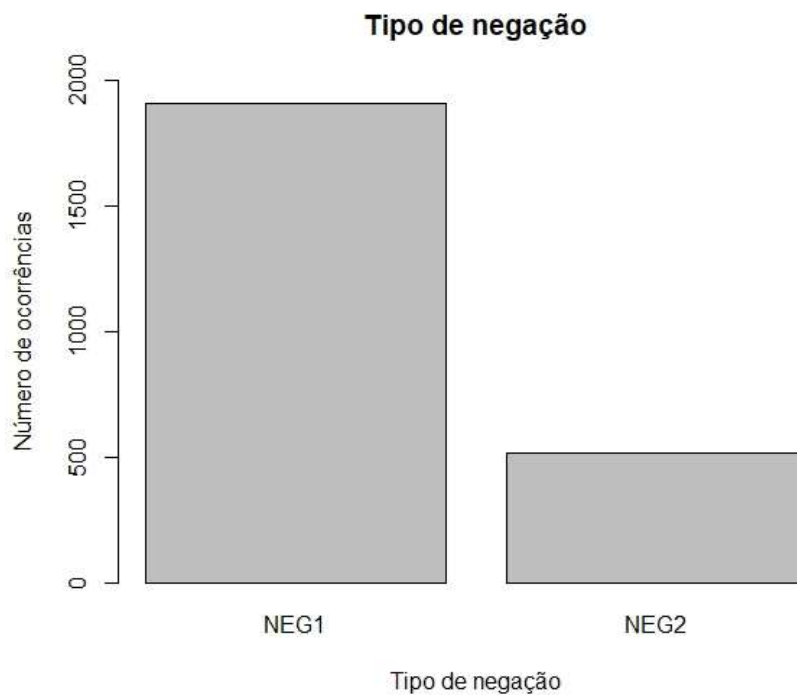
Tabela 4 – Distribuição geral dos dados sem NEG3

	Frequência	Porcentagem
NEG1	1909	78,7%
NEG2	516	21,3%

Fonte: Elaboração nossa.

Destarte, observamos que NEG1 ocorre em aproximadamente 78,7% das vezes e NEG2 em 21,3%. O gráfico abaixo apresenta a distribuição geral dos dados:

Gráfico 2 - Total de ocorrências



Fonte: Elaboração nossa.

Tal fato demonstra como este fenômeno de variação ocorre diferentemente em regiões distintas do país. Yacovenco e Nascimento (2016, p. 134-135) sintetizam a distribuição das ocorrências a partir de pesquisas realizadas em diferentes cidades brasileiras:

Tabela 5 – Distribuição da negação sentencial em diferentes variedades linguísticas

Cidades	Não V %	Não V Não %	V Não %
Fortaleza	625/774 77	149/774 18	39/774 5
Natal (conversacional)	308/466 66,1	96/466 20,6	62/466 13,3
Vitória	1751/2263 77,4	478/2263 21,1	34/2263 1,5
Mariana	1787/2505 71,5	489/2505 19,5	40/2505 1,5
São Paulo	5279/5607 94	354/5607 5,8	4/5607 0,2
Curitiba	1371/1408 97,4	37/1408 2,6	-
Florianópolis	1018/1065 95,6	47/1065 4,4	-
Porto Alegre	1402/1410 99,4	8/1410 0,6	-
Goiás-GO	1909/3029 75,9	516/3029 20,5	88/3029 3,6

Fonte: Adaptado de Yacovenco e Nascimento (2016, p. 134-135).

A partir desses dados, observamos uma distribuição dos dados parecida nas cidades localizados nas regiões Nordeste e Centro-Oeste. Ainda a respeito desse fenômeno de variação na região nordeste, Lopes (2021, p. 62) apresenta um quadro sintético das ocorrências encontradas em amostras de língua oral coletadas no estado da Bahia:

Tabela 6 – A negação nas mesorregiões baianas – análise geral de estratégias

	Dados/Total	Percentual
Pré-Verbal (Não V)	1500/2322	64,6%
Dupla (Não V Não)	532/2322	22,9%
Pós-Verbal (V Não)	290/2322	12,5%

Fonte: Adaptado de Lopes (2021, p. 62).

Observamos, aí, uma tendência maior de NEG3, ao passo que as ocorrências da negação pré-verbal e dupla negação são similares às encontradas em Goiás (GO). Ainda a respeito dos dados sintetizados por Yacovenco e Nascimento (2016), há, também, a variação de formas semelhante em Mariana (MG) e Vitória (ES), municípios que se localizam na região sudeste. Em contrapartida, ainda no Sudeste, em São Paulo (SP) a ocorrência das formas inovadoras é menos comum. Essa tendência continua rumo ao sul do país, havendo menor número de ocorrências de NEG2 em Porto Alegre (RS).

5.1.1 Variáveis linguísticas

Nesta subseção, apresentamos os resultados da análise quantitativa e visualização da distribuição dos dados.

5.1.1.1 Status da informação

O *status* discursivo da informação negada é bastante relevante, visto que, conforme argumentamos em 4.1.1, as variantes terão o mesmo valor de verdade em contextos específicos. Encontramos apenas duas ocorrências de NEG2 quando a informação negada é nova no discurso. Assim, diante do inexpressivo número de ocorrências, entendemos que esse nível dessa variável não deve constituir o envelope de variação. Dessa forma, as ocorrências das estratégias de negação se deram conforme a tabela a seguir:

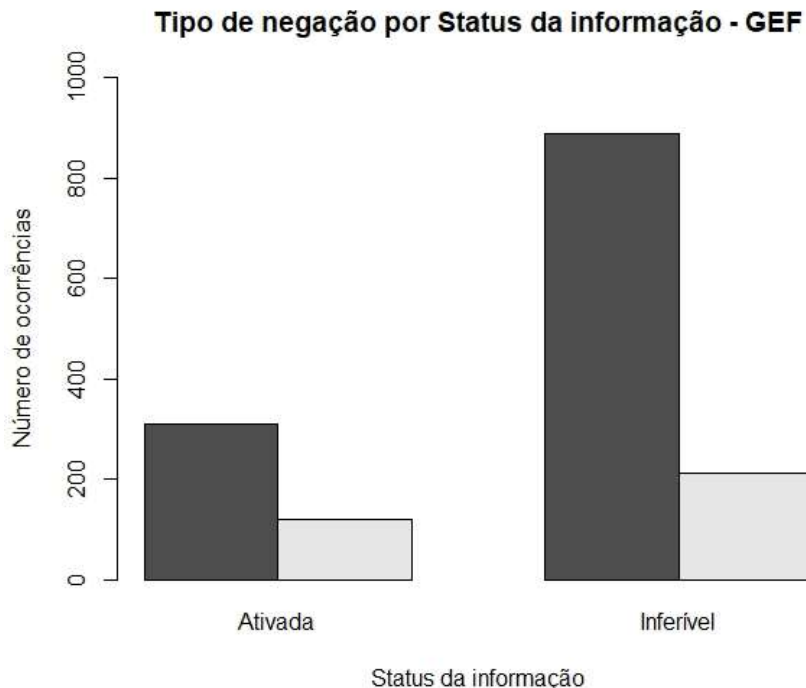
Tabela 7 – Distribuição dos dados pelo Status discursivo da informação negada - GEF

	GEF			
	Ativada		Inferível	
	Frequência	Porcentagem	Frequência	Porcentagem
NEG1	311	72.2%	888	80.7%
NEG2	120	27.8%	213	19.3%

Fonte: Elaboração nossa.

O gráfico a seguir permite visualizar melhor os dados:

Gráfico 3 - Total de ocorrências pelo Status da informação negada - GEF



Fonte: Elaboração nossa.

De pronto, é possível inferir, observando os números absolutos, que as ocorrências se concentram, principalmente, quando a informação negada no discurso é inferível, ou seja, a informação foi indiretamente ativada no discurso. O menor número de ocorrências na amostra GEF foi de NEG2 com a informação diretamente ativa. No entanto, ao observarmos os números proporcionais, verificamos que a distribuição dos dados é maior, entre NEG1 e NEG2, quando a informação é diretamente ativada no discurso.

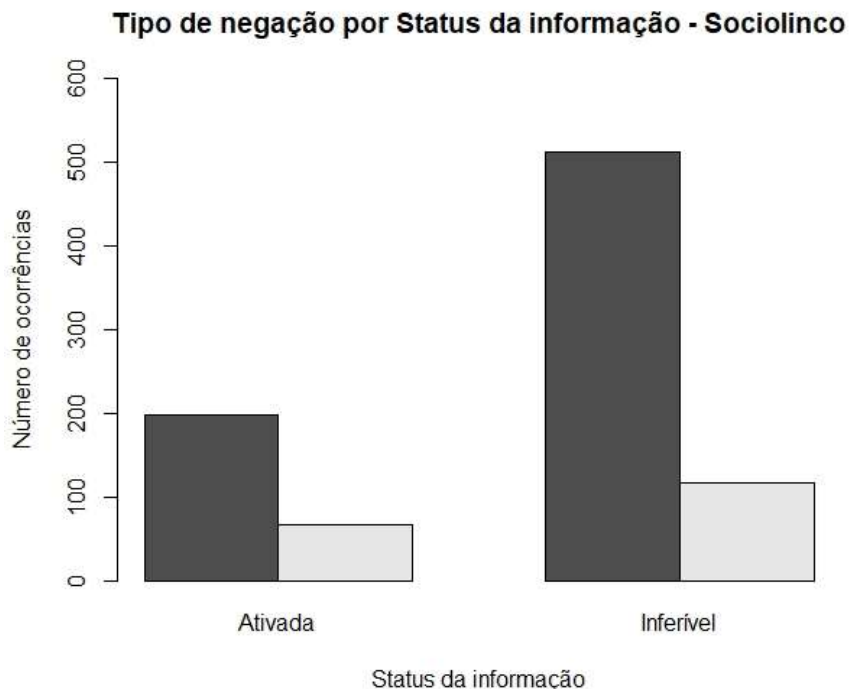
Tabela 8 – Distribuição dos dados pelo Status discursivo da informação negada - Sociolinco

	SOCIOLINCO			
	Ativada		Inferível	
	Frequência	Porcentagem	Frequência	Porcentagem
NEG1	198	74.7%	512	81.5%
NEG2	67	25.3%	116	18.5%

Fonte: Elaboração nossa.

O gráfico a seguir sintetiza o que apresentamos na tabela:

Gráfico 4 - Total de ocorrências pelo Status da informação negada - Sociolinco



Fonte: Elaboração nossa.

Observamos, na amostra Sociolinco, a mesma tendência observada na amostra GEF, ou seja, proporcionalmente NEG2 ocorre mais vezes com a informação diretamente ativada no discurso.

5.1.1.2 Tipo de sequência discursiva

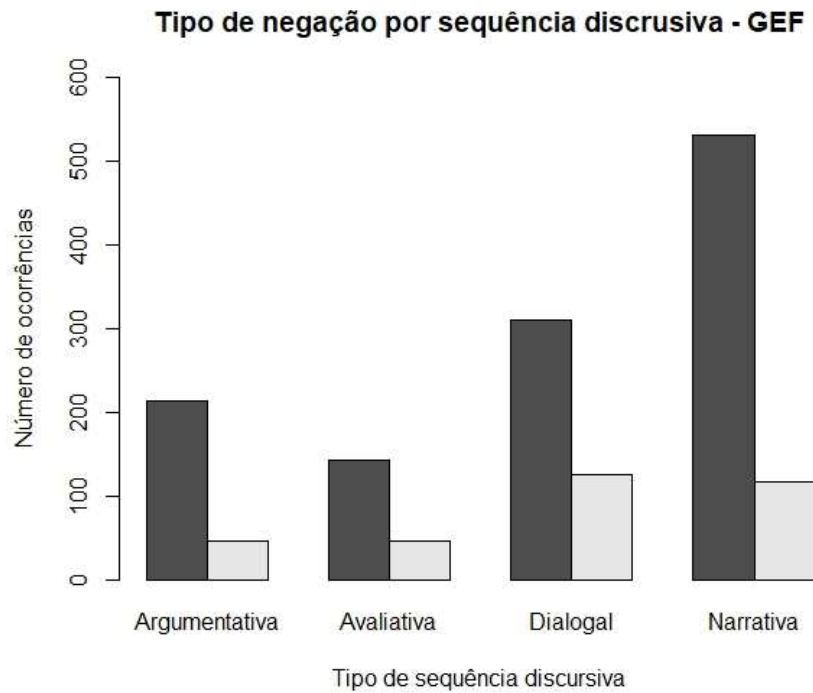
Na tabela a seguir, apresentamos a distribuição dos dados de acordo com o tipo de sequência discursiva em que as estratégias de negação ocorreram, considerando as amostras:

Tabela 9 – Distribuição dos dados pelo Tipo de sequência discursiva por amostra

		NEG1		NEG2	
		Frequência	Porcentagem	Frequência	Porcentagem
GEF	Argumentativa	214	82.3%	46	17.7%
	Avaliativa	143	75.7%	46	24.3%
	Dialógica	310	71.3%	125	28.7%
	Narrativa	532	82.1%	116	17.9%
SOCIOLINCO	Argumentativa	179	88.2%	24	11.8%
	Avaliativa	23	59 %	16	41.0%
	Dialógica	235	75.1%	78	24.9%
	Narrativa	273	80.8%	65	19.2%

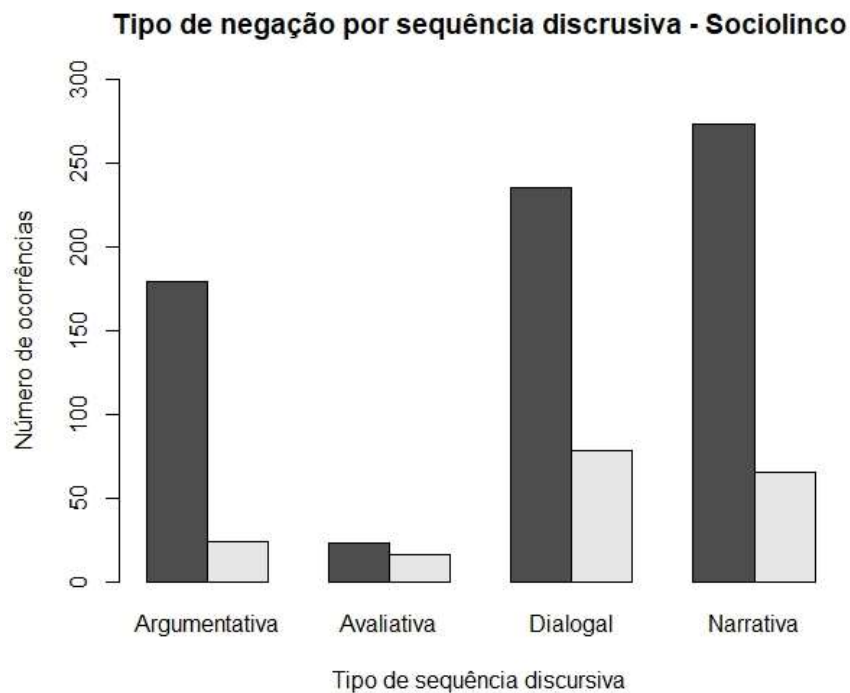
Fonte: Elaboração nossa.

Os gráficos a seguir apresentam as informações apresentadas na tabela, o primeiro apresenta as informações da amostra GEF:

Gráfico 5 - Total de ocorrências pelo tipo de sequência discursiva - GEF

Fonte: Elaboração nossa.

O gráfico a seguir apresenta as informações da amostra Sociolingo:

Gráfico 6 - Total de ocorrências pelo tipo de sequência discursiva - Sociolingo

Fonte: Elaboração nossa.

O primeiro ponto que nos chama a atenção é o fato de o tipo procedural possuir apenas ocorrências de NEG1 na amostra de 2003 e ter ocorrido tão poucas vezes e, por isso, será excluída da análise. Notamos que o tipo de sequência discursiva argumentativa, na amostra de 2019, concentra a menor proporção das ocorrências e, em contrapartida, na mesma amostra, temos a sequência avaliativa alcançando 41.1% de dupla negação. As sequências que apresentam maior proporção de NEG2 em relação a NEG1 são as dialógicas e procedural, seguidas pela sequência narrativa, que concentra o maior número de ocorrências – considerando números absolutos – de ambas as estratégias de negação. A sequência avaliativa tem o menor número de ocorrências, em números absolutos, em ambas as amostras e considerando as duas estratégias, possui, no entanto, uma distribuição mais equilibrada na amostra mais recente, se comparada à de 2003.

5.1.1.3 Tipo de oração

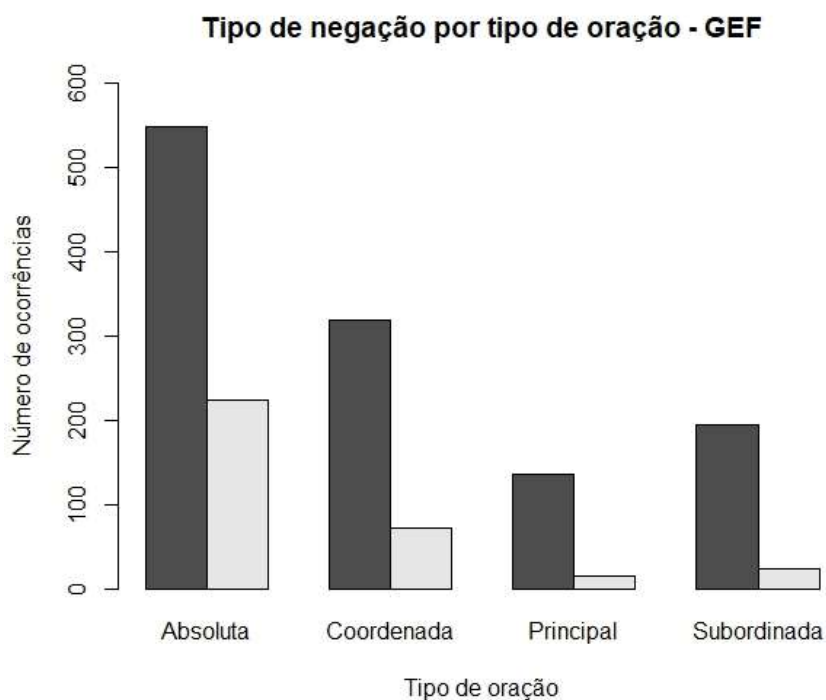
Na tabela abaixo apresentamos a distribuição das ocorrências de negação no *corpus* de acordo com o tipo de oração:

Tabela 10 – Distribuição dos dados pelo Tipo de oração

		NEG1		NEG2	
		Frequência	Porcentagem	Frequência	Porcentagem
GEF	Absoluta	549	71%	224	29%
	Coordenada	319	81.7%	71	18.2%
	Principal	136	90.7%	14	9.3 %
	Subordinada	195	89%	24	11%
SOCIOLINGO	Absoluta	350	72.5%	133	27.5 %
	Coordenada	192	85%	34	15%
	Principal	83	95.4%	4	4.6%
	Subordinada	85	87.6%	12	12.4 %

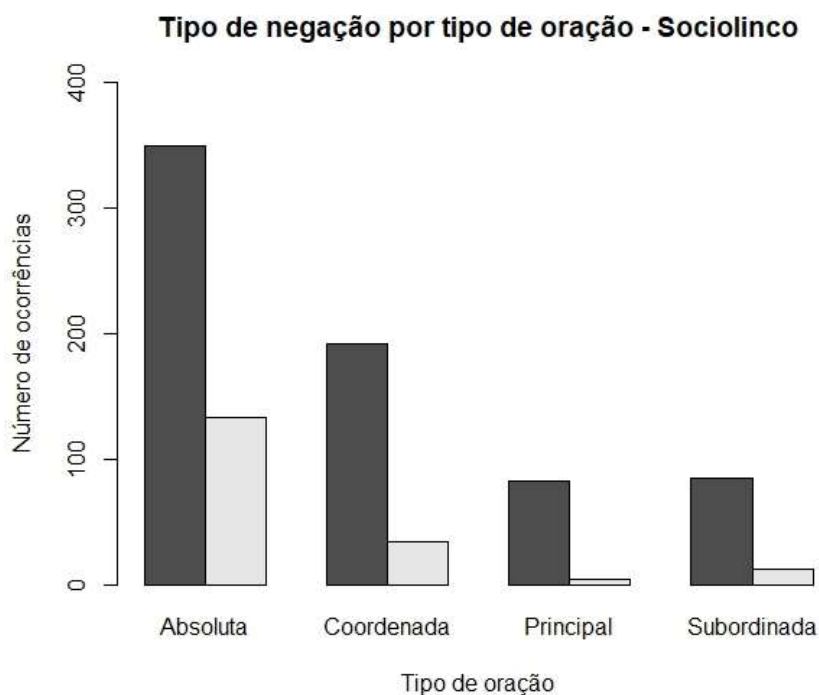
Fonte: Elaboração nossa.

Plotamos o gráfico a seguir com os números absolutos das ocorrências das estratégias de negação, de acordo com o tipo de oração, encontradas na amostra GEF:

Gráfico 7 - Total de ocorrências pelo tipo de oração - GEF

Fonte: Elaboração nossa.

No gráfico abaixo, observamos as ocorrências encontradas na amostra Sociolinco:

Gráfico 8 - Total de ocorrências pelo tipo de oração - Sociolinco

Fonte: Elaboração nossa.

Para esta variável previsora, observamos que há uma concentração maior de ocorrências de NEG2, em ambas as amostras, em um dos níveis: orações absolutas. Chama a atenção a baixa proporção de NEG2 em orações principais na amostra de 2019. Assim, tanto em número absolutos, quanto em proporcionais, NEG2 se tende a ocorrer mais em orações absolutas.

5.1.1.4 Marcador discursivo

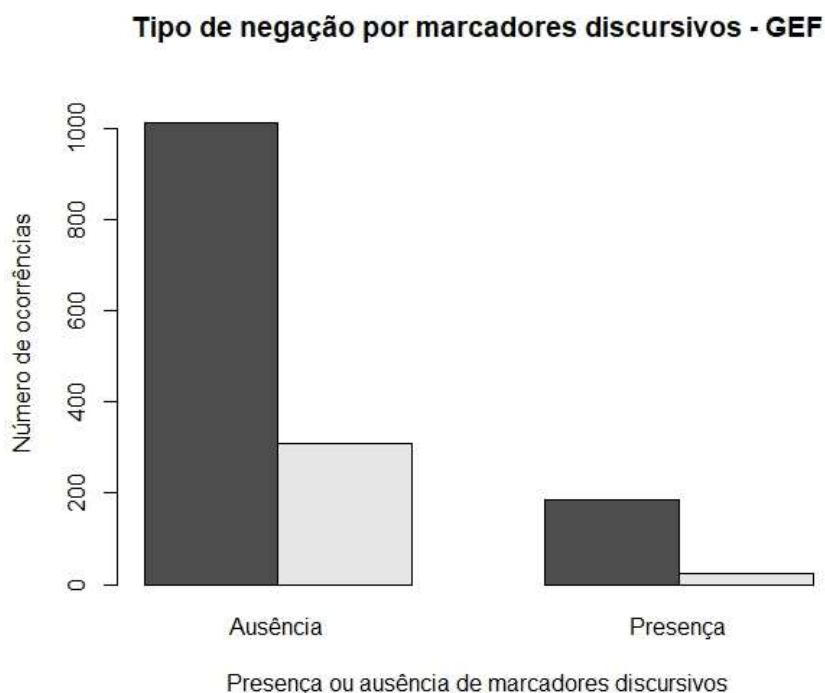
A presença de marcadores discursivos, conforme discutimos em 4.1.4, é um fator que, notadamente, diminui as ocorrências de negação, tanto NEG1 quanto NEG2.

Tabela 11 – Distribuição dos dados pela presença ou ausência de marcadores discursivos - GEF

	GEF			
	Presença		Ausência	
	Frequência	Porcentagem	Frequência	Porcentagem
NEG1	187	88.6%	1012	76.6%
NEG2	24	11.4%	309	23.4%

Fonte: Elaboração nossa.

Este gráfico apresenta a distribuição da negação anteposta ao verbo e da dupla negação de acordo com a presença e ausência de marcadores discursivos na amostra GEF:

Gráfico 9 - Total de ocorrências pela presença ou ausência de marcadores discursivos - GEF

Fonte: Elaboração nossa.

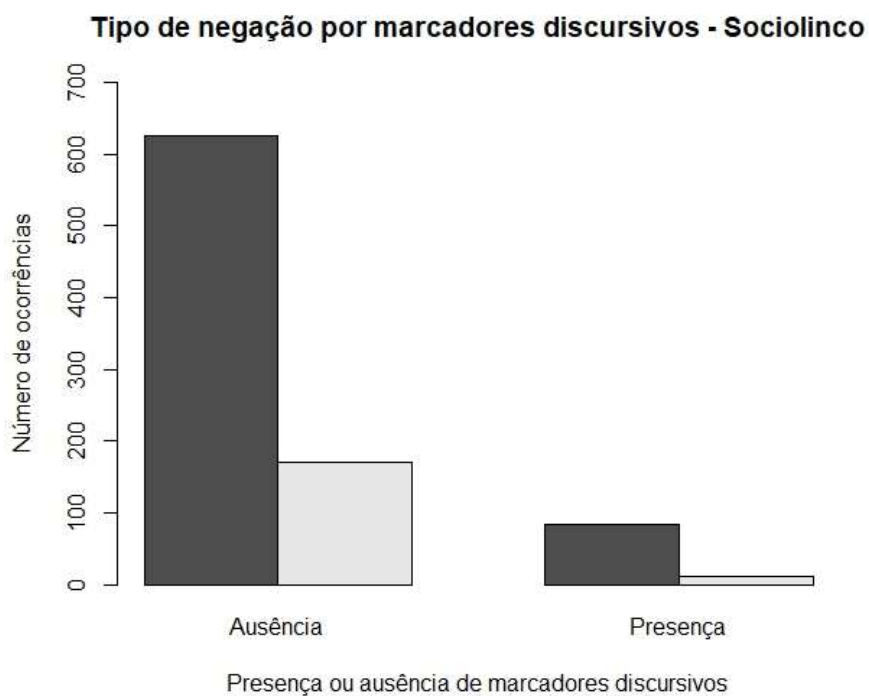
Conforme se observa, a presença de uma sentença negativa parece dispensar o uso de marcadores discursivos nesses contextos. Além da grande diferença em números absolutos, NEG2 ocorre proporcionalmente menos em contexto em que há um marcador discursivo.

Tabela 12 – Distribuição dos dados pela presença ou ausência de marcadores discursivos - Sociolinco

	Sociolinco			
	Presença		Ausência	
	Frequência	Porcentagem	Frequência	Porcentagem
NEG1	84	87.5%	626	78.5%
NEG2	12	12.5%	171	21.5%

Fonte: Elaboração nossa.

De acordo com os dados da tabela acima, podemos visualizar melhor a distribuição dos dados na amostra Sociolinco no gráfico abaixo:

Gráfico 10 - Total de ocorrências pela presença ou ausência de marcadores discursivos - Sociolinco

Fonte: Elaboração nossa.

O mesmo pode ser observado nos dados da amostra Sociolinco, com uma leve diminuição da proporção de ocorrências de NEG1 em contextos em que há marcadores discursivos e com um aumento proporcional em contextos sem marcadores.

5.1.1.5 Tipo de sujeito

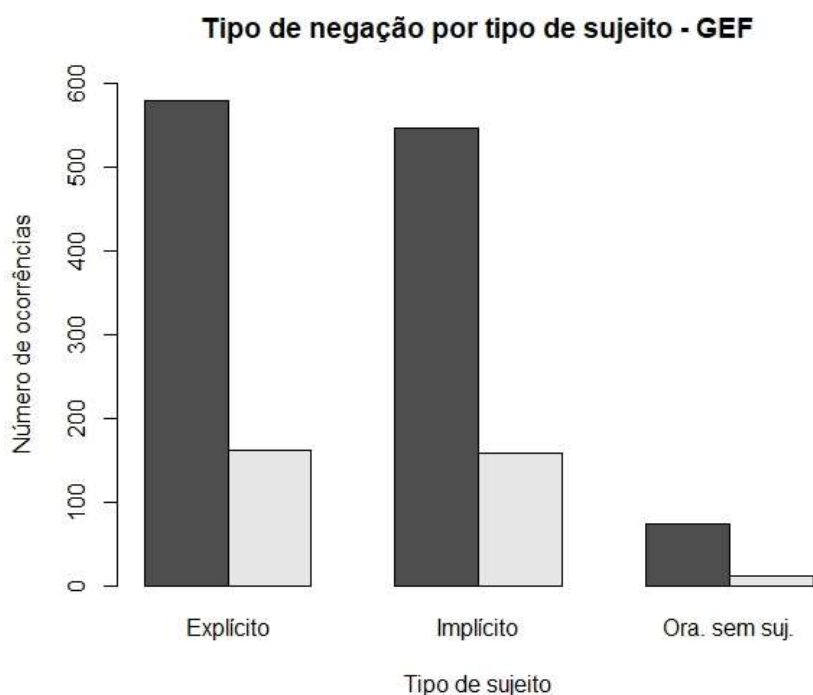
Na tabela a seguir, observamos a distribuição das ocorrências de negação de acordo com o tipo de sujeito das orações em que a negação ocorreu:

Tabela 13 – Distribuição dos dados por tipo de sujeito

		NEG1		NEG2	
		Frequência	Porcentagem	Frequência	Porcentagem
GEF	Implícito	546	77.4%	159	22.6%
	Explícito	580	78.2%	162	21.8%
	Sem sujeito	73	85.9%	12	14.1%
SOCIOLINCO	Implícito	314	77.9%	89	22.1%
	Explícito	365	80%	91	20%
	Sem sujeito	31	91.2%	3	8.8%

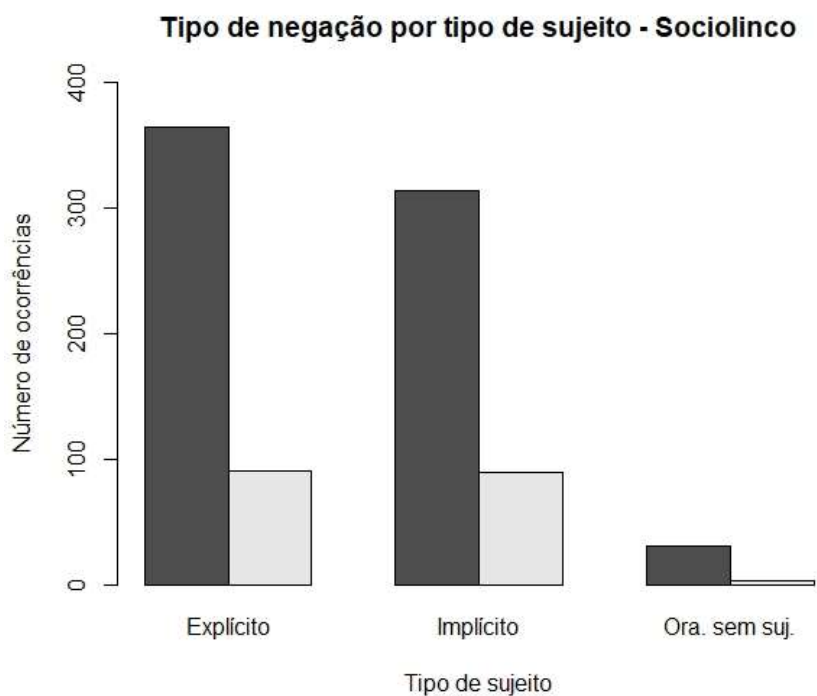
Fonte: Elaboração nossa.

Antes de comentarmos a distribuição dos dados, podemos observá-la nos dados na amostra GEF a partir do gráfico a seguir:

Gráfico 11 - Total de ocorrências pelo tipo de sujeito - GEF

Fonte: Elaboração nossa.

Da mesma forma, a distribuição das estratégias de negação pelo tipo de sujeito, para a amostra Sociolinco, pode ser observada no gráfico a seguir:

Gráfico 12 - Total de ocorrências pelo tipo de sujeito - Sociolinco

Fonte: Elaboração nossa.

O nível, desta variável previsora, que tem o menor número de ocorrências é o que representa as orações sem sujeito. A existência do sujeito, esteja ele implícito ou explícito, apresenta similaridades distributivas, como diferenças de menos de 1% para os dados da amostra GEF entre eles e pouco mais de 2% para os dados da amostra Sociolinco.

5.1.2 Variáveis sociais

Nesta subseção apresentamos os resultados da análise quantitativa dos efeitos das variáveis sociais sobre a variável resposta.

5.1.2.1 Amostra

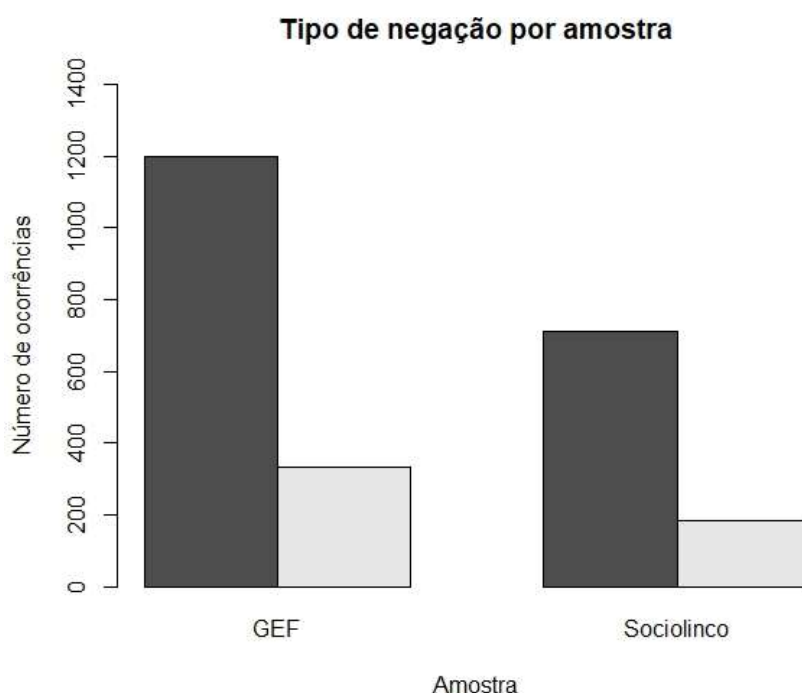
A amostra, como apresentado em 4.2.1, está dividida em dois grupos: GEF e Sociolinco. A distribuição de dados nas amostras ocorre da seguinte forma:

Tabela 14 – Distribuição dos dados pela amostra

	GEF		SOCIOLINCO	
	Frequência	Porcentagem	Frequência	Porcentagem
NEG1	1199	78.3%	710	79.5%
NEG2	333	21.7%	183	20.5%

Fonte: Elaboração nossa.

As ocorrências se concentram mais na amostra coletada em 2003, mas a proporção continua similar, com diminuição de 2% na amostra coletada em 2019. As ocorrências encontradas no Banco de Dados do GEF (UFG) se distribuem, quanto à forma de negação, da forma seguinte:

Gráfico 13 – Total de ocorrências por amostra

Fonte: Elaboração nossa.

Assim sendo, NEG1 surge 1269 vezes, somando 76%, ao passo que NEG2 ocorre 369 vezes, 22% das ocorrências, e, por fim, NEG3 ocorre 39 vezes, 2%. Encontramos, neste banco de dados, 1677 ocorrências no total.

Neste caso, NEG1 surge em 721 vezes, 75% das ocorrências, ao passo que NEG2 ocorre 184 vezes, somando 19% dos dados e, por fim, NEG3 ocorre 52 vezes, 6% do total. O total de ocorrências, encontradas nesse banco de dados, é de 957. É possível verificar uma diminuição das ocorrências de NEG3, aumentando 2% em cada uma das outras duas variáveis.

Além dessas ocorrências, também foram observados casos de *NÃO* que não se relacionavam com um sintagma verbal. Para os dados extraídos em 2003, foram encontradas 208 ocorrências desse tipo e, para os dados de 2019, o *NÃO* isolado ocorreu 317 vezes. Na maioria dos casos, esse advérbio surgia como resposta direta a uma pergunta do documentador, como em:

Doc. e aí aqui em Goiás. Cê já morou em outro lugar além desses dois aqui e lá no João Francisco?
 Inf. não...
 Doc. sempre morô nesse lá?
 Inf. esse... (SOCIOLINCO31FF-Sara)

Essas ocorrências foram descartadas, pois, apesar de ser possível recuperar o que está sendo negado, “[...] morou em outro lugar [...]”, não há como afirmar se o informante optaria por NEG1 caso preenchesse todas as casas argumentais do verbo em sua fala, ao invés de optar pela elipse.

5.1.2.2 Sexo

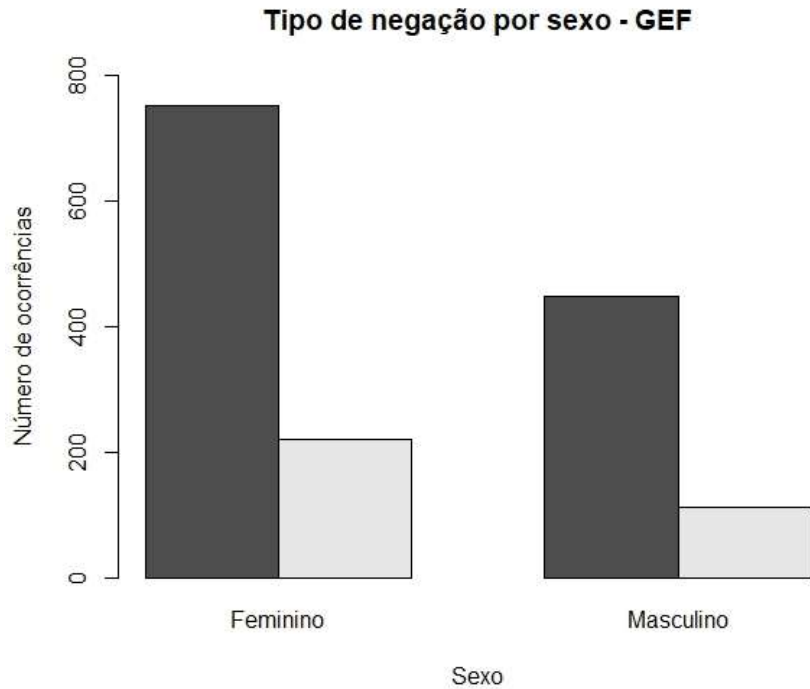
Nesta seção, apresentamos a distribuição dos tipos de negação de acordo com o sexo dos informantes.

Tabela 15 – Distribuição dos dados por sexo - GEF

	GEF			
	Feminino		Masculino	
	Frequência	Porcentagem	Frequência	Porcentagem
NEG1	752	77.2%	447	79.9%
NEG2	221	22.8%	112	20.1%

Fonte: Elaboração nossa.

A distribuição das ocorrências pelo sexo do informante, na amostra GEF, pode ser visualizada no gráfico que segue:

Gráfico 14 – Total de ocorrências pelo sexo do informante - GEF

Fonte: Elaboração nossa.

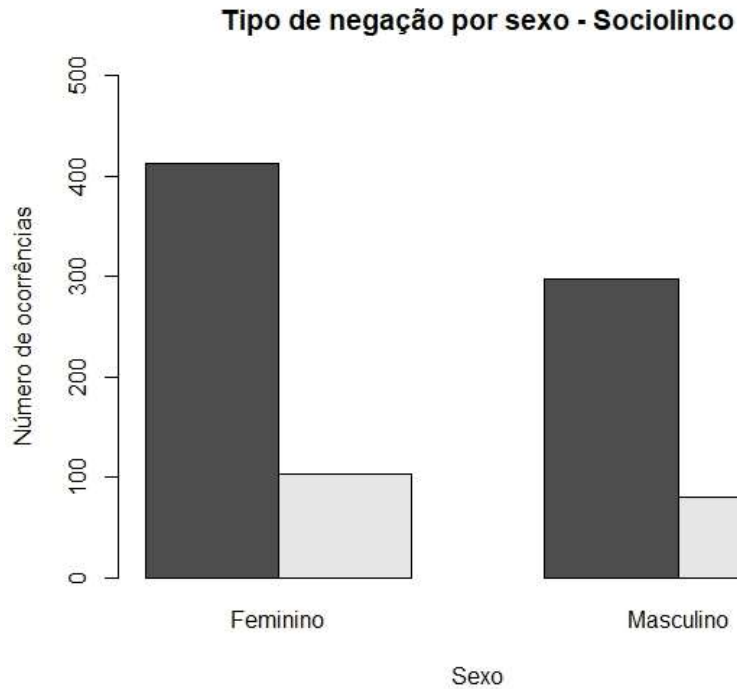
Na amostra do GEF, a maioria das ocorrências é de NEG1 e foram registradas em entrevistas com mulheres; já os homens utilizaram essa estratégia de negação 447 vezes. Da mesma forma, NEG2 foi mais frequente nas entrevistas realizadas com mulheres.

Tabela 16 – Distribuição dos dados por sexo - Sociolingo

	SOCIOLINCO			
	Feminino		Masculino	
	Frequência	Porcentagem	Frequência	Porcentagem
NEG1	413	80%	297	78.7%
NEG2	103	20%	80	21.3%

Fonte: Elaboração nossa.

O gráfico abaixo sintetiza as informações contidas na tabela acima:

Gráfico 15 – Total de ocorrências pelo sexo do informante - Sociolinco

Fonte: Elaboração nossa.

A distribuição dos dados observada para a amostra GEF se repete no do SOCIOLINCO, sendo que a maioria das ocorrências de negação foi encontrada nas entrevistas realizadas com mulheres, a proporção também é próxima. Isso é um indicador de que a variação entre as estratégias de negação não teve muita alteração ao longo dos 16 anos que separam a coleta das duas amostras.

Quanto à hipótese apresentada em 4.2.2, essa distribuição quantitativa não é suficiente para confirmá-la ou refutá-la, visto que as entrevistas realizadas com as informantes são mais longas e, portanto, contêm o maior número de ocorrências. Em 5.2.2.2, discutiremos novamente esta questão.

5.1.2.3 Faixa etária

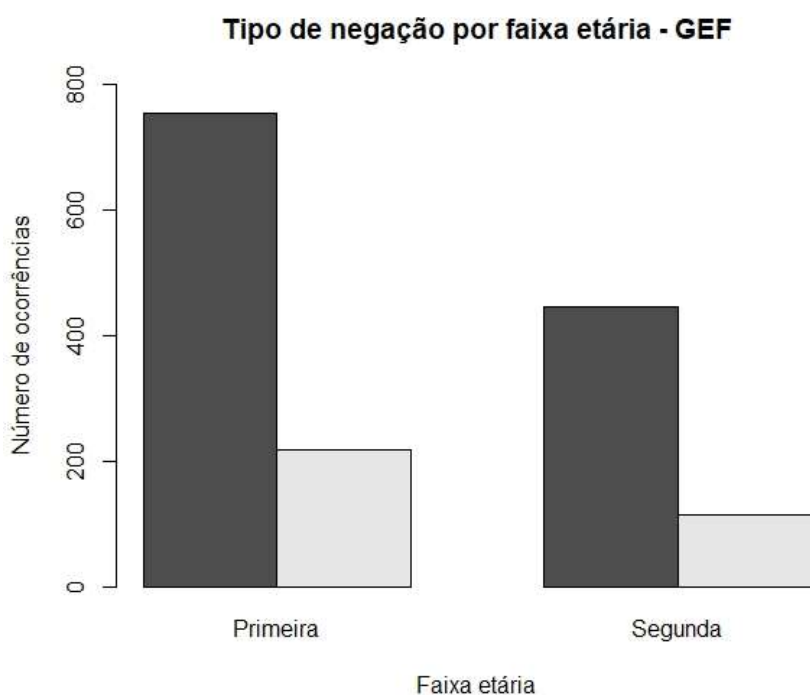
Ao abordarmos a faixa etária, deparamo-nos com a seguinte situação, a uma alteração na distribuição de dados entre uma faixa-etária e outra. Vejamos, primeiro, a distribuição dos dados dos GEF:

Tabela 17 – Distribuição dos dados por faixa etária - GEF

	GEF			
	1ª Faixa Etária		2ª Faixa Etária	
	Frequência	Porcentagem	Frequência	Porcentagem
NEG1	754	77.5%	445	79.5%
NEG2	218	22.5%	115	20.5%

Fonte: Elaboração nossa.

Essas informações podem ser mais bem visualizadas no gráfico a seguir:

Gráfico 16 – Total de ocorrências faixa etária - GEF

Fonte: Elaboração nossa.

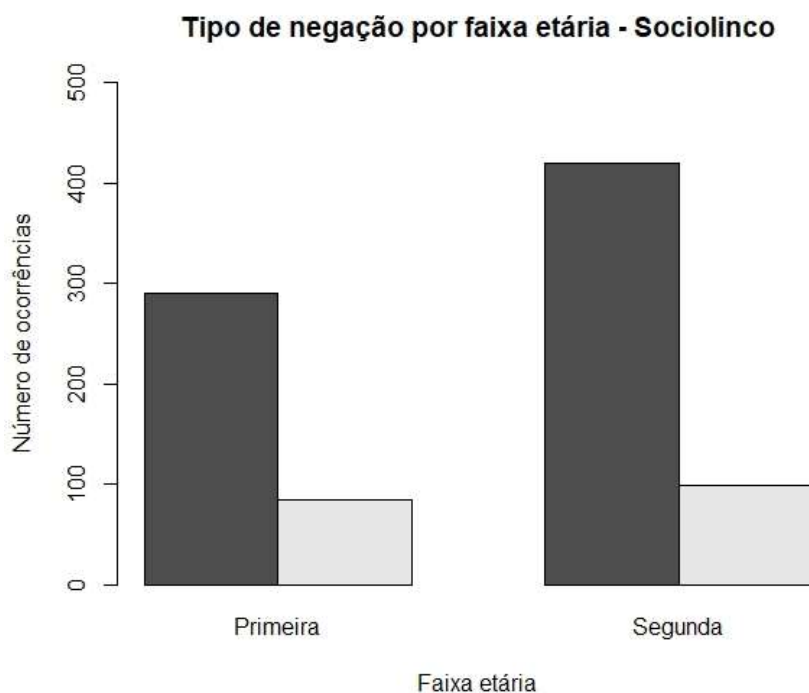
Nas entrevistas realizadas em 2003, a maioria das ocorrências de NEG1 foi encontrada nas entrevistas dos informantes de primeira faixa etária, o mesmo vale para NEG2.

Tabela 18 – Distribuição dos dados por faixa etária - Sociolingo

	SOCIOLINCO			
	1ª Faixa Etária		2ª Faixa Etária	
	Frequência	Porcentagem	Frequência	Porcentagem
NEG1	290	77.5%	420	81%
NEG2	84	22.5%	99	19%

Fonte: Elaboração nossa.

Sintetizando as informações, apresentamos o gráfico a seguir:

Gráfico 17 – Total de ocorrências faixa etária - Sociolinco

Fonte: Elaboração nossa.

Por outro lado, ao observarmos a distribuição dos dados na amostra do SOCIOLINCO, percebemos que a maioria das ocorrências, tanto de NEG1 quanto de NEG2, concentra-se na segunda faixa etária.

5.1.2.4 Informante

Na tabela a seguir, organizamos as ocorrências de acordo com os informantes. O que se evidencia, neste caso, é a maior concentração de ocorrências entre as entrevistas realizadas no ano de 2003.

Tabela 19 – Distribuição dos dados por informante

		NEG1		NEG2	
		Frequência	Porcentagem	Frequência	Porcentagem
GEF	A	181	78.6%	49	21.44%
	B	134	77.5%	39	22.5%
	C	90	67%	44	33%
	D	133	81.5%	30	18.5%
	E	112	79.5%	29	20.5%
	F	102	77.3%	30	22.7%
	G	44	81.5%	10	18.5%
	H	114	83.8%	22	16.2%
	I	123	82%	27	18%
	J	112	75.1%	37	24.9%
	K	14	70%	6	30%
	L	40	80%	10	20%
	SOCIOLINCO	M	57	67%	28
N		95	88.7%	12	11.3%
O		30	79%	8	21%
P		52	72.2%	20	27.8%
Q		72	90%	8	10%
R		107	80%	27	20%
S		6	66.6%	3	33.4%
T		44	66.6%	22	33.4%
U		58	84%	11	16%
V		27	77.1%	8	22.9%
W		97	88.1%	13	11.9%
X		65	73.8%	23	26.2%

Fonte: Elaboração nossa.

Observamos que a maioria dos informantes apresenta distribuição regular de ocorrências, ou seja, NEG2 não passa de 1/3 das ocorrências.

5.1.3 Regressão Logística

Para uma melhor compreensão dos dados e da influência das variáveis independentes sobre a variável previsora faremos alguns testes que permitiram entender melhor o que foi exposto até aqui. Segundo Oushiro (2017, s/p): “O interesse nas análises de regressão logística é 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”. Além do mais, os modelos serão a base para se calcular os pesos relativos de cada uma das variáveis, apresentados em 5.2.

O primeiro passo é a criação de um modelo com todas as variáveis – à exceção da variável “informante”, esta distorce o modelo. Na imagem abaixo é possível ver um resumo deste primeiro modelo:

Figura 9: Resumo do modelo geral

```
Call:
glm(formula = TIPO ~ SEXO + FAIXAETARIA + AMOSTRA + STATUS +
     TIPOSEQ + TIPOORA + MARCADORES + SUJ, family = binomial,
     data = dados)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.1236  -0.7380  -0.5645  -0.3438   2.6131

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)    -0.83578    0.18650  -4.481 7.42e-06 ***
SEXOMASC       -0.13544    0.10769  -1.258  0.20853
FAIXAETARIA2a -0.20879    0.10852  -1.924  0.05435 .
AMOSTRASOCIOLINCO -0.09212    0.11125  -0.828  0.40766
STATUSINF     -0.30349    0.11007  -2.757  0.00583 **
TIPOSEQAVA     0.67117    0.20672   3.247  0.00117 **
TIPOSEQDIA     0.66592    0.15822   4.209 2.57e-05 ***
TIPOSEQNAR     0.23186    0.15759   1.471  0.14121
TIPOORACOO    -0.61132    0.12688  -4.818 1.45e-06 ***
TIPOORAPRI    -1.58398    0.25581  -6.192 5.94e-10 ***
TIPOORASUB    -1.12076    0.19059  -5.881 4.09e-09 ***
MARCADORESS   -0.85634    0.18974  -4.513 6.38e-06 ***
SUJIMP         0.03660    0.10650   0.344  0.73113
SUJSEM/SUJ    -0.71158    0.29279  -2.430  0.01508 *
---
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 2510.5  on 2424  degrees of freedom
Residual deviance: 2338.7  on 2411  degrees of freedom
AIC: 2366.7

Number of Fisher Scoring iterations: 5
```

Fonte: elaboração nossa (*print* da tela).

A primeira informação que o *R* retorna é a função usada para a criação do modelo. A função que usamos é a “glm” (Modelos lineares generalizados²⁰) usada para verificar os efeitos das variáveis (OUSHIRO, 2017, s/p). Em seguida, ainda na fórmula, constam a variável resposta, “TIPO” – referindo-se, aqui, a tipo de negação – seguida do “~” – que significa “por” – e as variáveis predictoras: “STATUS + TIPOSEQ + TIPOORA + MARCADORES + SUJ + FAIXAETARIA + SEXO + AMOSTRA” – que significam, respectivamente: status da informação; tipo de sequência discursiva; tipo de oração; presença de marcadores discursivos; tipo de sujeito; faixa etária; sexo; e amostra. O termo “family = binomial” indica que a variável

²⁰ “Generalized Linear Models”.

dependente é binária – um dos motivos que nos levaram a excluir NEG3 das análises. Por fim, “data = dados” indica a tabela de dados que estamos consultando na seção do *R*.

Em seguida, o *R* apresenta os “Resíduos”, segundo Oushiro (2017, s/p), os resíduos são “a diferença entre o valor observado e o valor esperado [em uma distribuição] normal”. Temos então o mínimo, o primeiro quartil, a mediana, o terceiro quartil e o máximo. O ponto “médio” dessas medidas é a mediana, a partir dela, verificamos que há certo equilíbrio entre os quartis e o mínimo e o máximo.

Depois, observamos os coeficientes e seus respectivos valores de “estimativa”, “erro padrão” e “valor de significância”. Este resumo, que o *R* nos fornece, apresenta os coeficientes em *logodds*, ou seja, uma escala logarítmica que aponta as chances de a variação acontecer de acordo com o nível da variável dependente testada em escala logarítmica. Segundo Oushiro (2017, s/p):

[...] a escala de *logodds* vai de $-\infty$ a $+\infty$, com ponto neutro em zero. Em relação a *odds*, ela tem a vantagem de ser uma escala simétrica, com um mesmo intervalo entre o ponto neutro e suas extremidades. A operação *log* tem justamente o papel de transformar valores entre 0 e 1 em um valor negativo [...]. Desse modo, a interpretação de valores é muito mais intuitiva do que numa escala assimétrica, pois os intervalos em que há favorecimento ou desfavorecimento de um evento são diretamente comparáveis.

Assim, as análises demonstram como uma variável previsora favorece ou desfavorece a ocorrência da variável resposta tendo como ponto neutro o 0. O intercepto da análise é o ponto de partida para o entendimento dos coeficientes. Assim, considerando que NEG1 é mais comum nos dados, essa medida inicial apresenta os valores para NEG2 – ou seja, para a variação observada – no primeiro nível de cada uma das variáveis: São as chances – *logodds* – de NEG2 ocorrer se temos um falante do sexo feminino, da primeira faixa etária, do banco de dados do GEF, em um contexto em que a informação é ativada e a sequência discursiva é argumentativa, com orações absolutas, sem marcadores discursivos e o sujeito explícito.

Nesse caso, a estimativa é de -0.83578, ou seja, as chances de NEG1 ocorrer nesses contextos é maior do que as de NEG2 – lembrando que o ponto neutro para o *logodds* é 0. Para se obter os valores dos demais níveis, das demais variáveis, é necessário somar a primeira estimativa à que se quer obter. Por exemplo, se se quer obter a estimativa para as ocorrências de NEG2 em contextos em que o informante é do sexo masculino, temos de somar o valor do intercepto ao valor deste nível da

variável: $(-0.83578) + (-0.13544) = (-0.97122)$. Notamos, portanto, que um coeficiente negativo para uma determinada variável, depois da soma, apontará para um desfavorecimento da variação. Ao contrário, um coeficiente negativo apontará para algum favorecimento em relação ao intercepto, ainda que continue negativo.

Resta entender a significância. Neste estudo, utilizaremos apenas o “Pr”, ou seja, o valor P apontado pelo programa. Antes de mais nada, vale ressaltar que, quando os valores são muito baixos o R pode retornar o resultado em notação “e” (OUSHIRO, 2017, s/p). Desta forma, o valor $7.42e-06$ indica que o número que antecede o “e” deve ser multiplicado por 100 elevado ao número que sucede o “e” – um número, neste caso, negativo – -06 . É sabido que uma potência negativa resulta em um número negativo, neste caso teremos 0.0000742 . Há outros números que aparecem normalmente, ou seja, eles estão mais próximo de 1 que os que aparecem em notação “e”.

A partir dessas noções, podemos seguir com a interpretação do resumo do modelo apresentado acima. O valor p ideal deve ser ≤ 0.05 (OUSHIRO, 2017, s/p) para que possamos considerar que uma variável é significativa para o modelo. Assim, a partir dos resultados obtidos, notamos que os níveis referentes a “sexo masculino”, “2ª faixa etária”, “amostra Sociolinco”, “tipo de sequência narrativa” e “sujeito implícito” apresentam um valor p fora do intervalo ideal. Isso não indica, necessariamente que tenhamos que descartá-las, apenas que outros testes são necessários.

Por fim, o R fornece os valores dos desvios “nulo” e “residual” e o “AIC”. O desvio nulo se refere à variabilidade dos dados sem variáveis independentes e o desvio residual à mesma variabilidade com estas variáveis. O “AIC” (“Critério de informação Akaike”²¹) serve para verificar se o modelo apresenta simplicidade e qualidade para a descrição dos dados. Caso o AIC se afastasse muito dos desvios, o modelo não seria suficientemente simples para a interpretação correta dos dados.

A partir disso, constatamos que as variáveis mencionadas acima podem distorcer os dados, visto que seu valor p não permite descartar a hipótese nula. Assim, a partir deste modelo, não podemos precisar as chances de NEG2 ocorrer considerando os coeficientes obtidos. Para isso, fizemos três testes com o modelo para comparar a influência das variáveis. Segundo Oushiro (2017, s/p) a função usada para este teste é a “step” que pode ser feita de três maneiras, duas delas são: “uma

²¹ “Akaike information criterion”.

‘de baixo para cima’, que começa com um modelo sem variáveis predictoras e tenta adicioná-las uma a uma, e outra ‘de cima para baixo’, que começa com um modelo completo e tenta eliminar variáveis uma a uma”. Há ainda o teste em direção “both” por meio da qual, segundo Oushiro (2017, s/p): “o programa começa executando o mesmo que a direção “forward”, mas, toda vez que inclui uma nova variável, ele tenta excluir alguma variável que possa não mais estar contribuindo para o modelo”.

O primeiro teste realizado foi o “forward”, ou seja, o R adiciona as variáveis a fim de verificar os efeitos que cada uma delas na qualidade do modelo.

Figura 10: Teste “Forward”

```
Call: glm(formula = TIPO ~ TIPOORA + TIPOSEQ + MARCADORES + STATUS +
  SUJ + FAIXAETARIA, family = binomial, data = dados)

Coefficients:
(Intercept)      TIPOORAC00      TIPOORAPRI      TIPOORASUB      TIPOSEQAVA
   -0.93780        -0.61252        -1.57792        -1.11025         0.69947
TIPOSEQDIA      TIPOSEQNAR      MARCADORESS      STATUSINF      SUJIMP
   0.66256         0.23793         -0.84083        -0.29968         0.05211
SUJSEM/SUJ      FAIXAETARIA2a
  -0.68186        -0.21075

Degrees of Freedom: 2424 Total (i.e. Null); 2413 Residual
Null Deviance:      2510
Residual Deviance: 2341      AIC: 2365
```

Fonte: elaboração nossa (*print* da tela).

Sabendo o que este teste faz, importa apontar a fórmula que obtivemos após a rodagem: “TIPO ~ TIPOORA + TIPOSEQ + MARCADORES + STATUS + SUJ + FAIXAETARIA”. Fica evidente que as variáveis “Sexo” e “amostra” não foram incluídas no modelo final. O teste em “backwar” começa com todas as variáveis e o R vai excluindo uma a uma para testar a eficiência do modelo. Com este teste obtivemos o seguinte resultado:

Figura 11: Teste “Backward”

```
Call: glm(formula = TIPO ~ STATUS + TIPOSEQ + TIPOORA + MARCADORES +
  SUJ + FAIXAETARIA, family = binomial, data = dados)

Coefficients:
(Intercept)      STATUSINF      TIPOSEQAVA      TIPOSEQDIA      TIPOSEQNAR
   -0.93780        -0.29968         0.69947         0.66256         0.23793
TIPOORAC00      TIPOORAPRI      TIPOORASUB      MARCADORESS      SUJIMP
   -0.61252        -1.57792        -1.11025        -0.84083         0.05211
SUJSEM/SUJ      FAIXAETARIA2a
  -0.68186        -0.21075

Degrees of Freedom: 2424 Total (i.e. Null); 2413 Residual
Null Deviance:      2510
Residual Deviance: 2341      AIC: 2365
```

Fonte: elaboração nossa (*print* da tela).

Novamente, o modelo obtido exclui as variáveis “Sexo” e “amostra”. O último teste a ser realizado é o “both”, o resultado obtido por este é o que se apresenta a seguir:

Figura 12: Teste “Both”

```
Call: glm(formula = TIPO ~ TIPOORA + TIPOSEQ + MARCADORES + STATUS +
SUJ + FAIXAETARIA, family = binomial, data = dados)

Coefficients:
(Intercept)      TIPOORACOO      TIPOORAPRI      TIPOORASUB      TIPOSEQAVA
-0.93780         -0.61252         -1.57792         -1.11025         0.69947
TIPOSEQDIA      TIPOSEQNAR      MARCADORESS      STATUSINF      SUJIMP
0.66256          0.23793          -0.84083         -0.29968         0.05211
SUJSEM/SUJ      FAIXAETARIA2a
-0.68186         -0.21075

Degrees of Freedom: 2424 Total (i.e. Null); 2413 Residual
Null Deviance: 2510
Residual Deviance: 2341      AIC: 2365
```

Fonte: elaboração nossa (*print* da tela).

A fórmula para o modelo final se mantém, ou seja, excluem-se as variáveis sociais “Sexo” e “amostra” – as razões para isso serão exploradas em 5.2.2. Este modelo possui, portanto, a melhor qualidade para a análise dos dados. Abaixo apresentamos o resumo do modelo obtido depois dos testes:

Figura 13: Resumo do modelo final

```
Call:
glm(formula = TIPO ~ TIPOORA + TIPOSEQ + MARCADORES + STATUS +
SUJ + FAIXAETARIA, family = binomial, data = dados)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.0994 -0.7391 -0.5586 -0.3456  2.6007

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.93780    0.17444  -5.376 7.61e-08 ***
TIPOORACOO  -0.61252    0.12682  -4.830 1.37e-06 ***
TIPOORAPRI  -1.57792    0.25569  -6.171 6.77e-10 ***
TIPOORASUB  -1.11025    0.19022  -5.837 5.33e-09 ***
TIPOSEQAVA   0.69947    0.20489   3.414 0.00064 ***
TIPOSEQDIA   0.66256    0.15805   4.192 2.76e-05 ***
TIPOSEQNAR   0.23793    0.15723   1.513 0.13022
MARCADORESS -0.84083    0.18928  -4.442 8.90e-06 ***
STATUSINF    -0.29968    0.10998  -2.725 0.00643 **
SUJIMP        0.05211    0.10596   0.492 0.62285
SUJSEM/SUJ   -0.68186    0.29194  -2.336 0.01951 *
FAIXAETARIA2a -0.21075    0.10549  -1.998 0.04574 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 2510.5  on 2424  degrees of freedom
Residual deviance: 2341.1  on 2413  degrees of freedom
AIC: 2365.1

Number of Fisher Scoring iterations: 5
```

Fonte: elaboração nossa (*print* da tela).

Em comparação ao primeiro modelo, neste obtivemos maior significância das variáveis, apenas dois níveis de variáveis apresentam valor p alto. A partir disso, é possível imprimir uma tabela com outras medidas obtidas a partir desse modelo:

Tabela 20 – Tabela do modelo de regressão logística final

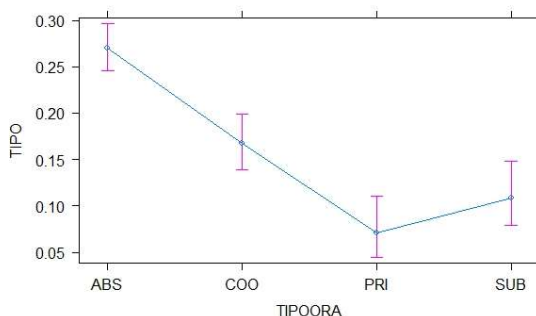
Preditores	Odds Ratio	CI	p
(Intercepto)	0.39	0.28 – 0.55	<0.001
Oração coordenada	0.54	0.42 – 0.69	<0.001
Oração principal	0.21	0.12 – 0.33	<0.001
Oração subordinada	0.33	0.22 – 0.47	<0.001
Sequência avaliativa	2.01	1.35 – 3.01	0.001
Sequência dialogal	1.94	1.43 – 2.66	<0.001
Sequência narrativa	1.27	0.94 – 1.73	0.130
Presença de marcadores	0.43	0.29 – 0.62	<0.001
Informação inferível	0.74	0.60 – 0.92	0.006
Sujeito implícito	1.05	0.86 – 1.30	0.623
Oração sem sujeito	0.51	0.28 – 0.87	0.02
2ª Faixa etária	0.81	0.66 – 1	0.046

Fonte: elaboração nossa.

Os resultados aqui apresentados usam como coeficiente os *odds*, ou seja, as chances de a variável acontecer de acordo com o nível da variável previsora testada variam de 0 a 1 se desfavorecem e de 1 a ∞ se favorecem (OUSHIRO, 2017, s/p). Na tabela acima, sabemos que o ponto de partida é NEG2, assim, observamos que os níveis de variáveis que favorecem sua ocorrência são os tipos de sequência avaliativa e dialogal, o status da informação inferível e a 2ª faixa etária – lembrando que o coeficiente do intercepto deve ser somado aos coeficientes de cada nível das variáveis. O tipo de sequência narrativa e o sujeito implícito possuem um coeficiente que favorece a variação, no entanto o valor p é maior que 0.05. Assim sendo, a sequência discursiva avaliativa é a que mais favorece a variação de formas. Esses resultados podem ser mais bem visualizados nos gráficos a seguir.

Gráfico 18 – Tipo de oração

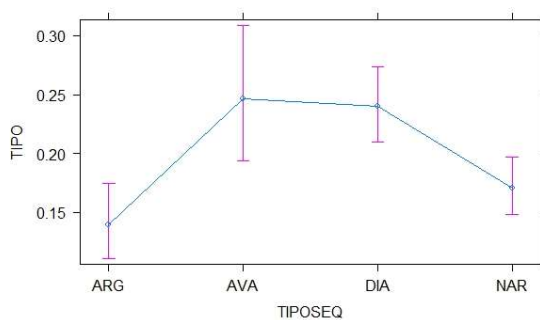
TIPOORA effect plot



Fonte: elaboração nossa.

Gráfico 19 – Tipo de sequência

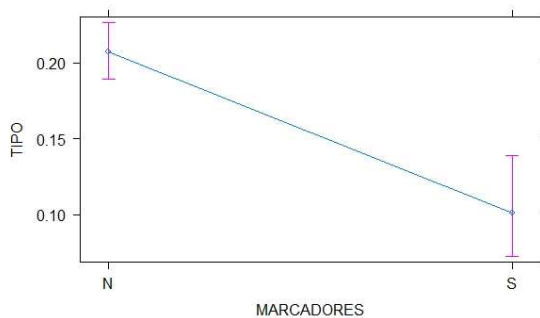
TIPOSEQ effect plot



Fonte: elaboração nossa.

Gráfico 20 – Presença ou ausência de marcadores discursivos

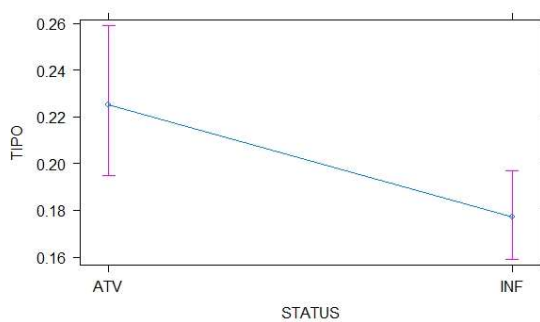
MARCADORES effect plot



Fonte: elaboração nossa.

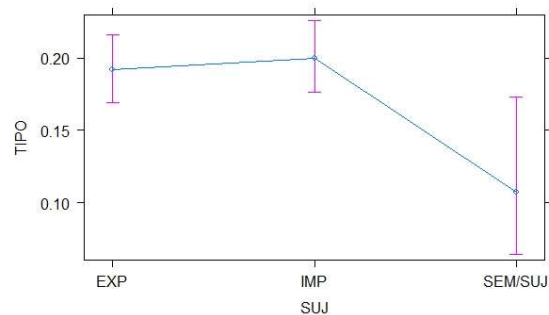
Gráfico 21 – Status discursivo da informação

STATUS effect plot



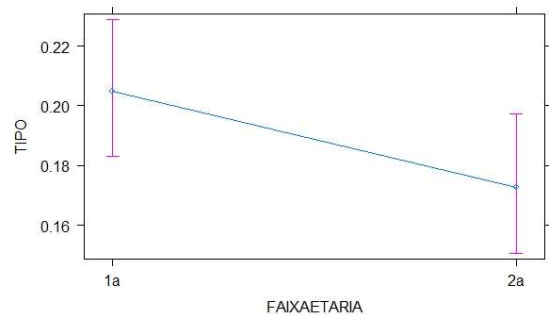
Fonte: elaboração nossa.

Gráfico 22 – Tipo de sujeito
SUJ effect plot



Fonte: elaboração nossa.

Gráfico 23 – Faixa etária
FAIXAETARIA effect plot



Fonte: elaboração nossa.

Para a interpretação desses gráficos é necessário considerar que as medidas apresentadas são de probabilidades em números proporcionais. Assim, quando se observa as medidas para a primeira faixa etária, nota-se que há pouco mais de 21% de chances de NEG2 ocorrer para primeira faixa etária e pouco menos de 18% para segunda. Ainda é necessário considerar o intervalo de confiança, expresso em rosa nos gráficos, que indicam que para a primeira faixa etária as chances estão entre 18 e 22%.

5.2 Análise qualitativa

Nesta subseção, exploraremos o efeito de cada uma das variáveis independentes na variável resposta a partir de outra medida: o peso relativo. Para a obtenção dos pesos relativos é necessária a criação de um modelo de regressão logística com efeitos mistos. Para fim de análise, neste modelo manteremos as variáveis sociais excluídas nas regressões anteriores para verificar o efeito que elas

podem ter na variável dependente. Para isso, então, criamos o modelo de regressão logística com efeitos mistos cujo resumo é apresentado a seguir:

Tabela 21 – Tabela do modelo de regressão logística com efeitos mistos

```

Generalized linear mixed model fit by maximum likelihood (Adaptive Gauss-Hermite
  Quadrature, nAGQ = 0) [glmerMod]
Family: binomial ( logit )
Formula: TIPO ~ STATUS + TIPOSEQ + TIPOORA + MARCADORES + SUJ + (1 | INFORMANTE) +
  SEXO * FAIXAETARIA + AMOSTRA
Data: dados
Control: glmerControl(optimizer = "nloptwrap")

      AIC      BIC    logLik deviance df.resid
2361.0   2453.7  -1164.5   2329.0     2409

Scaled residuals:
   Min       1Q   Median       3Q      Max
-1.1326 -0.5625 -0.4046 -0.2393  4.9345

Random effects:
 Groups      Name      Variance Std.Dev.
INFORMANTE (Intercept) 0.07693  0.2774
Number of obs: 2425, groups: INFORMANTE, 24

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)   -2.27649    0.16704  -13.628 < 2e-16 ***
STATUS1         0.15608    0.05565   2.804 0.005040 **
TIPOSEQ1       -0.38145    0.11353  -3.360 0.000779 ***
TIPOSEQ2         0.27881    0.12801   2.178 0.029405 *
TIPOSEQ3         0.28616    0.08754   3.269 0.001080 **
TIPOORA1         0.82880    0.09541   8.686 < 2e-16 ***
TIPOORA2         0.20763    0.11385   1.824 0.068188 .
TIPOORA3        -0.76758    0.19457  -3.945 7.98e-05 ***
MARCADORES1     0.42943    0.09603   4.472 7.76e-06 ***
SUJ1             0.23798    0.11071   2.150 0.031592 *
SUJ2             0.26405    0.11056   2.388 0.016924 *
SEXO1            0.03786    0.08137   0.465 0.641729
FAIXAETARIA1    0.08551    0.08214   1.041 0.297819
AMOSTRA1         0.05733    0.08225   0.697 0.485800
SEXO1:FAIXAETARIA1 0.07538    0.08217   0.917 0.358932
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Fonte: elaboração nossa (*print da tela*).

A função “glmer” (“Modelos Lineares Generalizados de Efeitos Mistos”²²) exige uma fórmula diferente, o termo (1| informante) indica que esta variável deve ser entendida como aleatória e o asterisco entre sexo e faixa etária indica que o modelo deve considerar, além do cálculo das variáveis separadamente, interação entre elas. Observamos que o valor *p* começa a ficar alto a partir da primeira variável social. Veremos com isso influenciara os resultados em seção própria – 5.2.2. A partir daí, calculamos o peso relativo das variáveis, cujo resultado apresentamos a seguir:

²² “Generalized Linear Mixed-Effects Models”.

5.2.1 Variáveis linguísticas

Nesta seção, apresentamos os resultados obtidos para as variáveis linguísticas de acordo com esta ordem: status da informação; tipo de sequência discursiva; tipo de oração; presença ou ausência de marcadores discursivos; e tipo de sujeito.

Para as variáveis linguísticas, foi possível explorar os dados mais profundamente. Assim, nossa análise se dará da seguinte forma: primeiramente analisaremos as ocorrências de forma geral (Amostra Geral), com os dados dos bancos de dados do GEF e do Sociolingo. Em seguida, analisaremos separadamente as ocorrências em cada um destes bancos de dados – “Amostra GEF” e “Amostra Sociolingo”.

5.2.1.1 *Status* da informação

Na tabela a seguir apresentamos os resultados obtidos para o status discursivo da informação negada:

Tabela 22 – NEG2 por *status* da informação

	Total	%	Peso relativo	Valor de <i>p</i>
Ativada	696	26.9	0.54	0.005
Inferível	1729	19.0	0.46	0.005

Fonte: Elaboração nossa.

A partir das informações contidas na tabela acima, podemos entender um pouco mais a respeito da distribuição dos dados. Primeiro, aprechem as ocorrências totais, em seguida, o percentual de ocorrência de NEG2. Ou seja, das 696 ocorrências de encontradas em que a informação negada foi diretamente evocada no discurso, apenas 26,9% são de NEG2, as demais são NEG1. A coluna seguinte contém os pesos relativos para cada um dos níveis da variável “Status da informação”. Assim sendo, é perceptível o fato de que em contextos nos quais a informação foi diretamente ativada favorecem as ocorrências da forma inovadora. A informação

contida na última coluna é o valor de “*p*”, considera-se que este valor deve ser inferior a 0,05 para que se possa descartar a hipótese nula²³.

Schwenter (2005) apresenta um quadro em que sintetiza seus achados ao pesquisar a influência de fatores pragmáticos sobre a negação no PB:

Tabela 23 – Negação no PB, por *status* da proposição negada

Forma	Informação nova	Inferível	Ativada
NEG1	OK	OK	OK
NEG2	#	OK	OK
NEG3	#	#	OK

Fonte: Adaptado de Schwenter (2005, p. 1452).

Nossa pesquisa confirma essa hipótese, visto que, em Goiás, encontramos apenas 2 ocorrências de NEG2 com a informação nova, o que não estatisticamente significativo. Segundo Nascimento (2014, p. 82), em Vitória (ES) o maior número de ocorrência de NEG2 se dá com a informação recuperável no contexto, não sendo necessária sua ativação direta. Todavia, o peso relativo para a ocorrência com a informação inferível é de 0.50 e quando a informação é diretamente ativa há um leve desfavorecimento, perceptível por seu peso relativo: 0.49.

Em São Paulo (SP), segundo Rocha (2013, p. 64), a informação diretamente ativada no discurso favorece a ocorrência de NEG2 na região, ao passo que a informação inferível a desfavorece. Goldnadel *et al.* (2013, p. 57) atesta o maior favorecimento de NEG2 em contextos em que a informação tenha sido ativa, seja de forma inferível ou direta. Em suma, mesmo em localidades onde se observa uma maior resistência às formas inovadoras de negação, o *status* da informação no discurso é fator determinante para a variação entre formas.

Isso pode ser observado, também na análise feita com os dados do GEF:

Tabela 24 – NEG2 por *status* da informação – Amostra GEF

	Total	%	Peso relativo	Valor de <i>p</i>
Ativada	431	27,8	0.55	0.007
Inferível	1101	19.3	0.45	0.007

Fonte: Elaboração nossa.

²³ Segundo Oushiro (2017, s/p): “[...] convencionalmente, a comunidade científica costuma usar o limite de 5% para considerar algo como muito pouco provável para acontecer ao acaso. Isso é chamado de nível α (nível alfa): o limite estabelecido pelo pesquisador para rejeitar a hipótese nula”.

Observamos que os pesos relativos se aproximam da amostra geral, mantendo o valor de p baixo. O teste de regressão logística feito com a amostra Sociolingo não selecionou esta variável como significativa para o modelo, por isso, não incluímos sua análise aqui.

5.2.1.2 Tipo de sequência discursiva

Os resultados que obtivemos para o tipo de sequência discursiva estão na tabela a seguir:

Tabela 25 – NEG2 por Tipo de Sequência discursiva

	Total	%	Peso relativo	Valor de p
Dialogais	748	27.1	0.57	0.001
Argumentativa	463	15.1	0.41	0.001
Avaliativa	228	27.2	0.57	0.029
Narrativa	986	18.4	0.45	0.032

Fonte: Elaboração nossa.

A sequência discursiva, conforme apresentado em 4.1.2, diz respeito a como o falante organiza seu discurso, de acordo com cada situação comunicativa (DI PALMA BACK *et al.*, 2004, s/p). Por se tratar de uma entrevista sociolinguística, é natural que as sequências mais comuns sejam a dialógica e a narrativa. É importante observar, no entanto, que apesar de uma maior frequência no geral, a narrativa apresenta números proporcionais de ocorrências de NEG2 inferiores aos que aparecem quando a sequência é avaliativa ou dialógica – o que por sua vez pode ser compreendido pelo fato de que em um diálogo a troca de turnos de fala favorece que o informante comente uma informação ativada no discurso do documentador.

Além disso, notamos que o software de análise estatística não selecionou as sequências procedurais, isso se deve ao fato de não haver nenhuma ocorrência de dupla negação com este tipo de sequência, ou seja, não há variação. As sequências que possuem o maior percentual de ocorrência da forma inovadora apresentam um maior peso relativo, ou seja, evidencia-se estatisticamente que elas favorecem a alternância de formas. As sequências narrativas e argumentativas apresentam um valor p baixo e, por isso, podemos entender que, em contextos em que estes tipos de sequências predominam, NEG2 é desfavorecida.

Goldnadel *et al.* (2013), ao investigar a variação das formas de negação na região sul do Brasil, não selecionou esta variável. Da mesma forma, Rocha (2013) não controla esta variável para a análise das estratégias de negação sentencial no PB paulistano. É possível fazer a comparação desta variável com os resultados obtidos por Nascimento (2014, p. 72). Neste caso, em Vitória (ES), as sequências discursivas dialogais e avaliativas favorecem a ocorrência da dupla negação, da mesma forma que em nossos dados. Porém, é preciso destacar o fato de as sequências dialogais apresentarem um peso relativo significativamente mais alto do que o encontrado em nossos dados, 0.78. Da mesma forma, as sequências argumentativa e narrativa desfavorecem as ocorrências. O trabalho de Nascimento (2014) não trata de sequências descritivas.

A regressão realizada com os dados da amostra GEF selecionou a variável “Tipo de sequência discursiva” nos testes realizados.

Tabela 26 – NEG2 por tipo de Sequência discursiva – Amostra GEF

	Total	%	Peso relativo	Valor de p
Dialogais	435	28.7	0.58	0.002
Argumentativa	260	17.7	0.44	0.101
Avaliativa	189	24.3	0.53	0.480
Narrativa	648	17.9	0.45	0.039

Fonte: Elaboração nossa.

Com os valores obtidos, podemos afirmar que as sequências dialogais favorecem a dupla negação – visto que o peso relativo é de 0.58 e o valor p é baixo – e as sequências narrativas desfavorecem – com peso relativo de 0.45. Além destes, os dados obtidos da “Amostra Sociolingo” nos permitem verificar que há similaridades com os dados abordados de forma geral:

Tabela 27 – NEG2 por tipo de Sequência discursiva – Amostra Sociolingo

	Total	%	Peso relativo	Valor de p
Dialogais	313	24.9	0.53	0.379
Argumentativa	203	11.8	0.31	0.000
Avaliativa	39	41.0	0.70	0.002
Narrativa	338	19.2	0.45	0.222

Fonte: Elaboração nossa.

O valor de p não nos permite afirmar que as sequências dialogais, de fato, favorecem as ocorrências de NEG2. Há, porém, indício suficientes para que isso seja afirmado a respeito das sequências avaliativas. É possível, também, afirmar que, na

amostra de 2019, as sequências argumentativas desfavoreceram a ocorrência da dupla negação.

5.2.1.3 Tipo de oração

Para a análise dos fatores linguísticos que influem na variação das formas de negação, selecionamos também o tipo de oração em que a negação ocorre. A tabela abaixo demonstra os resultados obtidos:

Tabela 28 – NEG2 por tipo de oração

	Total	%	Peso relativo	Valor de p
Absoluta	1256	28.4	0.70	0.000
Coordenada	616	17.0	0.55	0.068
Principal	237	7.6	0.32	0.000
Subordinada	316	11.4	0.43	0.075

Fonte: Elaboração nossa.

De início, notamos que o maior número de ocorrências se dá em orações absolutas, seguida por orações coordenadas e, depois, por períodos compostos por subordinação, sendo que a oração subordinada concentra o maior número de ocorrências entre estas – bem como maior percentual de ocorrências de NEG2. Há que se observar que o valor de p para as orações subordinadas é maior que o ideal 0.05 . Todavia, os períodos compostos por subordinação não favorecem a variação de formas, vide peso relativo observado para cada uma delas – principal: 0.32 ; subordinada: 0.43 . As orações absolutas, por sua vez, favorecem muito as ocorrências da dupla negação, visto que o peso relativo é de 0.70 . Os períodos compostos por coordenação favorecem levemente a variação da estratégia de negação.

Em contraste aos resultados por nós obtidos, podemos comparar aos demais estudos levantados na revisão de literatura. Lima e Silva (2016, p. 142) apresenta uma alternância entre as formas de negação a partir de dados coletados do *corpus* C-ORAL-BRASIL. O resultado obtido pelo pesquisador aponta para a prevalência de ocorrências de dupla negação em orações absolutas, seguida por ocorrências em períodos compostos por coordenação e subordinação. Nascimento (2014, p. 79) encontrou resultados semelhantes em Vitória (ES) aos obtidos em Goiás (GO), há diferenças no peso relativo das orações absolutas (0.55), mas ainda assim

favorecendo a variação. Nascimento (2014) aponta que os períodos compostos por subordinação favorecem a alternância de formas, diferindo dos resultados obtidos em nosso estudo. Goldnadel *et al.* (2013, p. 56) demonstra que, apesar de esse fenômeno ter menor força no sul do país, as orações absolutas concentram o maior número de ocorrências de dupla negação, seguidas das orações principais e das orações coordenadas.

Abaixo apresentamos os pesos relativos obtidos para a Amostra do GEF:

Tabela 29 – NEG2 por tipo de oração – Amostra GEF

	Total	%	Peso relativo	Valor de p
Absoluta	773	29	0.69	0.000
Coordenada	390	18.2	0.56	0.083
Principal	150	9.3	0.35	0.008
Subordinada	219	11	0.4	0.020

Fonte: Elaboração nossa.

A partir dos pesos relativos dos níveis dessa variável, percebemos que as orações absolutas favorecem a dupla negação e as orações principais e subordinadas desfavorecem. As orações coordenadas apresentam um peso relativo acima do ponto neutro, tendo, no entanto, o valor de p acima do ideal.

Abaixo os valores obtidos para esta variável a partir da Amostra Sociolinco:

Tabela 30 – NEG2 por tipo de oração – Amostra Sociolinco

	Total	%	Peso relativo	Valor de p
Absoluta	483	27.5	0.72	0.000
Coordenada	226	15.0	0.54	0.404
Principal	87	4.6	0.24	0.003
Subordinada	97	12.4	0.51	0.876

Fonte: Elaboração nossa.

As orações absolutas possuem o maior peso relativo e os casos em que a negação ocorre na oração principal de um período composto por subordinação apresentam o menor peso relativo. Porém, as orações coordenadas e as subordinadas, para esta amostra, apresentam um valor de p alto. Assim, fica evidente que a análise somente desses dados não seria suficiente para averiguar o favorecimento ou desfavorecimento de NEG2 por esta variável.

5.2.1.4 Marcador discursivo

Os marcadores discursivos interferem na escolha da estratégia de negação. Segundo Nascimento (2014, p. 88)

Além disso, os marcadores conversacionais do tipo interacional (checking): né?, entendeu?, sabe?, que são discursivos, são importantes no condicionamento da negação. Entretanto, para esse fenômeno, sua atuação pode também ser entendida como um fator de natureza sintática, uma vez que, por ocuparem a mesma posição do segundo não na oração, sua presença diminui a ocorrência de dupla negação.

Abaixo, observamos os resultados obtidos para a presença ou ausência de marcadores discursivos:

Tabela 31 – NEG2 por presença ou ausência de marcadores discursivos

	Total	%	Peso relativo	Valor de p
Sim	307	11.7	0.39	0
Não	2118	22.7	0.61	0

Fonte: Elaboração nossa.

Nascimento (2014) e Rocha (2013), conforme argumentamos em 4.1.4, apontam para o desfavorecimento dos marcadores discursivos para as variantes inovadoras – fala vitoriense e paulistana, respectivamente. Observamos que a ausência de marcadores discursivos favorece a dupla negação, com peso relativo de 0.61, ao passo que a presença deles inibe o aparecimento do segundo *NÃO*.

O mesmo ocorre caso analisemos separadamente as amostras de 2003 e de 2019. A amostra do GEF apresenta os seguintes valores:

Tabela 32 – NEG2 por presença ou ausência de marcadores discursivos – Amostra GEF

	Total	%	Peso relativo	Valor de p
Sim	211	23.4	0.38	0
Não	1321	23.4	0.62	0

Fonte: Elaboração nossa.

Notamos, inicialmente que o valor de p de ambos os níveis, como na amostra geral, é o ideal – tão baixo que o programa só retorna o número 0. Para esta amostra, a ausência de marcadores discursivos favorece um pouco mais que na Geral, ao passo que a amostra Sociolingo tem o peso relativo um pouco mais baixo, mas ainda assim indica o favorecimento:

Tabela 33 – NEG2 por presença ou ausência de marcadores discursivos – Amostra Sociolinco

	Total	%	Peso relativo	Valor de p
Sim	96	12.5	0.42	0.051
Não	797	21.5	0.58	0.051

Fonte: Elaboração nossa.

Em suma, a presença de marcadores discursivos desfavorece a ocorrência da dupla negação, ao passo que sua ausência favorece a variação entre as estratégias de negação.

5.2.1.5 Tipo de sujeito

Na tabela a seguir, apresentamos os resultados para a variável *Tipo de Sujeito*:

Tabela 34 – NEG2 por tipo de sujeito

	Total	%	Peso relativo	Valor de p
Explícito	1198	21.1	0.56	0.031
Implícito	1108	22.4	0.57	0.016
Sem sujeito	119	12.6	0.38	0.009

Fonte: Elaboração nossa.

A partir da regressão logística executada no *R*, podemos afirmar que essa variável é relevante para as ocorrências de NEG2. A realização do sujeito, seja de forma explícita – com peso relativo de 0.56 – ou implícita – com peso relativo de 0.57 –, favorece a variável inovadora. Todavia, as orações sem sujeito desfavorecem que essa estratégia seja escolhida, mas não impede totalmente. Nascimento (2014, p. 87) observa que esta variável não foi selecionada, apesar de o maior percentual de ocorrência das variáveis inovadoras ocorrerem em orações sem sujeito ou com o sujeito implícito. Este fato demonstra a diferença entre os dados de Vitória (ES) e de Goiás (GO), visto que a realização do sujeito, em nossos dados, favorece a dupla negação.

Rocha (2013, p. 85-59, grifo do autor) esclarece que, em seu estudo, a variável *sujeito* será controlada a partir de duas categorias: “*sujeito lexicalizado* ou *sujeito não lexicalizado*”. As orações com sujeito não lexicalizado compreendem os casos em que o sujeito é oculto, implícito ou inexistente. Além disso, devido ao fato de NEG2 ocorrer muito pouco na variedade paulista do PB, o autor seleciona, para sua análise, apenas as ocorrências em que a sequência discursiva é dialógica – isso

altera o percentual de ocorrências, visto que em geral a frequência era de 5,8% e, após selecionar somente este tipo de sequência, o pesquisador obtém 13,9% de dupla negação (ROCHA, 2013, p. 72). A variedade paulistana aponta neutralidade em orações em que o sujeito não é lexicalizado, com peso relativo de 0.5, e leve desfavorecimento nos casos de o sujeito ocorrer lexicalizado, com peso de 0.49.

Na região sul do Brasil (GOLDNADEL, 2013, p. 57) a dupla negação ocorre em contexto em que o sujeito é explícito, 4,8% das vezes, mas isto não é determinante, pois as ocorrências com o sujeito implícito são 4% do total e orações sem sujeito são 4,4%.

Observamos que diferentemente de outros estudos, esta variável é relevante em nossos dados, havendo favorecimento da dupla negação em orações que tenham sujeito, seja explícito ou implícito.

Ao analisar separadamente as amostras, deparamo-nos com o problema visto acima, os dados isolados não suficientes para explicar o favorecimento ou desfavorecimento de NEG2. Os pesos relativos que obtivemos para NEG2 estão dispostos na tabela a seguir:

Tabela 35 – NEG2 por tipo de sujeito – Amostra GEF

	Total	%	Peso relativo	Valor de p
Explícito	742	21.8	0.55	0.093
Implícito	705	22.6	0.55	0.098
Sem sujeito	85	14.1	0.39	0.051

Fonte: Elaboração nossa.

Vemos, então que os pesos relativos apontam para o favorecimento de NEG2 por orações que tenham sujeito e o desfavorecimento por aquelas que não o tenham. Todavia, o valor de p não nos permite aceitar essa conclusão, sendo necessária a análise das duas amostras juntas. Isso se evidencia na tabela abaixo, na qual apresentamos os valores da Amostra Sociolinco:

Tabela 36 – NEG2 por tipo de sujeito – Amostra Sociolinco

	Total	%	Peso relativo	Valor de p
Explícito	456	20	0.58	0.158
Implícito	403	22.1	0.60	0.089
Sem sujeito	34	8.8	0.33	0.087

Fonte: Elaboração nossa.

Vê-se, então, que essas variáveis analisadas em separado não apontam para uma mudança na preferência dos falantes por uma ou outra estratégia de negação sentencial. Pelo contrário, esses dados apontam para uma variação estável.

5.2.2 Variáveis sociais

Um fenômeno de variação discreto, como o que analisamos neste trabalho, pode exigir uma análise mais aprofundada. Primeiramente, fizemos a análises das variáveis linguísticas controladas – em 5.2.1. Nesta seção, analisaremos as variáveis sociais.

Um primeiro ponto a ser destacado é limitação imposta pelas amostras analisadas. Trata-se de entrevistas realizadas com um intervalo de 16 anos, em que a coleta de 2019 seguiu, rigorosamente, o perfil de informante abordado na coleta de 2003. Assim, variáveis sociais que são normalmente consideradas em estudos de sociolinguística – como escolaridade e perfil socioeconômico do informante – não são controladas aqui, visto que as amostras contam com informantes de perfil socioeconômico baixo e com baixo nível de escolarização. Além disso, não se pode apontar se alguma das variáveis é estigmatizada ou prestigiada socialmente.

Por se tratar de um fenômeno discreto, que ocorre nos níveis sintático e pragmático, não há ocorrências suficientes para que se possa determinar quais fatores sociais realmente influem na escolha do falante. As entrevistas sociolinguísticas, conforme Freitag (2009, p. 120), são realizadas a fim de neutralizar o paradoxo do observador e, como são amostras já constituídas, não foram pensadas para este fenômeno especificamente – por exemplo, para superar o paradoxo do observador, o documentador tenta conduzir o informante a sequências discursivas narrativas que, como discutimos em 5.2.1.1, desfavorecem a variação.

A este respeito, Rocha (2013, p. 66) argumenta que a variável social que apresenta maior significância para este fenômeno na fala paulistana é a escolaridade. Nesta pesquisa, o autor conta com informante que possuem o ensino médio completo ou o ensino superior, sendo que NEG2 é favorecida no primeiro caso. Goldnadel (2013, p. 59) encontra resultado semelhante em seus dados, em pesquisa realizada no sul do país. Neste caso, os informantes foram distribuídos em três níveis de escolaridade: “nível primário, nível ginásial e nível secundário” (GOLDNADEL, 2013,

p. 54). A conclusão desta pesquisa (GOLDNADEL, 2013) demonstra o favorecimento de NEG2 pelos informantes que com menor grau de instrução.

Furtado da Cunha (2001, p. 8) aponta para a mesma tendência em seus dados – apesar de não controlar outras variáveis sociais, observa um aumento das ocorrências de NEG2 e NEG3 de acordo com a diminuição do nível de escolaridade. Nascimento (2014) aponta que as variáveis sociais controladas em seu estudo da fala de Vitória (ES) não foram selecionadas na rodagem dos dados, nem mesmo a escolaridade, neste caso. Cavalcante (2007, p. 75) aponta que a dupla negação é favorecida por informantes que tenham um menor grau de instrução.

Diante disso, apresentamos a seguir a análise das variáveis sociais.

5.2.2.1 Amostra

Como evidenciamos na seção anterior, a diferença entre as estratégias de negação usadas pelos falantes vilaboenses em 2003 e em 2019 não diferem muito. A distribuição geral dos dados é bastante similar. Isso fica ainda mais claro se observarmos os valores dos pesos relativos obtidos para cada uma das amostras:

Tabela 37 – NEG2 por Amostra

	Total	%	Peso relativo	Valor de p
GEF	1532	21.7	0.51	0.476
Sociolinco	893	20.3	0.49	0.476

Fonte: Elaboração nossa.

Os valores de p para essas variáveis indicam que devemos aceitar a hipótese nula. Assim, assumimos que não houve mudança significativa nos usos da negação que pudessem ser verificados estatisticamente. Isso reforça a afirmação, que fizemos acima, de que o fenômeno em questão se trata de uma variação estável.

Vale ressaltar que, devido às características das amostras aqui utilizadas, não pudemos controlar outras variáveis sociais, como escolaridade e perfil socioeconômico. Sendo assim, podemos afirmar, a partir dos dados que temos à disposição, que se trata de uma variação estável para informantes de camadas populares e baixa escolaridade.

5.2.2.2 Sexo

Na tabela a seguir, observamos os resultados obtidos para NEG2 de acordo com o sexo do informante:

Tabela 38 – NEG2 por Sexo

	Total	%	Peso relativo	Valor de p
Feminino	1489	21.8	0.51	0.633
masculino	936	20.5	0.49	0.632

Fonte: Elaboração nossa.

À primeira vista, notamos que o peso relativo indicaria que essa variável não favorece ou desfavorece a variação. O valor de p evidencia isso, visto que confirma a hipótese nula.

Esse resultado, então, confirma nossa hipótese, de que não há diferença significativa para o sexo no que diz respeito à escolha da estratégia de negação. Ou seja, a depender dos fatores pragmáticos e linguísticos, discutido acima, a variação pode ocorrer igualmente, independente do sexo do falante.

Goldanadel (2013) observou a mesma tendência em seu estudo a respeito do uso das estratégias de negação na região sul do país, em que NEG2 ocorreu em 3.5% das vezes quando o informante era do sexo masculino e 4.8% para informantes do sexo feminino. Cavalcante (2007), por sua vez, investigou as estratégias de negação utilizadas em comunidades de fala de “antigos quilombos ou comunidades de escravos” (CAVALCANTE, 2007, p. 19) e seus dados também corroboram o que se observa em Goiás-GO: a variável “sexo” não é selecionada pelo software de análise estatística como determinante para a prevalência de uma ou outra variante. No estudo de Rocha (2013), assim como no de Cavalcante (2007), esta variável não foi selecionada como significativa.

Nascimento (2014), entretanto, observa que os homens favorecem a negação simples anteposta ao verbo – com peso relativo de 0.53 para homens e 0.47 para mulheres –, ao passo que as mulheres preferem a dupla negação – com peso relativo de 0.53 para mulheres e 0.47 para homens. Em todos os estudos observados, apenas este último apresentou esta variável como estatisticamente significativa. Todavia, a autora argumenta que o fato de esta ser a única variável social selecionada não é possível aprofundar a discussão.

5.2.2.3 Faixa etária

Os bancos de dados aos quais recorremos²⁴ controlaram a faixa etária apenas como primeira e segunda, não sendo possível analisar esta variável em três níveis, como se faz comumente. É, porém, ainda uma amostra randômica da mesma comunidade de fala, em períodos distintos, sendo, portanto, um estudo de tendências em tempo real. Na tabela a seguir, apresentamos os resultados que o *R Core Team* retornou para a variável faixa etária:

Tabela 39 – NEG2 por Faixa etária

	Total	%	Peso relativo	Valor de <i>p</i>
1 ^a	1346	22.4	0.52	0.292
2 ^a	1079	19.8	0.48	0.292

Fonte: Elaboração nossa.

Poderíamos dizer que os mais jovens favorecem o uso da forma duplamente marcada de negação. Observamos que para a primeira faixa etária, em um total de 1346 ocorrências, apenas 22.4% são de NEG2, números superiores aos da segunda faixa etária que soma 19.8% de dupla negação para um total de 1079 ocorrências, essa diferença percentual se reflete no peso relativo obtido, que indica que falantes mais novos preferem a variável inovadora, mas ambos os pesos relativos estão muito próximos do ponto neutro. No entanto, como em todas as variáveis sociais aqui exploradas, o valor de *p* é bastante alto, o que nos faz descartar a hipótese de que haja influência desta variável para o fenômeno em questão.

Da mesma forma, outras pesquisas também apontam para essa estabilização. Cavalcante (2007, p. 70) afirma que os dados não apontam para uma mudança em progresso. Rocha (2013, p. 66) aponta que esta variável, tal qual em nosso estudo, não demonstra uma tendência de substituição da negação pré-verbal pela dupla negação ou pela negação posposta ao verbo. Nascimento (2014, p. 86) também se deparou com um leve favorecimento de NEG2 por falantes mais jovens, mas seus resultados para esta variável também não foram estatisticamente significantes.

Se não há influência da faixa etária, podemos afirmar que a variação das estratégias de negação na fala vilaboense é um fenômeno de variação estável. Essa estabilidade, nos termos de Freitag (2005, p. 108), é observada na fala dos indivíduos

²⁴ Não foi possível coletar dados, estratificados em três faixas etárias, como seria o esperado em um estudo aos moldes labovianos, devido ao contexto da pandemia ocasionada pela proliferação do vírus SARS-CoV-2 (o novo coronavírus).

entrevistados – cujo nível socioeconômico é baixo e a escolarização também. Não foram encontrados indícios de gradação etária e a mudança não pode ser constatada.

CONSIDERAÇÕES FINAIS

Um dos pilares da sociolinguística é a ideia de que a língua é um sistema heterogêneo, havendo a possibilidade da ocorrência de fenômenos de variação e mudança linguística. A variação é aqui entendida como a possibilidade de se dizer a mesma coisa de formas alternativas, ou seja, é possível se referir ao mesmo estado de coisas usando formas linguísticas diferentes, mantendo o valor de verdade.

A alternância entre as formas de negação sentencial – pré-verbal (NÃO+SV); dupla negação (NÃO+SV+NÃO); e negação pós-verbal (SV+NÃO) – é um fenômeno de variação constatado em várias regiões do Brasil, conforme apontam os trabalhos de: Avelar; Silva; Almeida (2013); Cavalcante (2007); Cunha (2001); Nascimento (2014); Rocha (2013); e Sousa (2012).

A partir da revisão de literatura realizada, atestamos que a dupla negação e a negação sentencial perdem força na região Sul do país, mas, ainda assim, ocorrem, e são mais frequentes nas regiões Sudeste e Nordeste. Em nosso estudo, notamos que essa alternância de formas ocorre também na Cidade de Goiás (GO). No *locus* em questão, a estratégia de negação pré-verbal é mais frequente que a dupla negação e a negação pós-verbal tem poucas ocorrências, o que justificou que estas fossem descartadas para a análise dos dados. Os objetivos desta pesquisa foram alcançados, quais sejam: descrever a variação das estratégias de negação na cidade de Goiás-GO; verificar se o fenômeno em questão se trata de uma variação estável ou de mudança linguística; averiguar quais são as variáveis condicionantes para a variação de formas.

De forma geral, apesar de não ser possível comprovar isso por falta de dados, as formas alternantes de negação no PB parecem ter influência do contato com as línguas de troco banto – discutido em 1.6, visto que elas apresentam a dupla negação sentencial. Outro fator que se observa como determinante para a essa variação é o desgaste fonológico sofrido pelo NÃO pré-verbal, passando a ser apenas o NUM. Infelizmente, não foi possível controlar essa variável para que verificássemos se ela é significativa para a escolha dos falantes, pois seria preciso desenvolver uma investigação fonético-fonológica, que foge ao escopo da presente pesquisa. Esse desgaste, como discutimos em 2.3.2, ocorre por razões pragmáticas, economia e informatividade.

Encerramos este trabalho com a síntese dos resultados obtidos e discussões propostas. Nas amostras levantadas para a análise, encontramos 1909 ocorrências para NEG1 e 516 ocorrências de NEG2, sendo, então, 78.7% de ocorrências se deram com o *NÃO* anteposto ao verbo e 21.3% para a dupla negação.

Esse tipo de alternância de formas, como argumentamos em 1.6, tem origem no contato entre as línguas bantas e o português. Além disso, em 2.3.2, discutimos a influência da informatividade e a economia para esta variável.

Na seção 5, apresentamos a análise de dados. Com isso, verificamos que as chances de NEG2 ocorrer são maiores na amostra coletada em 2003, todavia, proporcionalmente, não obtivemos resultados estatísticos que apontem para uma diferença significativa entre as amostras. Este fato e os resultados obtidos para a variável faixa etária – não diferença significativa nas escolhas dos informantes das duas faixas etárias – apontam que o fenômeno em questão é, de fato, uma variação estável.

A variável sexo aponta para este fato, também. Conforme discutimos em 4.2.2, Scherre e Yacovenco (2011) mostram que estudos realizados demonstram que informantes do sexo feminino usam mais livremente variáveis inovadoras que estão abaixo do nível da consciência, que não sofram estigma social, ao passo que rejeitam formas que sofram algum tipo de estigma. O fato de não encontrarmos diferenças estatísticas significativas entre os usos de ambos os sexos apontam que, essa variação estável ocorre abaixo do nível da consciência.

No que se refere às variáveis linguísticas, verificamos que quando a informação é diretamente ativada no discurso as chances de que NEG2 ocorra aumentam. Em se tratando do tipo de sequência discursiva, a avaliativa e a dialogal aumentam as chances da variante inovadora.

Da mesma forma, orações absolutas são mais propícias para NEG2 se comparadas às orações principais, subordinadas e coordenadas. A ausência de marcadores discursivos é, também, fator determinante para que o falante escolha NEG2 em detrimento de NEG1. Por fim, o tipo de sujeito apresenta pouca diferença entre os níveis explícito e implícito, em consonância a orações sem sujeito, que apresentam a menor chance de realização da dupla negação para esta variável.

Para que se possa avaliar melhor esse fenômeno de variação, seria necessário que se fizessem novas entrevistas, abrangendo pessoas de outros extratos sociais, com diferentes níveis de escolaridade, visto que essas variáveis

sociais foram determinantes em outras pesquisas – Rocha (2013), Goldnadel (2013), Furtado da Cunha (2001) e Cavalcante (2007). O controle das ocorrências em que o *NÃO* pré-verbal for trocado pelo *NUM* também parece determinante para este estudo. Devido à limitação das amostras que utilizamos não pudemos explorar essas possibilidades.

Como discutimos em 5.2.1, as variáveis linguísticas influem na escolha do falante e, além disso, acrescentamos o fato discutido em 2.3.2, esse fenômeno ocorre devido a forças pragmático-discursivas envolvidas no momento da comunicação. Além disso, como nossas análises estatísticas revelam, as estratégias de negação na fala vilaboense caracterizam-se como variação estável, abaixo do nível da consciência.

REFERÊNCIAS

- AMARAL, A. **O dialeto caipira**. 3. ed. São Paulo: Hucitec-SCET-CEC, 1976. [1920].
- AVELAR, L. L. M. R. N.; SILVA, M. R.; ALMEIDA, T. P. As formas de negação com o item não no português falado em Santa Luzia: um estudo preliminar. *In*: AMARAL, Eduardo Tadeu Roque (Org.). **O português falado em Minas Gerais**. Belo Horizonte: Faculdade de Letras da UFMG, 2013. (Viva Voz). p. 27-36.
- BAGNO, M. **Dicionário crítico de sociolinguística**. 1. Ed. São Paulo: Parábola Editorial, 2017.
- BAGNO, M. O impacto das línguas bantas na formação do português brasileiro. **Cadernos de Literatura em Tradução**, n. 16, p. 19-31, 2016.
- BORGES, F. C. S. F. **A cidade de Goiás e o Turismo**: um estudo do patrimônio histórico e cultural e sua influência na organização da atividade turística no município. 2010. Dissertação (Mestrado em Turismo e Hotelaria) – Programa de Mestrado Stricto Sensu em Turismo e Hotelaria, Centro de Educação de Balneário Camboriú, Universidade do Vale do Itajaí, Balneário Camboriú-SC, 2010.
- BORTONI-RICARDO, S. M. **Português brasileiro**: a língua que falamos. São Paulo: Contexto, 2021.
- BORTONI-RICARDO, S. M. **Manual de Sociolinguística**. São Paulo: Contexto, 2014.
- BYBEE, J. L. **Língua, uso e cognição**. Tradução Maria Angélica Furtado da Cunha. São Paulo: Cortez, 2016.
- CASTILHO, C. M. M. de. **Processo de redobrimento sintático no português medieval**: a formação das perífrases com estar. 2005. Tese (Doutorado em Linguística) – Programa de Pós-Graduação em Linguística, Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2005.
- CASTILHO, A. T. **Nova gramática do português brasileiro**. 1. ed. São Paulo: Contexto, 2012.
- CASTILHO, A. A.; PRETI, D. (org.). **A linguagem falada culta na cidade de São Paulo**. São Paulo: FAPESP, 1987.
- CAVALCANTE, R. **A negação pós-verbal no português brasileiro: Análise descritiva e teórica de dialetos rurais afro-descendentes**. 2007. Dissertação

(Mestrado em Letras e Linguística). Instituto de Letras, Universidade Federal da Bahia. Salvador, 2007.

COELHO, I. L. *et al.* **Para conhecer Sociolinguística**. São Paulo: Contexto, 2015.

DA COSTA SOUZA, P. **A dupla negação pré-verbal no catalão e no português brasileiro**: história, variação e uso. 2017. Tese (Doutorado em Filologia e Língua Portuguesa) – Departamento de Letras Clássicas e Vernáculas da Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2017.

DI PALMA BACK, A. C. *et al.* Classificação das sequências discursivas em entrevistas sociolinguísticas. *In*: Encontro Celsul - Círculo de Estudos Lingüísticos do Sul, 6., 2004, Florianópolis. **Anais...** Universidade Federal de Santa Catarina.

FARACO, C. A. **Linguística histórica**: uma introdução ao estudo da história das línguas. São Paulo: Parábola Editorial, 2005.

FREITAG, R. M. K. **Documentação sociolinguística**: coleta de dados e ética em pesquisa. São Cristóvão: Editora UFS, 2017. 82 p.

FREITAG, R. M. K. Idade: uma variável sociolinguística complexa. **Línguas & Letras**, v. 6, p. 105-12, 2005.

FURTADO DA CUNHA, M. A. O Modelo das Motivações Competidoras no Domínio Funcional da Negação. **Revista D.E.L.T.A**, ano 17, n. 1, 2001.

GOLDNADEL, M. *et al.* Estratégias alternativas de negação sentencial na região sul do Brasil: análise da influência de fatores pragmáticos a partir de dados do projeto VARSUL. **Revista de Estudos da Linguagem**, v. 21, n. 2, p. 35-74, 2013.

GUY, G. R.; ZILLES, A. **Sociolinguística quantitativa**: instrumental de análise. São Paulo: Parábola, 2007.

HALL, S. **A identidade cultural na pós-modernidade**. 11. ed. Rio de Janeiro: DP&A, 2006.

HOUAISS, A.; VILLAR, M. S. **Dicionário Houaiss de Língua Portuguesa**. Rio de Janeiro: Objetiva, 2009.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). Disponível em: <https://cidades.ibge.gov.br/brasil/go/goias/panorama>. Acesso em 25 de abr. de 2021.

LABOV, W. Where does the linguistic variable stop? A response to Beatriz Lavandera. **Working papers in sociolinguistics**, Washington-DC, n. 44, 1978. Disponível em: <https://files.eric.ed.gov/fulltext/ED157378.pdf>. Acesso em: 12 fev. 2021.

LABOV, W. **Padrões sociolinguísticos**. Tradução de Marcos Bagno, Maria Mata Scherre, Caroline R. Cardoso. São Paulo: Parábola Editorial, 2008.

LAVANDERA, B. R. Los limites de la variable sociolingüística. In: LAVANDERA, B. R. **Variación y significado**. Buenos Aires: Libreria Hachette, 1984.

LIMA E SILVA, L. F. **Negação verbal no Português Brasileiro**: Aspectos teórico-metodológicos em estudo baseado em *corpus*. Dissertação (Mestrado em Estudos Linguísticos) – Programa de Pós-Graduação em Estudos Linguísticos, Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2016.

LIPSKI, J. Angola e Brasil: vínculos linguísticos afro-lusitanos. **Veredas**: Revista da Associação Internacional de Lusitanistas, n. 9, p. 83-98, 1 mai. 2008. Disponível em: https://digitalis-dsp.sib.uc.pt/jspui/bitstream/10316.2/34451/1/Veredas9_artigo6.pdf. Acesso em: 26 de jan. de 2020.

LUCCHESI, D. A periodização da história sociolingüística do Brasil. **D.E.L.T.A.**, 33.2, 2017. p. 347-382.

MACHADO, A. L. G. Relações sociais como fatores decisivos no uso de pronomes de tratamento de 2ª pessoa. In: VI Simpósio nacional Estado e poder: cultura, 2010, São Cristóvão: UFF. **Anais...** 2010. Disponível em: <https://www.historia.uff.br/estadoepoder/6snepc/GT8/GT8-ANA.pdf>. Acesso: 28 jun. 2021.

MENDONÇA, R. **A influência africana no português do Brasil**. Brasília: FUNAG, 2012.

NASCIMENTO, C. A. R. **A negação no português falado em Vitória/ES**. Dissertação (Mestrado em Linguística) - Centro de Ciências Humanas e Naturais, Universidade Federal do Espírito Santo, Vitória, 2014.

NEVES, M. H. M. **Gramática Funcional**: interação, discurso e texto. São Paulo: Contexto, 2018.

NEVES, M. H. M. **Gramática de usos do português**. 2. Ed. São Paulo: Editora Unesp, 2011.

OUSHIRO, L. **Introdução à Estatística para Linguistas**. Zenodo, 2017. Disponível em <http://doi.org/10.5281/zenodo.822070>. Acesso em 20 de mai. de 2021.

OUSHIRO, L. **Identidade na pluralidade**: avaliação, produção e percepção linguística na cidade de São Paulo. Tese (Doutorado em Letras) - Identidade na pluralidade: avaliação, produção e percepção linguística na cidade de São Paulo, Departamento de Linguística, Universidade de São Paulo. São Paulo: FFLCH/USP, 2015.

PAREDES SILVA, V. L. Forma e função nos gêneros de discurso. **Alfa revista de Linguística**, São Paulo, 1997, v. 41, p. 79-98.

PRUDENTE, T. C. A. **Cotidiano e preservação**: asilo São Vicente de Paulo da Cidade de Goiás. Trabalho de Conclusão de Curso (Mestrado Profissional em Gestão do Patrimônio Cultural) – Instituto Goiano de Pré-História e Antropologia, Universidade Católica de Goiás, Goiânia, 2006.

R CORE TEAM. **R**: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2013. Disponível em: <https://www.r-project.org/>. Acesso em 15 de dez. de 2020.

RIBEIRO, D. **O povo brasileiro**: a formação e o sentido do Brasil. São Paulo: Companhia das Letras, 1995.

RIBEIRO, L. C. de S. **Variação pronominal de primeira pessoa do plural**: *nós* e *a gente* na cidade de Goiás. 2020. Tese (Mestrado em Língua e Interculturalidade), Programa de Pós-Graduação Stricto Sensu em Língua, Literatura e Interculturalidade, Campus Cora Coralina, Universidade Estadual de Goiás, 2020.

ROCHA, R. S. **A negação dupla do português paulistano**. 2013. 97 f. Dissertação (Mestrado em Linguística)- Programa de Pós-Graduação em Semiótica e Linguística Geral, Universidade de São Paulo, São Paulo, 2013. Disponível em: www.teses.usp.br/teses/disponiveis/8/8139/tde.../2013_RafaelStoppaRocha_VCorr.pdf. Acesso em: 15 fev. 2021.

SAUSSURE, F. de. **Curso de Linguística Geral**. Tradução Antônio Chelini, José Paulo Paes, Isidoro Blikstein. 27. ed. São Paulo: Cultrix, 2006.

SCHERRE, M. M. P.; YACOVENCO, L. C. A variação linguística e o papel dos fatores sociais: os gêneros dos falantes em foco. **Revista da ABRALIN**, v. 10, n. 3, p. 121-146, 2011.

SCHWENTER, S. A. The pragmatics of negation in Brazilian Portuguese. **ScienceDirect**: *Lingua*, v. 115, n. 10, p. 1427-1456, out. 2005. Disponível em:

<https://www.sciencedirect.com/science/article/pii/S0024384104000889>. Acesso em 20 de mai. de 2021.

SCHWENTER, S. A. Fine-Tuning Jespersen Cycle. *In*: BIRNER, B. J.; WARD, G. W. (Orgs.). **Drawing the boundaries of meaning**. Neo-Gricean studies in honour of Laurence R. Horn. Amsterdã e Filadélfia: Benjamins. 327-344, 2006.

SILVA, L. A. **Os usos “até” na língua falada na cidade de Goiás**: funcionalidade e gramaticalização. Goiânia, 2005. 187 p. Dissertação (Mestrado em Letras) – Faculdade de Ciências e Letras da Universidade Federal de Goiás.

SOUSA, L. T. **Sintaxe e interpretação de sentenças negativas no português brasileiro**. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas-SP, 2012.

TARALLO, F. **A Pesquisa Sociolinguística**. São Paulo: Ática, 1986.

WEINREICH, U.; LABOV, W.; HERZOG, M. I. **Fundamentos empíricos para uma teoria da mudança linguística**. Tradução de Marcos Bagno. São Paulo: Parábola Editorial, 2006.

ANEXO A – ROTEIRO DE ENTREVISTA SEMIESTRUTURADA

1. Você tem quantos anos?
2. Você mora em Goiás desde quando?
3. Em que outro lugar que você morou?
4. Você nasceu de dia ou de noite?
5. Na fazenda ou na cidade?
6. Como que foi o dia do seu nascimento? As pessoas te contaram? Como foi o parto naquela época?
7. Quem colocou esse nome em você?
8. Tem algum motivo especial por que você tem esse nome?
9. E quando você era pequeno/a que brincadeira vocês faziam, com quem você brincava?
10. E nessas brincadeiras teve algum fato especial?
11. Então você brincou muito na infância?
12. Vocês ajudavam os pais também?
13. E uma arte que você fez quando era pequeno/a?
14. Você tem irmãos? Quantos são?
15. E os pais batiam muito?
16. Todo mundo (os irmãos) apanhava?
17. E nessas brincadeiras suas, você já se machucou alguma vez? Já quebrou braços, pernas?
18. Tem mais história da infância que você queira contar?
19. E como eram as brigas da infância? Brigava mais com quem?
20. E você estudou? Onde e como foi?
21. Seus professores, você se lembra?
22. E teve um colega com quem você fez mais amizade na escola ou não?
23. Você trabalha?
24. E você tem parentes aqui em Goiás ou na fazenda?
25. Sua juventude, como é que foi?
26. E os seus amigos?
27. E você namorou muito?
28. Casou? Como foi?
29. Você acha que os relacionamentos de hoje são muito diferentes?
30. Você teve um momento, porque todos nós temos um momento assim, que a gente fica meio com medo e acha que sente um frio, algum perigo, né? Você já passou por alguma situação de medo, perigo?
31. Você tem medo de alguma coisa ou não?
32. Você participou de muitas festas ou participa?
33. Que tipo de festa você frequentou/frequenta?
34. Você sabe dançar? Gosta de dançar?
35. O que você gosta mais de fazer que te deixa feliz...
36. E você já pensou no futuro?
37. E o convívio assim com seus irmãos?

38. Tem um/a com quem você dá mais certo?
39. Eles estudaram também?
40. E quanto à religião?
41. Sobre a cidade de Goiás, você gosta daqui? Você acha que a cidade mudou muito?
42. Você viu todo esse crescimento aí da cidade?
43. Você tem muito contato com político ou não?
44. E viagem, você fez muitas viagens?
45. Você acha que Goiás está meio descuidada?
46. Goiânia, você já foi ou vai bastante? Sabe andar por lá?