



**UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA
DEPARTAMENTO DE LETRAS E ARTES
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGUÍSTICOS**

MARCELO DOS SANTOS

**O APAGAMENTO DE /d/ EM MORFEMA DE GERÚNDIO NA FALA DE FEIRA DE
SANTANA-BA**

Feira de Santana-BA
2021

MARCELO DOS SANTOS

O APAGAMENTO DE /d/ EM MORFEMA DE GERÚNDIO NA FALA DE FEIRA DE SANTANA-BA

Dissertação apresentada ao Programa de Pós-Graduação em Estudos Linguísticos da Universidade Estadual de Feira de Santana como requisito para obtenção do título de Mestre em Estudos Linguísticos.

Orientadora: Profa. Dra. Josane Moreira de Oliveira

Coorientadora: Profa. Dra. Juliana Escalier Ludwig Gayer.

S233a Santos, Marcelo dos
O apagamento de /d/ em morfema de gerúndio na fala de Feira de Santana - Ba / Marcelo dos Santos. –, 2021.
117 f.: il.

Orientadora: Josane Moreira de Oliveira
Coorientadora: Juliana Escalier Ludwig Gayer
Dissertação (mestrado) – Universidade Estadual de Feira de Santana, Programa de Pós-Graduação em Estudos Linguísticos, 2021.

1. Sociolinguística - Feira de Santana - BA 2. Gerúndio. 3. Linguística - Variação e mudança - Feira de Santana - BA I. Oliveira, Josane Moreira de, orient. II. Gayer, Juliana Escalier Ludwig, coorient. III. Universidade Estadual de Feira de Santana. IV. Título.

CDU: 80:30 (814.22)

Ficha Catalográfica – Biblioteca Central Julieta Carteadó

Tatiane Souza Santos - Bibliotecária CRB5/1634

**TERMO DE APROVAÇÃO
DISSERTAÇÃO DE MESTRADO**

**O APAGAMENTO DE /d/ EM MORFEMA DE GERÚNDIO NA FALA DE FEIRA DE
SANTANA-BA
MARCELO DOS SANTOS**

Dissertação/Tese submetida ao Programa de Pós-Graduação em Estudos Linguísticos da Universidade Estadual de Feira de Santana, área de concentração Linguagem e Sociedade, Linha de Pesquisa Variação e Mudança Linguística no Português, como requisito para obtenção do título de Mestre em Estudos Linguísticos.

Aprovada em 31 de março de 2021

BANCA EXAMINADORA:



Profa. Dra. Josane Moreira de Oliveira
Universidade Estadual de Feira de Santana
Orientadora



Profa. Dra. Juliana Escalier Ludwig Gayer
Universidade Federal da Bahia
Coorientadora



Profa. Dra. Norma Lúcia Fernandes de Almeida
Universidade Estadual de Feira de Santana
Examinador Interno



Prof. Dr. Gredson dos Santos
Universidade Federal da Bahia
Examinador Externo

Às duas mulheres que partilharam comigo todos os momentos da escrita deste texto, os momentos de risos e felicidades e também aqueles de dificuldades e tristezas, sem, jamais, negligenciar o olhar atento e carinhoso, o sorriso, o beijo, o abraço: **Neide Pinto dos Santos**, minha esposa, e **Leticia Pinto dos Santos**, minha filha.

AGRADECIMENTOS

A Deus, expressão de amor supremo, pelo dom da vida, pela presença constante e pela graça que me concede em completar mais esta etapa de formação.

À UEFS, esse local transformador de vidas, local de desafios e de conquistas, por cumprir seu papel formador na missão de oferecer educação pública e de qualidade socialmente referendada.

À Coordenação do Programa de Pós-Graduação em Estudos Linguísticos, pela administração cuidadosa, pela eficiência, efetividade e presteza de sempre no trato com o serviço público.

À professora Dra. Josane Moreira de Oliveira, minha orientadora, por ter me apresentado o mundo fascinante da Sociolinguística e por ter me conduzido pelo caminho da pesquisa em Linguística, pela paciência e pela disponibilidade.

À Professora Dra. Juliana Ludwig Escalier Gayer, minha coorientadora, por ter aceitado de pronto a proposta de coorientação, pelas leituras sempre atentas e contribuições tão precisas.

À Professora Dra. Silvana Silva de Faria Araújo, que me recebeu de maneira muito gentil e acolhedora no meu primeiro contato com o PPGEL.

Aos professores com quem cursei disciplinas, pela excelência no exercício do magistério e pelas partilhas, em especial: Patrício Nunes Barreiros, Josane Moreira de Oliveira, Liliane Lemos Santana Barreiros, Norma Lúcia Fernandes de Almeida, Rita de Cássia Ribeiro de Queiroz e Zenaide de Oliveira Novais Carneiro. E aos demais docentes do Programa, que trabalham para que o PPGEL seja referendado social e cientificamente.

Novamente à Professora Dra. Norma Lúcia Fernandes de Almeida, pela leitura atenta do texto para a qualificação e pelas palavras de incentivo.

Ao Professor Dr. Gredson dos Santos, pela leitura atenta do texto para a qualificação, pelas contribuições tão precisas, pertinentes e enriquecedoras, e pelas indicações de leitura que em muito contribuíram para a minha pesquisa.

Aos colegas doutorandos da Turma 1 e mestrandos da Turma 9 do PPGEL. Em especial a Claudice F. Santos, Dayane M. Lemos, Jan Carlos D. de Santana, Josimar S. Silva, Marinez Aparecida A. S. Trindade e Simone P. de Jesus, pelo acolhimento, pelo apoio mútuo e pela amizade. Vocês fizeram a jornada mais leve!

RESUMO

Esta dissertação apresenta os resultados de uma pesquisa sobre o apagamento da oclusiva /d/ no morfema *-ndo*, indicador de gerúndio, no português brasileiro. Assume-se que o fenômeno pesquisado é uma regra de natureza variável em que a variável dependente binária se caracteriza pela manutenção ou apagamento de /d/ no morfema de gerúndio como em *cantando* (forma padrão) e *cantano* (forma não-padrão). O referencial teórico-metodológico é o da Sociolinguística Variacionista (WEINREICH; LABOV; HERZOG, 2006 [1968]; LABOV, 2008 [1972]). O estudo foi realizado a partir de 24 entrevistas sociolinguísticas da cidade de Feira de Santana-BA, oriundas do banco de dados do projeto “A Língua Portuguesa no Semiárido Baiano” (UEFS). O objetivo da pesquisa foi descrever como o fenômeno está condicionado de acordo com fatores estruturais e sociais, bem como observar, em tempo aparente, se há indícios de mudança linguística em curso na comunidade pesquisada. Foram controladas as variáveis sociais ‘escolaridade’, ‘sexo’ e ‘faixa etária’, e as variáveis linguísticas ‘conjugação verbal’, ‘extensão do verbo’, ‘contexto seguinte’ e ‘estrutura sintática’. Os dados foram analisados quantitativamente por meio da ferramenta computacional GoldVarb X (SANKOFF; TAGLIAMONTE; SMITH, 2005). Os resultados indicam taxa de uso de 67,7% de aplicação da variante não-padrão no *corpus*. O fenômeno variável é condicionado por fatores linguísticos e extralinguísticos, sendo que foram estatisticamente relevantes, nesta ordem de seleção, as variáveis ‘escolaridade’, ‘sexo’, ‘contexto seguinte’, ‘faixa etária’ e ‘extensão do verbo’.

Palavras-chave: Gerúndio; Sociolinguística; Variação e mudança linguística; Apagamento de /d/; Feira de Santana.

ABSTRACT

This dissertation presents results of a research about the deletion of the stop /d/ in the morpheme *-ndo*, indicator of gerund, in Brazilian Portuguese. It is assumed that the researched phenomenon is a rule of variable nature in which the binary dependent variable is characterized by the maintenance or deletion of /d/ in the morpheme of gerund as in *cantando* (standard form) and *cantano* (non-standard form). The theoretical-methodological framework is the Variationist Sociolinguistics (WEINREICH; LABOV; HERZOG, 2006 [1968]; LABOV, 2008 [1972]). The study was carried out based on 24 sociolinguistic interviews from the city of Feira de Santana-BA, selected from the database of the project “A Língua Portuguesa no Semiárido Baiano” (UEFS). The objective of the research is to describe how the phenomenon is conditioned according to structural and social factors, as well as to observe, in apparent time, if there are signs of linguistic change taking place in the researched community. The social variables controlled were ‘education’, ‘sex’ and ‘age group’, and the linguistic variables were ‘verbal conjugation’, ‘verb length’, ‘following context’ and ‘syntactic structure’. The data were analyzed quantitatively using the computational tool GoldVarb X (SANKOFF; TAGLIAMONTE; SMITH, 2005). The results indicate usage rate of 67.7% of the non-standard variant in the *corpus*. The variable phenomenon is conditioned by linguistic and extra linguistic factors and the variables ‘education’, ‘sex’, ‘following context’, ‘age group’ and ‘verb length’ were statistically relevant in this order of selection.

Keywords: Gerund; Sociolinguistics; Linguistic variation and change; Deletion of /d/; Feira de Santana.

LISTA DE GRÁFICOS

Gráfico 1. Apagamento de /d/ no morfema de gerúndio: Feira de Santana (BA)	81
Gráfico 2. As variantes em competição: Feira de Santana-BA e outras variedades do PB.....	82
Gráfico 3. Frequência de apagamento de /d/ em morfema de gerúndio: Escolaridade	89
Gráfico 4. Frequência de apagamento de /d/ em morfema de gerúndio por informante: norma popular	91
Gráfico 5. Frequência de apagamento de /d/ em morfema de gerúndio por informante: norma culta	92
Gráfico 6. Frequência percentual de apagamento por Faixa Etária.....	96
Gráfico 7. Probabilidade de apagamento em função da Faixa Etária: comparação entre Feira de Santana (BA) e São José do Rio Preto (SP)	98
Gráfico 8. Cruzamento das variáveis Faixa etária e Escolaridade	100
Gráfico 9. Cruzamento das variáveis Faixa etária e Sexo	102
Gráfico 10. Cruzamento das variáveis Escolaridade e Sexo – Apagamento.....	103

LISTA DE FIGURAS

Figura 1. Modelo da Fonologia Lexical do PB	36
Figura 2. Processo fonológico de apagamento de /d/ no morfema <i>-ndo</i>	37
Figura 3. Bloqueio da regra de apagamento de /d/ em ambientes não derivados.....	38
Figura 4. Localização de Feira de Santana no Estado da Bahia.....	61
Figura 5. Vista da cidade de Feira de Santana	63

LISTA DE QUADROS

Quadro 1. Formas nominais do verbo	17
Quadro 2. Estratificação social do o <i>corpus</i> da pesquisa	69
Quadro 3. Ordem de seleção das variáveis independentes pelo <i>software</i> GoldVarb X	83

LISTA DE TABELAS

Tabela 1. Percentual de apagamento por categoria gramatical: Rio de Janeiro-RJ.....	22
---	----

Tabela 2. Apagamento de /d/ no morfema de gerúndio: Brejo Paraibano-PB	24
Tabela 3. Apagamento de /d/ no morfema de gerúndio: São José do Rio Preto-SP	25
Tabela 4. Apagamento de /d/ no morfema de gerúndio: Taboco-MS	26
Tabela 5. Apagamento de /d/ no morfema de gerúndio: Fortaleza-CE	27
Tabela 6. Apagamento de /d/ no morfema de gerúndio: Maceió-AL.....	28
Tabela 7. Apagamento de /d/ no morfema de gerúndio: Cachoeirinha-PB.....	28
Tabela 8. Apagamento de /d/ no morfema de gerúndio por estrutura sintática: São Paulo-SP	29
Tabela 9. Apagamento de /d/ no morfema de gerúndio – orações reduzidas: São Paulo (SP).	30
Tabela 10. Apagamento de /d/ no morfema de gerúndio – perífrases: São Paulo (SP).....	30
Tabela 11. Apagamento de /d/ no morfema de gerúndio: Envira-AM.....	31
Tabela 12. Apagamento de /d/ no morfema de gerúndio em contexto de leitura: Aracaju (SE)	32
Tabela 13. Apagamento de /d/ no morfema de gerúndio: Fortaleza (CE).....	33
Tabela 14. Apagamento de /d/ no morfema de gerúndio na fala de Feira de Santana: Contexto Seguinte	85
Tabela 15. Apagamento de /d/ no morfema de gerúndio na fala de Maceió-AL: Contexto Seguinte	86
Tabela 16. Apagamento de /d/ no morfema de gerúndio na fala de Feira de Santana: Contexto Seguinte (segunda rodada)	86
Tabela 17. Apagamento de /d/ no morfema de gerúndio na fala de Feira de Santana: Extensão do Verbo	87
Tabela 18. Apagamento de /d/ no morfema de gerúndio na fala de Feira de Santana: Escolaridade.....	89
Tabela 19. Apagamento de /d/ no morfema de gerúndio na fala de Feira de Santana: Sexo ...	94
Tabela 20. Apagamento de /d/ no morfema de gerúndio na fala de Feira de Santana: Faixa Etária.....	96
Tabela 21. Apagamento de /d/ no morfema de gerúndio: comparação entre Feira de Santana- BA e São José do Rio Preto-SP	97
Tabela 22. Apagamento de /d/ no morfema de gerúndio na fala de Feira de Santana: cruzamento Faixa Etária e Escolaridade.....	99
Tabela 23. Apagamento de /d/ no morfema de gerúndio na fala de Feira de Santana cruzamento Faixa Etária e Sexo	101
Tabela 24. Apagamento de /d/ no morfema de gerúndio na fala de Feira de Santana: cruzamento Escolaridade e Sexo	102

Tabela 25. Apagamento de /d/ no morfema de gerúndio na fala de Feira de Santana: Conjugação Verbal	105
Tabela 26. Apagamento de /d/ no morfema de gerúndio na fala de Feira de Santana: Estrutura Sintática	106

SUMÁRIO

INTRODUÇÃO	14
1.1 DEFININDO O GERÚNDIO	17
1.2 ESTUDOS SOBRE O APAGAMENTO DE /D/ NO MORFEMA DE GERÚNDIO	22
1.3 ANÁLISE FONOLÓGICA DO APAGAMENTO DE /d/ EM /Ndo/.....	34
1.3.1 A teoria da Fonologia Lexical	34
1.3.2 A análise de Ferreira (2010)	37
2 LÍNGUA, SISTEMA, VARIAÇÃO E MUDANÇA	40
2.1 DA LINGUÍSTICA À SOCIOLINGUÍSTICA.....	40
2.1.1 A Sociolinguística Variacionista	44
2.1.1.1 O problema dos fatores condicionantes.....	46
2.1.1.2 O problema da transição	47
2.1.1.3 O problema do encaixamento	47
2.1.1.4 O problema da avaliação	48
2.1.1.5 O problema da implementação	49
2.1.1.6 O tempo real e o tempo aparente	49
2.1.1.7. O estudo de Labov em Martha's Vineyard.....	51
2.2 AS DIMENSÕES INTERNA E EXTERNA DA VARIAÇÃO LINGUÍSTICA	54
2.2.1 A variação diatópica	55
2.2.2 A variação diastrática	57
2.2.3 A variação estilística	58
2.2.4 A variação diamésica	59
3 O CAMINHO METODOLÓGICO	61
3.1 A COMUNIDADE DE FALA	61
3.2 O BANCO DE DADOS	67
3.3 O <i>CORPUS</i>	68
3.4 A VARIÁVEL DEPENDENTE.....	69

3.5 AS VARIÁVEIS INDEPENDENTES SOCIAIS.....	70
3.5.1 Escolaridade.....	70
3.5.2 Sexo.....	71
3.5.3 Faixa etária.....	71
3.6 AS VARIÁVEIS INDEPENDENTES LINGUÍSTICAS.....	72
3.6.1 Conjugação verbal.....	72
3.6.2 Extensão do verbo.....	73
3.6.3 Contexto seguinte.....	74
3.6.4 Estrutura sintática.....	75
3.7 O MODELO ESTATÍSTICO APLICADO NO TRATAMENTO DOS DADOS	77
3.7.1 A ferramenta computacional para a análise estatística	78
4 APRESENTAÇÃO E ANÁLISE DOS DADOS	81
4.1 DADOS GERAIS	81
4.2 ANÁLISE DAS VARIÁVEIS ESTATISTICAMENTE SIGNIFICATIVAS.....	83
4.2.1 Análise das variáveis linguísticas	84
4.2.1.1 Contexto seguinte	84
4.2.1.2 Extensão do verbo	87
4.2.3 Análise das variáveis sociais	88
4.2.3.1 Escolaridade.....	88
4.2.3.2 Sexo	94
4.2.3.3 Faixa etária	95
4.3 CRUZAMENTOS DAS VARIÁVEIS SOCIAIS	99
4.4 VARIÁVEIS DESCARTADAS NA ANÁLISE ESTATÍSTICA	104
4.4.1 Conjugação verbal.....	104
4.4.2 Estrutura sintática.....	106
CONSIDERAÇÕES FINAIS.....	108
REFERÊNCIAS	113

INTRODUÇÃO

Esta dissertação tem como objeto de análise o apagamento da oclusiva /d/ no morfema de verbos no gerúndio na fala de Feira de Santana-BA. O gerúndio no português brasileiro (doravante PB) é gramaticalmente marcado pelo morfema *-ndo*, adicionado ao tema verbal, como em *cantando*, *dizendo* e *partindo*. O ouvinte atento, contudo, pode atestar que formas equivalentes do tipo *cantano*, *dizeno* e *partino*, em que a consoante /d/ é apagada no morfema *-ndo*, são comumente usadas na língua falada, como está exemplificado a seguir.

- (1) *Acrescenta um pouco de água, aí depois você joga a canela e o cravo e mexe como se você tivesse fazeno mingau.* (M1B33)
- (2) *Eu ia passano no caminho quando eu olhei, ela tava vino atrás de mim.* (M2S49)
- (3) *Às vez o pai fala, vai passano a mão pela cabeça, o filho vai aprontano, e aí quando vai vê já é tarde demais.* (H1B33)

O apagamento de /d/ está documentado em estudos dialetológicos como o de Amaral (1982 [1920]), que aborda esse tema ao caracterizar o dialeto caipira, e o de Marroquim (1934), que procurou descrever a fala do Nordeste. Inúmeras pesquisas (FERREIRA, 2010; VIEIRA, 2011; NASCIMENTO; ARAÚJO; CARVALHO, 2013; ALMEIDA; OLIVEIRA, 2017; MACEDO et al., 2018; ARAÚJO, 2019; GONÇALVES, 2018; ALMEIDA; PEREIRA; ARAÚJO, 2020) indicam se tratar de uma regra variável, que é condicionada por fatores internos e externos à língua, de modo que os fatores de ordem social condicionam mais fortemente a regra do que aqueles de ordem linguística.

A temática vem sendo bastante discutida nos últimos anos, sobretudo do ponto de vista da Sociolinguística. Não há, contudo, pesquisa anterior à apresentada nesta dissertação que tenha se debruçado ao estudo sobre o apagamento de /d/ no morfema de gerúndio com dados de Feira de Santana-BA. A pesquisa contribui para a descrição da fala feirense e permite comparar esta variedade de fala do semiárido baiano com outras já estudadas, algumas das quais apresentadas neste texto.

A concepção teórica que dá base à pesquisa é a da Sociolinguística. As lentes pelas quais se observou e analisou o objeto de pesquisa estiveram sob uma ótica de língua como fenômeno social que carrega as características sócio-históricas e culturais dos sujeitos falantes. Dentro dessa abordagem teórica de compreensão dos fatos da língua, entende-se a

importância de analisar como um fenômeno linguístico está condicionado não só pela estrutura interna, mas também pela estrutura social da comunidade de fala.

Por se tratar de uma pesquisa em que se trabalhou com uma grande quantidade de dados, a metodologia adotada foi a da Sociolinguística Variacionista (LABOV, 2008 [1972]), também conhecida como Teoria da Variação ou Sociolinguística Quantitativa. Tal metodologia lida com fenômenos variáveis, que é o caso do apagamento de /d/ no morfema de gerúndio no PB. A Teoria da Variação dispõe de um instrumental bastante sofisticado para análise estatística de dados linguísticos, e de um arcabouço teórico que permite a interpretação de resultados numéricos com vistas a explicar como fenômenos variáveis são condicionados ou não, simultaneamente, por fatores internos e externos à estrutura da língua.

A pesquisa apresentada nesta dissertação é, portanto, de natureza quantitativa e qualitativa, uma vez que procura interpretar resultados numéricos com base em uma teoria da linguagem. De tal modo, utilizou-se o programa computacional GoldVarb X (SANKOFF; TAGLIAMONTE; SMITH, 2005) para a análise quantitativa dos dados. Esse *software* opera análise estatística a partir do modelo de regressão logística multinível e fornece resultados probabilísticos de como as variáveis independentes influenciam a variável dependente.

Na pesquisa sobre o apagamento de /d/ no morfema de gerúndio na fala feirense, foram controladas as variáveis independentes de ordem linguística ‘conjugação verbal’, ‘extensão do verbo’, ‘contexto seguinte’ e ‘estrutura sintática’ e as variáveis independentes de ordem social ‘escolaridade’, ‘faixa etária’ e ‘sexo’. O *corpus* da pesquisa foi constituído a partir de inquéritos do banco de dados do projeto “A Língua Portuguesa no Semiárido Baiano”, sediado no Departamento de Letras e Artes da Universidade Estadual de Feira de Santana (UEFS).

Tomando como base o pressuposto teórico e metodológico já apresentado acima, a pesquisa teve como objetivo responder às seguintes perguntas: a) como o fenômeno de apagamento de /d/ no morfema de gerúndio está encaixado nas estruturas linguística e social da comunidade feirense? e b) o fenômeno configura um caso de variação estável ou de mudança em curso na comunidade estudada? Acredita-se que a resposta a tais perguntas contribui para caracterizar a variedade de fala investigada no tocante ao fenômeno analisado.

No primeiro capítulo deste texto, faz-se uma breve apresentação do que vem a ser a forma verbal de gerúndio na língua portuguesa, ocasião em que se relata uma breve história do gerúndio, desde o latim até português. Nessa mesma seção, apresenta-se também uma série de trabalhos que pesquisaram o mesmo fenômeno, com destaque para os estudos de natureza variacionista. Apresenta-se ainda, no primeiro capítulo, uma discussão acerca da natureza do

processo de apagamento de /d/ na sequência fonológica /Ndo/, a partir do estudo realizado por Ferreira (2010), que toma como aporte teórico a Fonologia Lexical.

No segundo capítulo do texto, é apresentado o referencial teórico que embasa a pesquisa, onde é traçado um percurso sucinto sobre a origem da Sociolinguística, apresentando alguns conceitos básicos dessa disciplina da Linguística geral até chegar à corrente Variacionista, em que se buscou instrumental metodológico para realização da pesquisa.

No terceiro capítulo, onde se define o caminho metodológico da pesquisa, são apresentados a comunidade de fala, a amostra, o *corpus*, a variável dependente, as variáveis independentes, a escolha do modelo estatístico para o tratamento dos dados e o programa computacional que subsidiou a análise quantitativa.

No quarto capítulo, estão os dados e as análises, iniciando-se a discussão pelas variáveis linguísticas para depois as sociais, finalizando com o cruzamento de variáveis sociais na busca de melhor compreensão sobre a existência de uma possível mudança linguística em curso.

Por fim, apresentam-se as conclusões da pesquisa, em que os resultados alcançados na seção de análises são retomados de maneira resumida com base na teoria que fundamentou a pesquisa e nos estudos anteriores apresentados na primeira seção do texto.

1 O GERÚNDIO NA LÍNGUA PORTUGUESA

Procura-se nesta primeira seção caracterizar a forma verbal de gerúndio na língua portuguesa. Na seção 1.1 apresenta-se o conceito de gerúndio, fazendo um breve histórico da origem dessa forma verbal do latim à língua portuguesa. Exemplos dos diversos usos do gerúndio no PB são apresentados com base em pesquisas realizadas em textos contemporâneos. Na segunda parte da seção, seção 1.2, apresenta-se um estado da arte, dentro do escopo da Sociolinguística, sobre o apagamento da oclusiva /d/ no morfema *-ndo*, indicador de gerúndio no PB.

1.1 DEFININDO O GERÚNDIO

A forma verbal de gerúndio constitui o ambiente linguístico principal do objeto da pesquisa que ora se apresenta, visto que a pesquisa consiste na investigação do apagamento da oclusiva /d/ no morfema *-ndo* nas formas verbais de gerúndio, com fins de descrever como o fenômeno está caracterizado na comunidade estudada quanto a condicionadores linguísticos e sociais. Os verbos no gerúndio são, portanto, o ambiente lexical, morfológico, fonológico e gramatical da investigação.

Segundo Castilho (2019), as formas nominais do verbo, dentre as quais está o gerúndio, encontram-se na fronteira entre os verbos propriamente ditos e os nomes, de modo que separando de um lado os verbos e do outro os nomes, as formas nominais do verbo ficam na coluna do meio. O autor representa este esquema no quadro a seguir:

Quadro 1. Formas nominais do verbo

/+verbo/	/±verbo/	/-verbo/
Verbos	Infinitivo Particípio Gerúndio	Substantivo Adjetivo Advérbio

Fonte: Castilho (2019, p. 408) (adaptado).

Segundo Castilho (2019, p. 408), “basta você colecionar usos dessas formas para notar que infinitivo, particípio e gerúndio não são nem prototipicamente verbais, nem prototipicamente substantivos, adjetivos ou advérbio”. O autor afirma que, ao se dar conta desse aspecto não definido das formas nominais do verbo, os gramáticos latinos a puseram,

sabiamente, na classe dos *participios*; sendo que o termo participio, do latim *partis + capio*, significa “tomar parte”. As formas nominais do verbo, portanto, tomam parte nas duas categorias: a dos nomes e a dos verbos. Segundo Castilho (2019), embora o termo “participio” tenha sido bem empregado pelos latinos, foi restringido na nomenclatura gramatical tradicional, à forma que conhecemos hoje por participio, aquela terminada em *-do/ -to*.

Mais especificamente sobre o gerúndio, Castilho (2019) argumenta que essa forma verbal pode ser usada ora com estatuto categorial mais próximo à classe dos nomes ora com estatuto categorial mais próximo à classe dos verbos. Assim, o gerúndio pode funcionar como núcleo de sentenças simples, núcleo de sentenças subordinadas, adjunto adnominal e como adjunto adverbial. Listam-se abaixo os exemplos desses usos do gerúndio anotados pelo autor (p. 409).

Como núcleo de sentença simples (+verbo)

(4) *Você sempre ameaçando!*

(5) *Circulando, circulando!*

Como núcleo de sentença subordinada (+verbo)

(6) *Dizendo/Tendo dito aquelas palavras, despediu-se.*

(7) *Fazendo-se tarde/Tendo-se feito tarde, saimos.*

Como adjunto adnominal restritivo (-verbo)

(8) *Queimou-se com água fervendo.*

(9) *Vi seu pai tomando um refresco.*

Como adjunto adnominal explicativo (-verbo)

(10) *Lá vai o pelotão dos recrutas marchando sob as ordens do sargento.*

(11) *A luz do sol, passando pela fresta da porta, atingia o espelho.*

Como adjunto adverbial (-verbo)

(12) *Saiu gritando.*

(13) *Durante aqueles dias nos distraímos brincando de pega-pega.*

Do ponto de vista histórico, Souza (2003) afirma que o gerúndio em latim era um substantivo verbal neutro caracterizado morfológicamente pelo sufixo *-nd-*, de sentido neutro

e que apresentava as terminações de acusativo, ablativo, dativo e genitivo singular, o que era uma característica dos nomes neutros da segunda declinação. Segundo a autora, o gerúndio latino era uma forma de natureza anfíbia, voltada ao mesmo tempo para a classe dos nomes e para a classe dos verbos.

Listam-se abaixo exemplos dos usos do gerúndio latino para os quatro casos citados. Os exemplos seguidos de suas respectivas traduções são retirados de Besselaar (1960 apud CARVALHO, 2018, p. 98):

Gerúndio genitivo:

(14) *Ars scribendi utilis est.* (A arte de escrever é útil)

Gerúndio dativo:

(15) *Amicus meus aptus est scribendo.* (Meu amigo é apto a escrever)

Gerúndio acusativo:

(16) *Amicus meus domum rediit ad scribendum.* (Meu amigo voltou à casa para escrever)

Gerúndio ablativo:

(17) *Amicus meus scribendo certiore me feci de adventu suo.* (Meu amigo, por escrito, me informou da sua chegada.)

Ao descrever as funções do gerúndio em cada um dos quatro casos exemplificados acima, Besselaar (1960 apud CARVALHO, 2018) assinala que, em cada um dos casos, o gerúndio latino assume uma função sintático-semântica distinta. Segundo o autor, no genitivo, o gerúndio acompanha adjetivos, substantivos e posições com sentido genérico, particularizando-os, sendo que há regras distintas de combinação a depender da classe da palavra que o gerúndio acompanha. No caso dativo, o gerúndio se associa a substantivos, adjetivos e também a verbos, dando sentido de finalidade quando ligado a nomes ou assumindo função de objeto direto quando ligado a verbos. No ablativo, o gerúndio assumia sentido de circunstância quando não preposicionado e sentido de simultaneidade, origem, e outros, quando preposicionado. No dativo, caso assumido em ocasiões em que o verbo acompanhado era preposicionado, a forma de gerúndio assumia função de objeto indireto.

Segundo Campos (1972), o gerúndio como é conhecido hoje originou-se da forma ablativa do gerúndio latino, que tinha valor adverbial e referia-se diretamente ao verbo da

oração principal. No entanto, o gerúndio na língua portuguesa expandiu-se para além do ablativo latino para expressar outros significados e funções sintáticas além daqueles típicos dos advérbios. Ao estudar textos arcaicos, a autora afirma que o gerúndio no português tinha fundamentalmente uso adverbial, circunstancial, adjetival, coordenado e perífrases.

Com base em textos atuais do PB, Campos (1972) afirma que os usos atestados no português arcaico se desenvolveram ainda mais no português contemporâneo, além de terem surgido, neste último, novos usos não atestados no primeiro. Os exemplos seguintes retirados de Campos (1972) ilustram os novos usos do gerúndio que não foram encontrados em textos arcaicos:

Gerúndio narrativo:

(18) *A casa no escuro, os meninos em redor do fogo, a cachorra Baleia vigiando.* (p. 158)

Gerúndio exclamativo:

(19) *Ô diacho! E a gente precisando tanto de cobre, hem, Marcolino!* (p. 161)

Gerúndio interrogativo:

(20) *Leleco: Que é isso, Zé Gato? Dando pulo de costas por causa de uma violinha?* (p. 162)

Gerúndio coordenado:

(21) *A velha correu logo para a Tia Maria, ajoelhando-se aos seus pés.* (p. 154)

Braga e Coriolano (2007), ao investigarem os vários usos do gerúndio, identificaram que as construções integradas pela forma verbal com morfema *-ndo* exibem diferentes graus de gramaticalização, sendo o grau mais baixo representado pelo predicado verbal de orações desgarradas, e o mais alto aquele em que a forma verbal no gerúndio funciona como marcador discursivo.

Os exemplos seguintes, retirados das referidas autoras, ilustram esses usos.

Gerúndio como predicado verbal de sentenças desgarradas:

(22) *Passarinho cantando, cachorro latindo, criança chorando. O ambiente era de aparente normalidade no morro do chapéu Mangueira, no Leme, terça-feira passada. Mas era só prestar atenção para perceber que os olheiros do tráfico acompanhavam de perto a movimentação.* (p. 177)

Gerúndio em orações complexas integradas:

(23) *Ao sair do veículo, um bandido rendeu a mulher dele, Nilda Ferreira, roubando o cordão de ouro e a aliança que ela usava. Knoller, que não estava armado, reagiu e levou um tiro no ouvido.* (p. 177)

Gerúndio em perífrases verbais:

(24) *Carro, todos eles são perigosos. Isso depende muito da pessoa que está dirigindo, entendeu?* (p. 178)

Gerúndio como elemento conector:

(25) *É preciso que todos cheguem a um mesmo patamar para o Fluminense alcançar objetivos em médio prazo. Isso não quer dizer que não se tenha de trabalhar visando a Taça Rio.* (p. 179)

Gerúndio como marcador discursivo:

(26) *Daqui a pouco, a gente está atuando em cima disso. Normalmente, tudo que eu gosto está... está sempre em cima de comunicação, está sabendo? uma... uma filmagem, uma entrevista, um... sabe?* (p. 179)

Essa breve apresentação sobre o gerúndio buscou definir essa forma verbal na língua portuguesa, identificar a sua origem latina e exemplificar seus diversos usos no PB. Como se pôde perceber, o gerúndio, que o português herdou da forma ablativa do latim, vem adquirindo novas funções sintáticas e novos significados no PB, chegando inclusive a gramaticalizar-se como marcador discursivo, como observaram Braga e Coriolano (2007).

Não houve nesta seção a intenção de fazer uma discussão aprofundada do gerúndio quanto a seu aspecto histórico ou sintático-semântico, pois que isto fugiria ao escopo da pesquisa ora apresentada. A apresentação e caracterização do gerúndio, contudo, fez-se necessária uma vez que, como já foi posto acima, a forma verbal de gerúndio é o ambiente linguístico da investigação apresentada neste texto.

Na seção seguinte é apresentado o estudo da arte sobre o apagamento de /d/ no morfema de gerúndio, tomando como base estudos de natureza variacionista, visto que a pesquisa apresentada nesta dissertação segue os pressupostos teórico-metodológicos da Sociolinguística Variacionista.

1.2 ESTUDOS SOBRE O APAGAMENTO DE /D/ NO MORFEMA DE GERÚNDIO

Amaral (1982 [1920]), na descrição do dialeto caipira, registrou que era comum naquela variedade de fala a queda da oclusiva /d/ em formas verbais do tipo andando e vendo (formas no gerúndio dos verbos andar e ver, respectivamente), que eram pronunciadas como *andano* e *veno*. Também Marroquim (1934), na sua descrição da língua do Nordeste, definiu o apagamento da oclusiva /d/ no morfema de gerúndio como uma das características da variedade de fala nordestina, que de acordo com o autor, tratava-se de um caso de assimilação.

Cristófarro Silva (1996) afirma que o apagamento da oclusiva /d/ no grupo /Ndo/ é um processo que se aplica em informação dada pelo componente fonológico e morfológico e que se aplica à desinência verbal de gerúndio, representada pelo morfema *-ndo*, de tal modo que a regra de apagamento é bloqueada quando a sequência /Nd/ está na raiz da palavra, como *lindo* (adjetivo), *Fernando* (nome próprio) e *mando* (verbo na 1ª pessoa do presente do indicativo).

Pesquisas feitas em diversas variedades do PB têm confirmado que o fato de /Ndo/ fazer parte da base da palavra constitui barreira para o apagamento de /d/, contudo, a regra opera, ainda que em frequência muito baixa, nesses ambientes não derivados, como foi demonstrado por Mollica e Mattos (1992).

Pesquisando o apagamento de /d/ em /Ndo/ na fala do Rio de Janeiro, Mollica e Mattos (1992) constataram que o fenômeno não se restringe à classe dos verbos no gerúndio, embora seja esse o ambiente mais propício ao apagamento. O estudo foi realizado com um *corpus* formado por 64 entrevistas de cariocas não universitários, da Amostra Senso. Os resultados por classe de palavras podem ser vistos na Tabela 1.

Tabela 1. Percentual de apagamento por categoria gramatical: Rio de Janeiro-RJ

Nome Próprio	Adjetivo	Numeral	Verbo pres.	Nome comum	Conector	Gerúndio
12/12	28/28	1/62	7/111	26/362	193/1049	1037/2673
0%	0%	2%	6%	7%	18%	39%

Fonte: Mollica e Mattos (1992, p. 58) (adaptada).

Mollica e Mattos (1996) constataram também que as variáveis ‘contexto seguinte’ e ‘extensão do vocábulo’ estão relacionadas ao apagamento de /d/, sendo que a frequência de

apagamento é maior quando o vocábulo tem mais de duas sílabas e quando há segmento no contexto seguinte, ou seja, não há pausa na fala depois da pronúncia do vocábulo com /Ndo/.

Nascimento e Mota (2004) realizaram, a partir de inquéritos experimentais do Projeto Atlas Linguístico do Brasil (ALiB), uma pesquisa em que analisaram o papel das variáveis ‘sexo’ e ‘grau de monitoramento’ no fenômeno de apagamento de /d/ no morfema de gerúndio no PB. As autoras compararam os resultados obtidos a partir de inquéritos de 12 falantes baianos, sendo quatro deles de Jequié, quatro de Santo Amaro e quatro de Salvador, com os resultados obtidos de oito falantes, cada um de uma região diferente do país, e obtiveram resultados semelhantes nas duas amostras, o que leva a concluir que o fenômeno está espreado na fala dos brasileiros. As pesquisadoras constataram, a partir da amostra tomada, que a taxa de apagamento de /d/ no morfema *-ndo*, era mais alta entre os falantes do sexo masculino, com percentual de 68%, do que entre os falantes do sexo feminino, com percentual de 47%.

Nascimento e Mota (2004) constataram que, quando o grau de monitoramento da fala aumentou, a taxa de apagamento desceu para 39% para os homens e 27% para as mulheres. Os dados para aferição do grau de monitoramento foram obtidos por meio de questionário fonético-fonológico e os resultados levaram as pesquisadoras a concluir que os falantes têm alto grau de consciência linguística do valor da forma padrão e da forma inovadora, sendo esta última avaliada de maneira estigmatizada. As autoras concluíram que o fenômeno não produz mudança no léxico, pois que, de acordo com o *corpus* da pesquisa, atinge apenas o morfema de verbos no gerúndio.

Lucena e Vasconcelos (2007) estudaram o apagamento da oclusiva /d/ no grupo /Ndo/ na fala culta na região do Brejo Paraibano-PB a partir de seis entrevistas sociolinguísticas de informantes universitários estratificados em dois sexos (os autores não informam precisamente o número de homens e mulheres na amostra). Foram controladas na pesquisa, além da variável social ‘sexo’, as variáveis linguísticas ‘classe de palavra’, ‘conjugação verbal’, ‘contexto seguinte’ e ‘extensão do vocábulo’.

Lucena e Vasconcelos (2007) constataram que das 281 ocorrências de gerúndio na amostra, 92 (32,7%) foram de apagamento. Foram selecionadas como significativas pelo programa GoldVarb X, em ordem de relevância, as variáveis ‘classe de palavra’, ‘sexo’ e ‘conjugação verbal’. O fato de a variável ‘classe de palavra’ ter sido a primeira a ser selecionada apontou para o fato de serem os verbos no gerúndio o contexto mais provável de ocorrer o apagamento.

O estudo realizado com a amostra do Brejo Paraibano-PB mostrou que as mulheres são mais conservadoras, pois dão preferência à forma padrão, enquanto os homens favorecem o apagamento. Quanto à conjugação verbal, os verbos de segunda conjugação mostraram-se favorecedores do apagamento e os da 1ª e 2ª conjugações inibidores. A Tabela 2 resume os resultados obtidos por Lucena e Vasconcelos na amostra de fala do Brejo Paraibano-PB.

Tabela 2. Apagamento de /d/ no morfema de gerúndio: Brejo Paraibano-PB

Grupos de fatores		Apl./Tot.	%	PR
Classe de palavras	Verbo no gerúndio + verbos em geral	90/147	61,2	0,93
	Conjunção + substantivos	2/134	1,5	0,07
Sexo	Masculino	51	37,2	0,64
	Feminino	41	29	0,38
Conjugação verbal	1ª	63/184	34,2	0,45
	2ª	18/41	43,9	0,71
	3ª	11/56	19,6	0,48

Fonte: Elaborada pelo autor a partir dos dados de Lucena e Vasconcelos (2007, p. 237-238).

Em sua pesquisa de Mestrado, Ferreira (2010) realizou uma análise sociolinguística do mesmo fenômeno buscando descrever como fatores linguísticos e sociais condicionam a o apagamento de /d/ no morfema de gerúndio na fala de São José do Rio Preto, no interior do Estado de São Paulo. A autora controlou três variáveis sociais e quatro variáveis linguísticas, sendo elas ‘sexo’, ‘faixa etária’, ‘escolaridade’, ‘contexto sintático’ (perífrase, oração reduzida, justaposição), ‘tipo de verbo auxiliar da perífrase’ (v1), ‘existência ou não de material interveniente entre v1 e o verbo no gerúndio’ e ‘tamanho do material interveniente’.

A partir dos valores obtidos por meio do *software* de análise estatística GoldVarb X, Ferreira (2010) constatou um quantitativo de 716 ocorrências de apagamento de um total de 999 dados, o que equivale a uma taxa de 72% de apagamento. Os resultados apontaram para um alto índice de aplicação da regra de apagamento de /d/ no morfema de gerúndio no *corpus* estudado. A análise quantitativa mostrou que os fatores sociais foram mais decisivos para a compreensão do fenômeno do que os fatores linguísticos, uma vez que a ordem de importância das variáveis selecionadas no programa computacional GoldVarb X foram, nesta ordem, ‘sexo’, ‘escolaridade’, ‘faixa etária’ e ‘estrutura sintática’.

Ferreira (2010) constatou que a forma variante em que há o apagamento de /d/ no morfema de gerúndio é mais frequente na fala dos homens (77%) do que na fala das mulheres (63%), que o fenômeno é favorecido pelos mais jovens, sendo 83% a taxa de apagamento para a faixa de 7 a 15 anos contra 36% para a faixa acima de 55 anos. O fenômeno é também favorecido pelos sujeitos menos escolarizados, sendo 86% a taxa de apagamento entre os

informantes que cursaram apenas o primeiro ciclo do Ensino Fundamental contra 49% para os indivíduos que cursaram o nível universitário.

Para as variáveis linguísticas, a pesquisa de Ferreira (2010) mostrou a estrutura sintática de perífrase como ambiente motivador do apagamento. Nesse ambiente, a taxa de aplicação foi de 72% e o peso relativo foi de 0,51. A estrutura sintática de oração reduzida mostrou-se inibidora (PR 0,45) e o contexto de justaposição foi retirado da rodada por representar menos de 5% dos dados. As variáveis tipo de V1 da perífrase e material interveniente entre V1 e *V+ndo* foram descartadas. Os resultados apontaram, segundo a autora, para uma mudança em curso, visto que a forma variante em que não há a realização da oclusiva /d/ no morfema *-ndo* é a forma preferida entre os falantes mais jovens e tem baixa aplicação entre os mais velhos. A Tabela 3 resume os resultados obtidos por Ferreira (2010).

Tabela 3. Apagamento de /d/ no morfema de gerúndio: São José do Rio Preto-SP

Grupos de fatores		Apl./Tot.	%	PR
Sexo	Feminino	315/497	63	0,40
	Masculino	327/426	77	0,59
Escolaridade	Fundamental I	111/129	86	0,72
	Fundamental II	177/232	76	0,57
	Ensino Médio	241/332	73	0,48
	Universitária	113/230	49	0,32
Faixa etária	7 a 15 anos	168/201	83	0,63
	16 a 25 anos	155/201	77	0,56
	26 a 35 anos	138/174	79	0,64
	36 a 55 anos	130/204	63	0,44
	Acima de 55 anos	51/143	35	0,17
Estrutura sintática	Perífrase	466/643	72	0,51
	Oração reduzida	223/322	69	0,45
	Justaposição	23/25	92	0,80

Fonte: Elaborada pelo autor a partir dos dados de Ferreira (2010, p. 97-108).

Em Taboco, distrito pertencente à cidade de Corguinho, no noroeste de Mato Grosso do Sul, Vieira (2011) analisou a fala de 16 informantes, também sob a perspectiva da Sociolinguística, e constatou frequência de 79,6% de aplicação da regra variável no dialeto estudado. Foram 386 casos de apagamento de um total de 485 ocorrências de verbos no gerúndio. Vieira (2011) verificou que os homens apagam mais a oclusiva /d/ no morfema *-ndo* do que as mulheres. A autora constatou também que o apagamento é proporcional ao tamanho do verbo, sendo que quanto maior o número de sílabas, maior a possibilidade de ocorrer o apagamento de /d/ no morfema *-ndo*.

Vieira (2011) não constatou diferença no comportamento entre jovens e mais velhos, sendo a taxa de aplicação da regra de apagamento de 75% para ambas as faixas etárias. Os dados da pesquisa foram apresentados em frequência de uso. Os resultados alcançados na pesquisa sobre a fala do distrito de Taboco podem ser visualizados na Tabela 4.

Tabela 4. Apagamento de /d/ no morfema de gerúndio: Taboco-MS

Grupos de fatores		Apl./Tot.	%
Sexo	Feminino	180/255	70
	Masculino	182/229	80
Faixa etária	15 a 35 anos	163/218	75
	+ de 50 anos	199/268	75
Conjugação	1 ^a	299/334	68
	2 ^a	92/105	88
	3 ^a	41/45	92
Extensão do vocábulo	Dissílabo	106/209	50
	Trissílabo	156/167	94
	Polissílabo	100/108	93
Contexto seguinte	Vogal	129/194	64
	Consoante	150/195	76
	Pausa	78/95	82

Fonte: Elaborada pelo autor a partir dos dados de Vieira (2011, p. 17-22).

Nascimento, Araújo e Carvalho (2013) analisaram dados de 24 entrevistas de falantes fortalezenses constantes do banco de dados do Projeto Norma Oral do Português Popular de Fortaleza (NORPOFOR). No referido estudo, as autoras controlaram as variáveis linguísticas ‘conjugação verbal’, ‘contexto seguinte’ e ‘extensão do verbo’ e as variáveis sociais ‘faixa etária’, ‘escolarização’ e ‘sexo’ dos informantes.

As autoras realizaram análise estatística a partir do *software* GoldVarb X e obtiveram um total de 477 dados dos quais 357 (75%) foram de apagamento. Dentre as variáveis controladas pelas autoras, o programa estatístico GoldVarb X selecionou como mais relevante o ‘grau de escolaridade’, seguido das variáveis ‘faixa etária’, ‘conjugação verbal’, ‘sexo’ e ‘contexto seguinte’; a variável ‘extensão do verbo’ não se mostrou relevante no estudo. O estudo mostrou que o apagamento é favorecido por falantes mulheres, com baixo nível de escolaridade e de faixa etária intermediária. Os resultados obtidos por Nascimento, Araújo e Carvalho (2013) podem ser visualizados na Tabela 5.

Tabela 5. Apagamento de /d/ no morfema de gerúndio: Fortaleza-CE

Grupos de fatores		Apl./Tot.	%	PR
Escolaridade	0 a 4 anos	170/191	89	0,75
	9 a 11 anos	175/274	64	0,31
Faixa etária	15 a 25 anos	105/137	77	0,54
	26 a 49 anos	144/166	87	0,71
	A partir de 50 anos	96/162	59	0,25
Conjugação verbal	1 ^a	247/320	77	0,53
	2 ^a	80/110	73	0,52
	3 ^a	18/35	48	0,19
Sexo	Feminino	179/244	81	0,57
	Masculino	166/244	68	0,43
Contexto fonológico posterior	Consoantes	206/240	84	0,63
	Pausa	43/75	57	0,31
	Vogal	96/150	64	0,31

Fonte: Elaborada pelo autor a partir dos dados de Nascimento, Araújo e Carvalho (2013, p. 407-408).

Almeida e Oliveira (2017) estudaram o apagamento de /d/ no morfema de verbos no gerúndio na fala de Maceió-AL. Os autores constituíram o *corpus* da pesquisa a partir de 30 entrevistas selecionadas da amostra do Projeto “Variação Linguística do Português Alagoano” (PORTAL), da Universidade Federal de Alagoas (UFAL). Foram controladas as variáveis sociais ‘sexo’ e ‘faixa etária’ e as variáveis linguísticas ‘extensão do verbo’, ‘conjugação verbal’ e ‘contexto seguinte’.

Os resultados da análise realizada por Almeida e Oliveira (2017) mostraram um total de 524 ocorrências de verbos no gerúndio, das quais 219 (42%) foram de apagamento. Foram significativas as variáveis ‘sexo’, ‘extensão do vocábulo’ e ‘contexto seguinte’. O estudo mostrou que, na fala maceioense, indivíduos do sexo masculino tendem a apagar mais a oclusiva /d/ no morfema de gerúndio (52%) do que as mulheres (30%), e que no contexto seguinte as vogais favorecem o apagamento.

Sobre a variável ‘contexto seguinte’, Almeida e Oliveira (2017) assumem que, embora a análise estatística indique ser o contexto de haplologia o grande favorecedor do apagamento, a haplologia é um processo fonético-fonológico causado pela presença de segmentos semelhantes ou iguais contíguos, e por isso é foneticamente motivado e não tem relação direta com o apagamento de /d/ no morfema *-ndo*. Os autores, portanto, descartaram o contexto de haplologia e consideraram ser o contexto de vogais o ambiente favorecedor da regra de apagamento no caso estudado, conforme Tabela 6, a seguir.

Tabela 6. Apagamento de /d/ no morfema de gerúndio: Maceió-AL

Grupos de fatores		Apl./Tot.	%	PR
Sexo	Feminino	75/247	30,4	0,39
	Masculino	144/277	52	0,61
Extensão do vocábulo	Dissílabo	16/54	29,6	0,37
	Trissílabo	128/324	39,5	0,51
	Polissílabo	75/146	51,4	0,63
Contexto seguinte	Haplologia	77/88	87,5	0,91
	Vogal	85/168	50,6	0,57
	Consoante	43/144	29,9	0,37
	Pausa	14/124	11,3	0,12

Fonte: Elaborada pelo autor a partir dos dados de Almeida e Oliveira (2017, p. 206-207).

Os valores da Tabela 6 mostram que o ambiente de haplologia é o que tem peso relativo mais alto dentre os fatores da variável ‘contexto seguinte’. Porém os autores consideram que o apagamento de /d/ no morfema de gerúndio é um fenômeno morfofonológico, e que a haplologia age naturalmente em qualquer ambiente foneticamente propício.

Macêdo et al. (2018) investigaram o apagamento da oclusiva /d/ no morfema de gerúndio na fala de Cachoeirinha, cidade localizada no agreste pernambucano. Os autores realizaram a pesquisa a partir de um corpus constituído por 36 entrevistas sociolinguísticas estratificadas em sexo, escolaridade e três faixas etárias (de 15 a 30 anos, de 31 a 45 anos, de 46 a 61 anos). Foram controladas no estudo as variáveis linguísticas ‘conjugação verbal’ e ‘extensão do verbo’.

Após a análise dos dados, Macêdo et al. (2018) constataram 386 ocorrências de verbos no gerúndio, das quais 244 foram de apagamento (63%). Quanto às variáveis controladas na pesquisa, apenas a variável social ‘escolaridade’ mostrou-se estatisticamente relevante, revelando que falantes que cursaram até o Ensino Médio favorecem o apagamento (PR 0,58) e que aqueles que cursaram o Nível Universitário o inibem (PR 0,38). A Tabela 7 mostra os resultados para a variável escolaridade na comunidade de Cachoeirinha-PE.

Tabela 7. Apagamento de /d/ no morfema de gerúndio: Cachoeirinha-PB

Grupo de fatores		Apl./Tot.	%	PR
Escolaridade	Ensino Médio	160/225	71	0,58
	Ensino Universitário	84/161	52	0,38

Fonte: Elaborada pelo autor a partir dos dados de Macedo et al. (2018, p. 252).

Gonçalves (2018) pesquisou o apagamento de /d/ no morfema de gerúndio na fala da cidade de São Paulo-SP e constatou que de 4.071 ocorrências de verbos no gerúndio, 1.070 (26,3%) era de apagamento. Para a realização da pesquisa, foram controladas as variáveis linguísticas ‘verbo auxiliar da perífrase’, ‘tamanho do verbo auxiliar *estar* na perífrase’ (se o auxiliar *estar* ocorre em forma reduzida, como em *tá*, ou em sua forma plena, como em *está*), ‘material interveniente entre o verbo auxiliar e o verbo principal no gerúndio’ (existência ou não existência de material), ‘conjugação verbal’ e ‘extensão do verbo’. Foram controladas as variáveis sociais ‘sexo’, ‘faixa etária’ (25 a 34 anos, 35 a 59 anos, 60 anos ou mais), ‘escolaridade’ (até o Ensino Médio, Ensino Universitário), ‘classe social’ (A, B1, B2, C1, C2), dentre outras. A análise estatística dos dados foi feita com auxílio do *software* R (R Core Team, 2018).

Gonçalves (2018) dividiu os dados de acordo com a estrutura sintática em que ocorre o gerúndio. Para tanto, o autor assumiu, conforme Ferreira (2010) e Ferreira, Gonçalves e Tenani (2012), que o gerúndio ocorre em três estruturas, sendo estas a justaposição (dois ou mais verbos no gerúndio justapostos), as orações reduzidas e as perífrases. De tal forma o autor distribuiu os dados por três fatores, conforme a Tabela 8, a seguir.

Tabela 8. Apagamento de /d/ no morfema de gerúndio por estrutura sintática: São Paulo-SP

Grupo de fatores	Apl./Tot.	%
Justaposição	22/55	40,7
Orações reduzidas	231/1.076	21,5
Perífrase	1.070/4.071	26,3

Fonte: Gonçalves (2018, p. 45) (adaptada).

Pelo fato de os dados de justaposição corresponderem a um percentual muito pequeno na amostra, não foi possível realizar análise estatística para esse ambiente sintático. Após análise estatística dos dados de fala de orações reduzidas e perífrases, o autor constatou resultados muito próximos para os dois ambientes sintáticos. Mostraram-se significativas as variáveis ‘sexo’, ‘escolaridade’, ‘faixa etária’ e ‘conjugação verbal’, sendo que o apagamento é motivado por falantes mais jovens, do sexo masculino e que estudaram até o Ensino Médio.

As Tabelas 9 e 10 mostram, separadamente, os dados para as estruturas sintáticas de oração reduzida e perífrase na fala paulistana.

Tabela 9. Apagamento de /d/ no morfema de gerúndio – orações reduzidas: São Paulo (SP)

Grupos de fatores		Apl./Tot.	%	Significância
Faixa etária	20 a 34 anos	108/395	27	p = 0,0010
	35 a 59 anos	107/345	31	
	60 anos ou +	16/336	5	
Escolaridade	Média	144/533	27	p = 0,0010
	Universitária	87/543	16	
Conjugação	1ª	164/703	23	p = 0,0382
	2ª	54/269	20	
	3ª	13/104	13	

Fonte: Elaborada pelo autor a partir dos dados de Gonçalves (2018, p. 50, 52, 57).

Tabela 10. Apagamento de /d/ no morfema de gerúndio – perífrases: São Paulo (SP)

Grupos de fatores		Apl./Tot.	%	Significância
Sexo	Feminino	296/1.334	22	p = 0,0010
	Masculino	521/1.607	32	
Faixa etária	20 a 34 anos	332/973	34	p = 0,0010
	35 a 59 anos	419/1.149	37	
	60 anos ou +	66/819	8	
Escolaridade	Média	520/1.568	33	p = 0,0010
	Universitária	297/1.373	22	
Conjugação	1ª	474/1.622	29	p = 0,0010
	2ª	287/1.013	28	
	3ª	56/306	18	

Fonte: Elaborada pelo autor a partir dos dados de Gonçalves (2018, p. 60, 61, 62, 66, 68).

Segundo Gonçalves (2018), a variável ‘conjugação verbal’ condicionou os dados de maneira diferente nos dois grupos sintáticos. Nas orações reduzidas apenas a 1ª conjugação favoreceu o apagamento, já nas perífrases a 1ª e a 2ª conjugações favoreceram e a 3ª inibiu o apagamento. A variável ‘sexo’ foi significativa apenas no contexto de perífrases, mostrando que os homens estão à frente no uso da forma não-padrão.

Araújo (2019) investigou o apagamento de /d/ no morfema de gerúndio na fala de Envira, cidade do interior do estado do Amazonas. A pesquisa foi feita a partir de *corpus* constituído pela própria pesquisadora, que gravou 16 informantes, sendo oito da zona urbana e oito da zona rural do município, e um de seus objetivos foi comparar as variedades rural e urbana de Envira no tocante ao apagamento de /d/ no morfema de verbos no gerúndio.

Para a realização das gravações, Araújo (2019) usou um questionário fonético com 49 questões que tinham como respostas verbos no gerúndio. Os dados coletados somaram 784 ocorrências e foram tratados estatisticamente por meio do *software* GoldVarb X. Foram controladas as variáveis sociais ‘sexo’, ‘faixa etária’ (de 18 a 30 anos e de 50 a 65 anos),

‘escolaridade’ (Ensino Fundamental 1 e Ensino Fundamental 2), ‘região’ (urbana, rural) e as variáveis linguísticas ‘conjugação verbal’ e ‘extensão do verbo’.

Os resultados obtidos por Araújo (2019) mostraram um percentual de apagamento de 62%. Foram 484 casos de aplicação de um total de 784 ocorrências de verbos no gerúndio. Foram relevantes as variáveis independentes ‘região da cidade’ (rural/urbana), ‘sexo’ e ‘escolaridade’. As variáveis independentes ‘extensão do verbo’ e ‘conjugação verbal’ foram descartadas.

A Tabela 11 resume os resultados obtidos na pesquisa realizada por Araújo (2019). Tais resultados mostraram que o apagamento de /d/ no morfema de verbos no gerúndio, na fala de Envira, é favorecida pelo fator região, sendo mais produtivo na fala rural, por indivíduos do sexo masculino e que estudaram até o Ensino Fundamental II.

Tabela 11. Apagamento de /d/ no morfema de gerúndio: Envira-AM

Grupos de fatores		Apl./Tot.	%	PR
Região	Urbana	206/393	48	0,40
	Rural	278/391	71	0,60
Sexo	Feminino	185/392	47	0,35
	Masculino	299/392	76	0,65
Escolaridade	Fund. I	211/392	54	0,42
	Fund. II	273/392	70	0,58

Fonte: Elaborada pelo autor a partir dos dados de Araújo (2019, p. 118-124).

Os resultados obtidos por Araújo (2019) para a variável social ‘escolaridade’ destoam daqueles encontrados nos outros estudos citados até aqui, visto que os dados de Envira mostraram que indivíduos mais escolarizados apagam mais. Sobre esse resultado, a autora afirma:

Evidentemente esse resultado deixa uma lacuna que merece ser investigada de forma mais aprofundada, levando em consideração aspectos relevantes como selecionar informantes de classes sociais diferentes, em contextos de fala diversos e outros níveis de escolaridade. Dessa maneira, provavelmente encontraríamos elementos que pudéssemos compreender melhor esse resultado. (ARAÚJO, 2019)

Esse resultado para a variável ‘escolaridade’ surpreendeu a autora, contrariando a hipótese inicial de pesquisa e deixa uma questão muito importante a ser investigada, como bem pontua a autora, na citação acima, sobre o efeito da escolaridade sobre as formas variantes do gerúndio naquela comunidade.

Cardoso, Pinheiro e Silva (2019) observaram o apagamento de /d/ no segmento /Ndo/ em contextos de leitura. O estudo não abordou especificamente o morfema de gerúndio, estendendo-se a outras classes de palavras. Os autores compuseram o *corpus* da pesquisa a partir de gravações de leituras de um mesmo texto realizadas por 20 informantes residentes em Aracajú, sendo dez deles com escolaridade até o Ensino Médio e outros 10 de com Ensino Universitário. O texto foi escrito especificamente para o referido estudo e a análise estatística foi realizada com auxílio do *software* R (R CORE TEAM, 2018). Foram controladas também a variável social ‘sexo’ e as variáveis linguísticas ‘classe de palavra’ (verbos no gerúndio, nomes e conectivos), ‘extensão do vocábulo’ e ‘conjugação verbal’, esta última, no caso, para a classe dos verbos no gerúndio.

Foram contabilizadas por Cardoso, Pinheiro e Silva (2019) 260 ocorrências de palavras com final /Ndo/ no *corpus*, das quais 176 (32%) eram de apagamento. As variáveis sociais ‘escolaridade’ e ‘sexo’ não se mostraram significativas, indicando que, em contexto de leitura, tanto falantes com Ensino Médio quanto aqueles com Ensino Universitário, independentemente do sexo, inibem o apagamento de /d/ no segmento /Ndo/.

No que diz respeito às variáveis linguísticas, mostraram-se significativas a ‘conjugação verbal’ e a ‘extensão do vocábulo’, cujos resultados estão expostos na Tabela 12, adiante. Favoreceram o apagamento a classe dos verbos da 1ª (17%) e 2ª (4%) conjugações, a 3ª conjugação mostrou-se desfavorecedora (1%). A variável extensão do vocábulo mostrou que os vocábulos com três ou mais sílabas (23%) são mais propensos ao apagamento que os dissílabos (9%). Os autores não especificaram o número exato de ocorrências por classe de palavras, apenas explicitaram que, no *corpus*, o apagamento da oclusiva /d/ na sequência /Ndo/ foi muito mais produtivo em verbos no gerúndio (22%) do que em nomes (5%) e conectivos (5%).

Tabela 12. Apagamento de /d/ no morfema de gerúndio em contexto de leitura: Aracaju (SE)

Grupos de fatores		%	Significância
Conjugação verbal	1ª	17	p = 0,033310
	2ª	4	
	3ª	1	
Extensão do vocábulo	Dissílabos	9	p = 0,002697
	Três ou mais sílabas	23	

Fonte: Elaborada pelo autor a partir dos dados de Cardoso, Pinheiro e Silva (2019, p. 186-187).

Almeida, Pereira e Araújo (2020) procurou aprofundar a compreensão sobre o apagamento de /d/ em morfema de gerúndio em Fortaleza-CE, investigando a norma culta

fortalezense. As autoras realizaram a pesquisa a partir de 23 inquéritos, extraídos do banco de dados do Projeto “Português Oral Culto de Fortaleza” (PORCUFORT). A amostra foi estratificada em sexo, faixa etária (22 a 35 anos, 36 a 55 anos e a partir de 56 anos) e tipo de registro (DID e EF). Foram controladas também as variáveis linguísticas ‘conjugação verbal’, ‘contexto seguinte’ e ‘extensão do verbo’. Os dados foram tratados por meio de GoldVarb X. Foram contabilizadas 387 ocorrências de verbos no gerúndio no *corpus*, das quais 80 eram de apagamento (20,7%).

Os resultados da análise sociolinguística realizada por Almeida, Pereira e Araújo (2020) mostraram ser significativas, nesta ordem de relevância, as variáveis ‘extensão do verbo’ e ‘sexo’ dos informantes, sendo que os polissílabos favoreceram (PR 0,57) e os dissílabos (PR 0,14) e trissílabos (PR 0,48) desfavoreceram o apagamento; quanto à variável ‘sexo’, as mulheres favoreceram o apagamento (PR 0,61), enquanto os homens inibiram-na (PR 0,41). As autoras seguiram refinando a análise por meio do cruzamento das variáveis extralinguísticas e constataram ainda que o apagamento é motivado por falantes do sexo feminino em registros informais.

A Tabela 13 mostra os resultados para as duas variáveis significativas na pesquisa realizada na fala culta de Fortaleza.

Tabela 13. Apagamento de /d/ no morfema de gerúndio: Fortaleza (CE)

Grupos de fatores		Apl./Tot.	%	PR
Sexo	Feminino	47/123	27,6	0,61
	Masculino	33/184	15,2	0,41
Extensão do verbo	Dissílabos	1/22	4,3	0,14
	Trissílabos	37/154	19,4	0,49
	Polissílabos	42/131	24,3	0,57

Fonte: Elaborada pelo autor a partir dos dados de Almeida, Pereira e Araújo (2020, p. 527-528).

Foram apresentados nesta seção alguns de estudos que abordaram o fenômeno de apagamento da oclusiva /d/ no morfema de gerúndio no PB e que serviram como ponto de partida para a pesquisa realizada sobre o mesmo fenômeno na fala de Feira de Santana. Esses estudos serviram de parâmetro, na medida em que indicaram quais variáveis linguísticas e sociais podem ajudar a compreender como o fenômeno está encaixado nas estruturas linguística e social da comunidade de fala investigada. Além disso, eles permitiram estabelecer um quadro comparativo da fala de Feira de Santana com diversas variedades do PB no que concerne ao fenômeno investigado.

Na seção seguinte deste primeiro capítulo, procura-se desenvolver uma discussão sobre a natureza morfofonológica do apagamento de /d/ no morfema de gerúndio. O objetivo da exposição é de melhor compreender o fenômeno investigado do ponto de vista morfológico e fonológico. A discussão é levantada a partir da apresentação da pesquisa de Ferreira (2010), em que a autora explica o processo morfofonológico de apagamento tomando como base a Teoria da Fonologia Lexical.

1.3 ANÁLISE FONOLÓGICA DO APAGAMENTO DE /d/ EM /Ndo/

Apresenta-se nesta seção uma discussão sobre a natureza fonológica do apagamento de /d/ no morfema de verbos no gerúndio. O objetivo da exposição é de ampliar a compreensão de como o ao apagamento da oclusiva /d/ na sequência /Ndo/ está condicionado morfofonologicamente. A discussão é realizada com base na análise Ferreira (2010), em que a autora procurou explicar a natureza do fenômeno com base na Teoria da Fonologia Lexical. Primeiro faz-se uma breve apresentação sobre a teoria que embasou a análise de Ferreira (2010) e, em seguida, uma exposição com base no trabalho da autora.

1.3.1 A teoria da Fonologia Lexical

A Fonologia Lexical é, nas palavras de Bisol (2017),

Uma teoria fonológica de base gerativa que tem por objeto a palavra desde a forma mais simples à mais complexa, em todas as minúcias de sua formação e subsequentes alterações, e a frase com suas características fonológicas. (BISOL, 2017, p. 81)

A Fonologia Lexical atua no campo da interação entre a morfologia e a fonologia. E seu modelo de análise compreende dois componentes, um componente lexical e um componente pós-lexical. O primeiro componente tem como domínio a palavra, onde ocorre a interação entre a morfologia e a fonologia. Nesse componente as regras são cíclicas e de aplicação categórica. O segundo componente tem como domínio a frase; nele as regras são acíclicas e de natureza variável. De acordo com Lee (1996),

Para a Fonologia lexical, portanto, o léxico é considerado como a estrutura composta de alguns níveis ordenados, que são os domínios de algumas regras fonológicas, de modo que o componente fonológico existe não só depois da sintaxe, mas também no léxico. (LEE, 1996, p. 131)

Kiparsky (1983) propõe que o léxico de uma língua está organizado em uma série de níveis, sendo estes níveis os domínios para a aplicação de regras morfológicas e fonológicas. O autor define quatro princípios básicos para a Fonologia Lexical, os quais estão sintetizados em Ferreira (2010) nos quatro pontos seguintes:

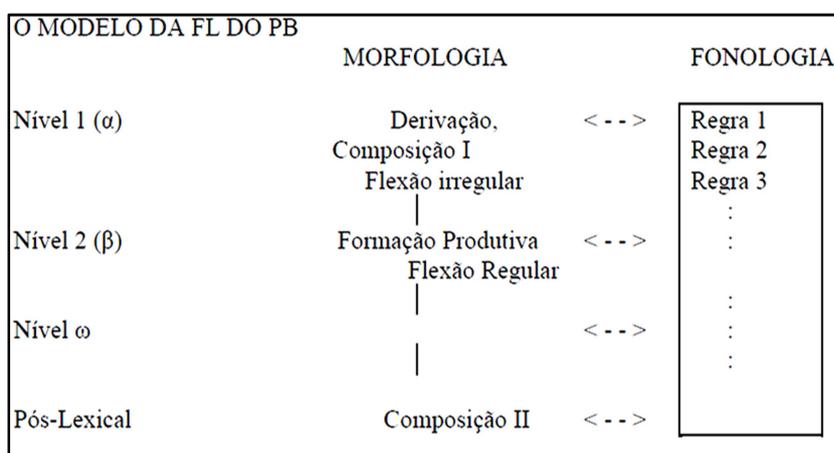
1. Hipótese do domínio fonte: princípio segundo o qual todas as regras fonológicas aplicam-se no nível mais alto do léxico;
2. Preservação da estrutura: segundo o qual somente segmentos contrastivos da representação subjacente da língua (os fonemas) podem ocorrer durante as operações lexicais, de modo que a preservação da estrutura determina os tipos de regras que podem se aplicar no léxico;
3. Condição de ciclo estrito: princípio segundo o qual as regras fonológicas cíclicas, aquelas de natureza lexical, aplicam-se somente em cada ciclo próprio;
4. Hipótese de Referência Indireta: princípio pelo qual se explica a falta de isomorfia entre as estruturas morfológicas e fonológicas, introduzindo a noção de domínio prosódico, ou seja, as regras fonológicas aplicam-se no domínio prosódico e não no domínio morfológico.

Assume-se na teoria da Fonologia Lexical que há dois tipos de regras, as regras lexicais, que envolvem a interface entre os componentes morfológico e fonológico; e as regras pós-lexicais, que envolvem a interface entre os componentes sintático e fonológico. As características das regras lexicais e das regras pós-lexicais podem ser resumidas da seguinte maneira, conforme Kiparsky (1983 apud LEE, 1996):

- 1 As regras lexicais podem se referir à estrutura interna das palavras, enquanto as regras pós-lexicais não podem;
- 2 As regras lexicais são cíclicas, enquanto as regras pós-lexicais não o são;
- 3 As regras lexicais submetem-se à Preservação da Estrutura, enquanto as regras pós-lexicais não se submetem;
- 4 As regras lexicais devem preceder todas as aplicações das regras pós-lexicais;
- 5 As regras lexicais podem ter exceções, enquanto as regras pós-lexicais não podem;
- 6 As regras lexicais devem preceder todas as aplicações das regras pós-lexicais, enquanto as regras pós-lexicais devem ser precedidas de todas as aplicações das regras lexicais;
- 7 As regras lexicais sujeitam-se à ordem disjuntiva, enquanto que as regras pós-lexicais sujeitam-se à ordem conjuntiva. (KIPARSKY, 1983 apud LEE, 1996, p. 131)

Lee (1995) afirma que há dois níveis ordenados no PB, em que regras fonológicas e morfológicas interagem, o nível α e o nível β . O nível α inclui a derivação, a flexão irregular e alguns processos de composição aos quais podem-se adicionar sufixos derivacionais; o nível β inclui os processos de flexão do verbo e do não-verbo e a formação produtiva de palavras, a exemplo das formadas a partir de sufixos como *-inho* e *-zinho*, na formação de diminutivo, e *-mente*, na formação de advérbios. Segundo o autor, há dois tipos de compostos no PB, os compostos lexicais, que se sujeitam aos princípios da Fonologia Lexical, e os compostos pós-lexicais, que não se sujeitam a tais princípios. A figura seguinte representa o modelo proposto pelo autor para a Fonologia Lexical do PB.

Figura 1. Modelo da Fonologia Lexical do PB



Fonte: Lee (1995, p. 11).

A Figura 1 representa a distribuição do léxico do PB em dois níveis. Para exemplificar os processos do nível α , Lee (1995, p. 12) lista os casos de “[feliz], [[felic]idade]; [descobrir], [descoberta]; [[radio-tax]icista], [[puxa-saqu]ismo]”. Já para os processos de nível β , o autor dá como exemplos “falo, falava; flor, flores; cafezinho”.

O nível ω representa a saída do léxico e entrada para a sintaxe, ou seja, a entrada no pós-léxico. O esquema da Figura 1 indica que cada vez que uma regra morfológica se aplica, surge um ambiente para a aplicação de uma regra fonológica que, por sua vez, possibilita a entrada de uma nova regra morfológica, evidenciando assim o caráter cíclico das regras lexicais.

Essa breve exposição dos conceitos básicos da Teoria da Fonologia Lexical não tem o intuito de aprofundar os conhecimentos sobre essa teoria fonológica, mas tão somente de facilitar a compreensão da análise de Ferreira (2010), a qual é apresentada a seguir.

1.3.2 A análise de Ferreira (2010)

Ao estudar o apagamento da oclusiva /d/ na sequência /Ndo/ na fala de São José do Rio Preto-SP, Ferreira (2010) compreendeu que o fenômeno não era condicionado apenas por fatores de natureza fonológica mas também por fatores de natureza morfológica, uma vez que o apagamento não se aplica a todas as classes de palavras, ainda que o ambiente fonológico seja propício¹. Por compreender que esse processo fonológico ocorre em contexto morfológico, a autora propôs uma análise com base na teoria da Fonologia Lexical.

Baseando-se no modelo proposto por Lee (1995) para a Fonologia Lexical do PB, Ferreira (2010) propõe que o gerúndio pode ser visto como pertencente ao nível β , o nível da flexão; e que palavras como *lindo*, *quando*, *mundo* e *segundo*, que são, respectivamente, adjetivo, advérbio, substantivo e numeral, pertencem ao nível α , o nível da derivação.

Ao estudar a variedade de fala de São José do Rio Preto-SP, Ferreira (2010) não encontrou caso de apagamento de /d/ na sequência /Ndo/ em outra classe de palavras que não fosse a dos verbos, especificamente a forma de gerúndio. Segundo a autora, pode-se assumir, com base na Fonologia Lexical, que as formas terminadas em /Ndo/ podem estar em dois níveis do léxico, no nível α ou no nível β ; e que, sendo o nível α o local da derivação e o nível β o local da flexão, as formas verbais de gerúndio estão no nível β , pois *-ndo* é uma forma flexional do verbo. Segundo a autora, o processo de apagamento de /d/ no morfema *-ndo* ocorre conforme representação na Figura 2.

Figura 2. Processo fonológico de apagamento de /d/ no morfema *-ndo*

[falar]	
[fal[a[ndo]]]	sufixação
[fal[ã[ndo]]]	nasalização
[fal[ã[no]]]	apagamento de “d”
[falãnu]	

Fonte: Ferreira (2010, p. 91).

¹ Alguns estudos, contudo, têm documentado casos de apagamento de /d/ em /Ndo/ em palavras que não sejam verbos no gerúndio. Mollica e Mattos (1992), por exemplo, documentaram *munno* para *mundo*, na fala carioca, e Gonçalves (2018) documentou vários casos de *quano* para *quando* na fala paulistana. Esses casos foram esparsos, mas tais ocorrências podem indicar que a regra que apaga a oclusiva /d/ na sequência /Ndo/ pode estar se espalhando para outras classes de palavras. Esses casos foram em menor número, indicando que o fenômeno é característico de verbos no gerúndio. É por ser um caso típico de verbos no gerúndio que se faz pertinente a análise de Ferreira (2010), ainda que os casos de apagamento de /d/, mesmo que timidamente, sejam documentados em outras classes de palavras.

Ferreira (2010) salienta que o processo fonológico é bloqueado para palavras de outras classes e para verbos terminados na sequência /Ndo/ que não estejam no gerúndio, como *vendo* e *mando*, formas dos verbos “vender” e “mandar” na primeira pessoa do singular do presente do indicativo. A Figura 3, a seguir, representa o bloqueio da regra de apagamento em palavras que não são formas verbais de gerúndio. Nesses vocábulos a regra não se aplica, embora o ambiente fonológico seja propício.

Figura 3. Bloqueio da regra de apagamento de /d/ em ambientes não derivados

[lindo]	[quando]	[mandar]	
[lĩndo]	[kuãndo]	[mand]o]	sufixação
*[lĩno]	*[kwãno]	[mãnd]o]	nasalização
*[lĩnu]	*[kwãnu]	*[mãno]	apagamento do /d/
		*[mãnu]	representação fonética
	Nível α	Nível β	

Fonte: Ferreira (2010, p. 91).

De acordo com o esquema apresentado na Figura 3, a regra de apagamento é bloqueada em *lindo*, *quando* e em *mando*. Segundo Ferreira (2010), isso acontece porque a oclusiva /d/ nesses três casos faz parte da raiz da palavra. A regra, portanto, não se aplica à sequência /Ndo/ como um todo, mas sim ao morfema *-ndo*, indicador de gerúndio no PB. Com base nessas informações, a autora chegou às seguintes conclusões:

1. O apagamento de /d/ na sequência /Ndo/ não se aplica no nível α e nem a todos os itens do nível β do léxico, uma vez que a regra de apagamento é bloqueada pelo *princípio da preservação da estrutura*;
2. A regra de apagamento é condicionada pelo *princípio do ciclo estrito*, uma vez que ela só se aplica no nível β , sendo bloqueada em ambientes derivados;
3. O fenômeno de apagamento de /d/ em /Ndo/ é uma regra de natureza variável, embora localizada no nível β do léxico.

A autora argumenta que o *princípio da preservação da estrutura* impede que a regra de apagamento de /d/ na sequência /Ndo/ se aplique em raiz de palavras, do contrário, isso causaria uma mudança estrutural no léxico da língua. Mesmo no nível β , domínio da regra, a oclusiva /d/ não é apagada na sequência /Ndo/ quando /d/ faz parte da raiz do verbo, como em *mando* e *vendo*, por exemplo, formas dos verbos mandar e vender no presente do indicativo.

O ambiente de aplicação da regra é restrito ao nível β , e, dentro deste, a aplicação da regra é restrita ao morfema *-ndo* de verbos no gerúndio.

A regra de apagamento não atinge todas as formas de gerúndio, configurando uma regra de natureza variável, o que constitui um problema teórico para a análise, visto que a teoria da Fonologia Lexical propõe que as regras lexicais são de natureza categórica, cabendo a variação ao pós-léxico. Ferreira (2010), contudo, define a regra de apagamento de /d/ no morfema *-ndo* como uma regra lexical de natureza variável.

Embora a análise de Ferreira (2010) seja específica para a amostra de São José do Rio Preto-SP, em que não houve ocorrência de apagamento da oclusiva /d/ em /Ndo/ em ambientes que não constituíssem morfema de gerúndio, a análise fonológica empreendida pela autora é bastante explicativa do fenômeno e se coaduna com Cristóvão Silva (1996), quando esta autora propõe que o processo fonológico requer informação dada pelo componente fonológico (sequência de consoantes /Nd/) e morfológico (morfema de gerúndio), e que não há reorganização lexical (não há indícios de mudança estrutural no léxico).

Deve-se ainda considerar, contudo, o fato de que algumas pesquisas têm indicado que, em algumas variedades do PB, o apagamento da oclusiva /d/ na sequência /Ndo/ pode ocorrer em ambiente que não seja o morfema *-ndo*, de verbos no gerúndio. Pesquisando a fala da cidade de São Paulo, por exemplo, Gonçalves (2018) documentou casos de apagamento de /d/ na conjunção “quando”, o mesmo já havia sido documentado por Vieira (2011) ao estudar a fala de Taboco, distrito de Corguinho-MS. Os registros da forma *quano* para “quando” documentados por Vieira (2011) e Gonçalves (2018) podem indicar o fato de que a regra pode se aplicar, ainda que timidamente, a outras classes de palavras.

Este capítulo objetivou apresentar a forma verbal de gerúndio na língua portuguesa. De tal modo, abordou-se o uso da forma verbal no latim e no português, apresentaram-se estudos que abordam o fenômeno variável de apagamento de /d/ no morfema *-ndo*, indicador de gerúndio e, por fim, apresentou-se uma análise que aborda o fenômeno variável do ponto de vista morfofonológico, com base na Fonologia Lexical. O capítulo seguinte apresenta o referencial teórico que sustenta a pesquisa.

2 LÍNGUA, SISTEMA, VARIAÇÃO E MUDANÇA

Esta pesquisa se fundamentou nos pressupostos teóricos da Sociolinguística estabelecidos por Weinreich, Labov e Herzog (2006 [1968]) e Labov (2008 [1972]), que concebem a língua como um sistema heterogêneo e ordenado que está em constante variação e mudança. Dentro desse modelo teórico, a variação e a mudança linguística são intrínsecas ao sistema; em outras palavras, variabilidade e mutabilidade são sistemáticas e fazem parte da competência linguística do falante.

Neste capítulo são apresentados os conceitos teóricos da Sociolinguística, que deram base à pesquisa sobre o apagamento de /d/ no morfema de verbos no gerúndio na fala de Feira de Santana-BA. Traça-se um breve relato dos caminhos trilhados pela Linguística moderna, desde sua ascensão ao status de ciência, alcançado a partir das ideias de Ferdinand de Saussure, até o nascimento da Sociolinguística e da corrente Variacionista, com as ideias de William Labov. São apresentados ainda os principais conceitos básicos no campo da Sociolinguística Variacionista que deram base para a interpretação do fenômeno pesquisado e que são de suma importância para a compreensão da pesquisa apresentada nesta dissertação.

2.1 DA LINGUÍSTICA À SOCIOLINGUÍSTICA

O status de ciência foi alcançado pela Linguística a partir das ideias publicadas no *Curso de linguística geral*, organizado e publicado em 1916, por Charles Bally e Albert Schehaye, a partir de anotações feitas por eles durante três cursos ministrados pelo linguista suíço Ferdinand de Saussure entre 1907 e 1911. O livro, que lançou as bases do pensamento estruturalista, foi publicado três anos após de Saussure.

Para Saussure (1995 [1916]), a língua é um sistema estruturado e autossuficiente no qual os elementos que o compõem estão em estrita relação de oposição e interdependência. Tais elementos obedecem a certos princípios de funcionamento que são regulados pelo próprio sistema e formam um todo coeso e coerente. Cabe à Linguística como ciência, na visão de Saussure, descrever como esse complexo sistema se estrutura; daí a teoria ser chamada de Estruturalismo.

Ao definir o objeto de estudo da Linguística, Saussure (1995 [1916]) estabelece algumas dicotomias para melhor delimitá-lo. Dentre elas destacam-se *langue* (língua) x *parole* (fala) e sincronia x diacronia. Ao separar a *langue* da *parole*, o autor definiu a primeira como o verdadeiro objeto de estudo da Linguística. Nas palavras dele, a *langue* é essencialmente

social, um sistema homogêneo de signos, um código comum virtual depositado na mente dos falantes de uma língua. A *parole*, por sua vez, é o uso individual – a materialização da *langue*. Por ser de natureza individual, externa, fragmentada, heterogênea e assistemática, a *parole*, na concepção de Saussure, não era passível de descrição, visto que é fluida, ficando, pois, fora da delimitação do objeto.

Quanto à dicotomia sincronia *x* diacronia, o autor estabeleceu que o estudo da língua poderia ser feito em dois eixos. No eixo da sincronia, pode-se observar a língua em um estado curto de tempo, em que as mudanças são mínimas – um recorte que pode ser feito no presente ou no passado, mais ou menos como uma fotografia de um estado do sistema. No eixo da diacronia, como era comum nos estudos históricos desenvolvidos pelos neogramáticos, comparam-se duas sincronias a fim de apreender a mudança no tempo. Para Saussure (1995 [1916]), é o eixo da sincronia que interessa à Linguística moderna, visto que descrever o sistema significa, do ponto de vista do autor, compreendê-lo em seu uso.

Em sua famosa metáfora do xadrez, Saussure (1995 [1916]), ao comparar o sistema linguístico com o jogo, argumentou que em nada adiantaria aos jogadores da partida saber qual a posição anterior da peça para a próxima jogada, muito menos importaria a substância da peça – se de madeira, porcelana ou marfim –, mas sim o seu valor simbólico. Essa metáfora aponta para a prevalência da sincronia sobre a diacronia no estruturalismo saussuriano.

Tendo assentado as bases da Linguística moderna e definido seu objeto de estudo, Saussure (1995 [1916]) compreendeu que a variação estava na *parole* e não na *langue*, e tal como num jogo de xadrez, Saussure compreendeu que não interessava ao falante conhecer os estados anteriores do sistema, visto que em nada tal conhecimento ajudaria para o uso do estado atual desse mesmo sistema. Ficaram de fora da teoria saussuriana a variação, observável por meio da *parole*, e a mudança linguística, observável no eixo diacrônico.

A teoria de Ferdinand de Saussure, nas palavras de Camacho,

Deu sustentação à interpretação da linguagem como um objeto sincrônico em si mesmo e por si mesmo. Essa sustentação representou, na realidade, um gesto de criação que propiciaria a construção de um estatuto de autonomia para a linguística no conjunto das ciências humanas. (CAMACHO, 2013, p. 25)

Segundo Camacho (2013), o estruturalismo além de conferir à linguística o status de ciência, traz em seus postulados um novo olhar sobre as diferenças socioculturais que subjazem as línguas. Se segundo a teoria de Saussure a língua é um sistema plenamente

estruturado, não há lugar dentro da teoria para se acreditar que haja línguas mais ou menos desenvolvidas. As questões relativas à superioridade de uma língua qualquer sobre outra não tem sustento científico no Estruturalismo.

Embora Saussure tenha reconhecido o caráter social da língua ao estabelecer as bases teóricas da Linguística moderna, quando afirma que a língua é “um tesouro depositado pela prática da fala em todos os indivíduos pertencentes à mesma comunidade” (1995 [1916], p. 21), parece não ter havido lugar na teoria para os fatos sociais referentes à língua, já que esta deveria ser estudada em si mesma.

Os postulados de Saussure tomaram força, inclusive, em novos enfoques sobre o estudo da língua, a exemplo da teoria da gramática gerativa, de Noam Chomsky. Não houve avanços, na teoria de Chomsky, na direção de superar a tradição da dicotomia *langue x parole*, mantendo-se a ideia de sistema imanente e autossuficiente frente ao caráter social da linguagem. Quanto a isso, Camacho (2013) argumenta que

Quando no final da década de 1950, surgiu Chomsky (1956) com seu *Syntactic Structures*, o recorte se manteve sob outra denominação e sob nova direção teórica. Com efeito, o sistema linguístico se enquadrou na moldura do conhecimento intuitivo do falante-ouvinte, um objeto de natureza psicológica ou cognitiva, denominada *competência*, com o descarte simultâneo dos atos de fala, infinitamente variáveis e variados, que, relegados ao conceito de desempenho, ficaram destituídos de qualquer natureza teórico-metodológica. (CAMACHO, 2013, p. 34)

Como a teoria da gramática gerativa não se propunha a lidar com as relações entre língua e sociedade, não superando, portanto o divórcio entre o estado homogêneo do sistema, de natureza imanente, e o estado “caótico” da fala, de natureza social, outros pesquisadores procuraram sanar essa dicotomia propondo uma ciência linguística que compreendesse os fatos da língua a partir de uma relação mais próxima entre o seu sistema estruturado interno e o contexto social de uso da língua bem como os diversos usos dessa mesma língua. A busca pela relação entre língua e sociedade deu origem à corrente teórica denominada de Sociolinguística.

Para a Sociolinguística, há uma relação bastante estreita entre língua e sociedade, de maneira que não há como conceber aquela separada desta. Toda e qualquer sociedade é composta de grupos que compartilham de maneiras diferenciadas bens e serviços essenciais. Dividida em variedades, umas mais prestigiadas e outras nem tanto, a língua é, de certa maneira, um espelho que reflete os arranjos sociais das diversas sociedades que delas fazem uso.

Mollica (2004, p. 9) define a Sociolinguística como a “ciência que se faz presente num espaço interdisciplinar, na fronteira entre língua e sociedade, focalizando precipuamente os empregos linguísticos concretos, em especial os de caráter heterogêneo”. Segundo a autora, a Sociolinguística é, de fato, um dos ramos da linguística, seu objetivo é estudar a língua em uso e seu método de investigação lança mão de aspectos linguísticos e sociais, o que é ratificado por Oliveira (2006), ao afirmar que,

[...] sendo a língua um meio de interação entre o indivíduo e a sociedade em que ele atua, está claro que há forças sociais que agem sobre essa interação. É esse relacionamento casual entre língua e sociedade que constitui o objeto de análise da Sociolinguística. (OLIVEIRA, 2006, p. 45)

A Sociolinguística surge como um ramo da Linguística geral que oferece ao pesquisador aporte teórico para que se possam incluir em seu método de análise, além de fatores linguísticos internos ao sistema, fatores de ordem social que fazem com que a língua seja o que ela é e como ela é. A Sociolinguística concebe a língua como uma instituição social, falada por sujeitos reais, localizados num tempo e num espaço específicos. O objetivo primordial da Sociolinguística é descrever como as formas linguísticas se relacionam com os distintos grupos sociais que as usam, e como esses grupos são percebidos por meio dessas manifestações linguísticas.

Desde a instituição da Sociolinguística como ramo de estudo da linguagem, oficialmente em 1964, em um congresso organizado por William Bright, na Universidade da Califórnia, nos EUA, essa disciplina da Linguística tem auxiliado os linguistas a compreender e explicar as diversas relações existentes entre língua e sociedade. Nas palavras de Camacho (2013), desde o surgimento da Sociolinguística, o interesse pelo estudo da língua no contexto social de uso tem sido crescente. O autor argumenta que “os diversos enfoques que se abriram sob o rótulo de sociolinguística vem cobrindo, desde o início, uma grande variedade de assuntos” (CAMACHO, 2-13, p. 35).

Ainda de acordo com Camacho (2013), a sociolinguística compreende, portanto, a três grandes ramos de estudo, são eles a Sociologia da Linguagem, desenvolvida por Joshua Fishman e que lida com questões relacionadas a políticas linguísticas; a Etnografia da Comunicação, desenvolvida por Dell Hymes, com foco em análise de eventos de fala e que desemboca na Análise da Conversação e na Sociolinguística Interacional; e a Sociolinguística Variacionista, desenvolvida por William Labov, que trata de problemas teóricos linguísticos a partir do estudo da língua em seu contexto social de uso.

É dentro da corrente da Sociolinguística denominada de Sociolinguística Variacionista, desenvolvida por William Labov, que se desenvolve a pesquisa sobre o apagamento da oclusiva /d/ no morfema de gerúndio no português falado em Feira de Santana-BA. Por esse motivo, apresenta-se uma breve introdução sobre esse campo da Sociolinguística na subseção seguinte.

2.1.1 A Sociolinguística Variacionista

A Sociolinguística Variacionista compreende que as línguas humanas têm natureza inerentemente histórica, que as línguas tendem à constante variação e à mudança. Sabe-se que cada estado de uma língua é o resultado de um conjunto de transformações ocorridas ao longo do tempo. As transformações por que passam as línguas humanas são lentas e graduais, se concretizam após um longo período de tempo e não atingem o sistema como um todo de uma só vez. Isso não significa dizer, contudo, que a existência da variação necessariamente implique mudança, embora toda mudança resulte de um estado anterior de variação.

Há formas linguísticas que convivem em variação estável por anos sem que nenhuma delas se torne obsoleta. Segundo Coelho (2010, p. 96), a “variação é uma propriedade inerente ao sistema linguístico, sendo motivada por condicionamentos internos e externos, e o falante tem competência linguística para lidar com regras variáveis”. Mendes (2013) também observa a importância de refletir sobre o fato de que existem duas (ou mais) formas de dizer a mesma coisa e de que deve o linguista procurar descobrir o que leva os falantes a empregarem ora uma, ora outra forma; e, indo além, o autor acrescenta que é necessário compreender qual o significado social da variação linguística.

Para analisar e descrever como a variação e a mudança linguística estão relacionadas com os aspectos socioculturais e históricos de uma comunidade de fala, a Sociolinguística Variacionista traz como proposta estudar e descrever a língua tal como ela acontece dentro da comunidade em que é usada, levando em conta os fatores sócio-históricos e culturais de tal comunidade e considerando que tais fatores extralinguísticos agem junto a fatores estruturais para que ocorram a variação e a mudança nas línguas. Ao propor os fundamentos para a abordagem variacionista dos fatos da língua, Weinreich, Labov e Herzog (2006 [1968]) afirmam que

[...] fatores linguísticos e sociais estão intimamente inter-relacionados no desenvolvimento da mudança linguística. Explicações confinadas a um ou outro 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 [1968], p. 126)

Ao apresentar a sua conhecida pesquisa sobre a mudança sonora em Martha's Vineyard, dentro do escopo da Sociolinguística Variacionista, Labov (2008 [1972]) reafirma que

[...] não se pode entender o desenvolvimento de uma mudança linguística sem levar em conta a vida social da comunidade em que ela ocorre. Ou, dizendo de outro modo, as pressões sociais estão operando continuamente sobre a língua, não de algum ponto remoto no passado, mas como uma força social imanente agindo no presente vivo. (LABOV, 2008 [1972], p. 21)

Dentro desses preceitos, Labov (2008 [1972]) propõe uma ciência linguística que estabelece uma relação indissociável entre sistema e estrutura social, sem, contudo, preconizar que se dê demasiado peso a aspectos sociais em detrimento de aspectos estruturais. Discutindo a grande importância que tem a análise estrutural para a descrição do sistema linguístico, Labov (2008 [1972]) chama a atenção para a grande contribuição que tem dado o Gerativismo para a compreensão das relações invariantes dentro da estrutura, ainda que despreze fatores sociais, mas afirma parecer claro “que não se pode fazer nenhum avanço importante rumo ao entendimento do mecanismo da mudança linguística sem o estudo sério dos fatores sociais que motivam a evolução linguística” (LABOV, 2008 [1972], p. 291).

A grande virada que sofre a Linguística, com os postulados de Weinreich, Labov e Herzog (2006 [1968]) e Labov (2008 [1972]), está no fato de que a variação, até então vista como algo fora do sistema, como propunha o Estruturalismo saussuriano, passa a ser compreendida como sistemática e diretamente relacionada com a mudança. Labov (2008 [1972] p. 125) assevera que “não existe mudança sem que haja antes variação e heterogeneidade”. Se a mudança é inerente ao sistema e não há mudança sem que haja variação, logo a variação é parte do sistema, sendo este, portanto, heterogêneo.

Ainda sobre a relação entre variação e sistema, Labov (2008 [1972], p. 87) questiona a proposição estruturalista de que a língua precisa ser um sistema estruturado e homogêneo para que possa funcionar eficientemente, pois que o sistema continua funcionando enquanto a estrutura muda. Diante de tal realidade, ele propõe que a variação “não é um produto da mistura dialetal irregular, mas uma propriedade inerente e regular do sistema” (LABOV, 2008 [1972], p. 262). Ao compreender a variação como característica interna do sistema linguístico, a Sociolinguística Variacionista concebe o sistema linguístico como heterogêneo e ordenado.

Weinreich, Labov e Herzog (2006 [1968]) afirmam que variação, mudança e sistema estão em íntima relação e que uma melhor compreensão dos fenômenos de variação e mudança linguística não deve prescindir da compreensão de como fatores linguísticos e sociais operam na evolução das línguas. Os autores formulam cinco questões teóricas, dispostas em forma de problemas, para que se possa dar conta da variação e da mudança linguística dentro da teoria variacionista. São eles: i) o problema dos fatores condicionantes; ii) o problema da transição; iii) o problema do encaixamento; iv) o problema da avaliação; e v) o problema da implementação; que passam a ser apresentados a seguir.

2.1.1.1 O problema dos fatores condicionantes

Weinreich, Labov e Herzog (2006 [1968], p. 121) sugerem que “um possível objetivo para uma teoria da mudança é determinar o conjunto de mudanças possíveis e condições possíveis para a mudança [...]”. Ao definir o problema dos fatores condicionantes gerais, os autores dão dois exemplos observados em mudanças linguísticas e que se mostram agir sempre na mesma direção:

Um desses fatores condicionantes parece se aplicar a áreas em que um sistema de dois fonemas está em contato com um sistema de um fonema fundido: argumentamos que, exceto sob determinadas condições especiais, a direção da mudança será em direção do sistema de um fonema. Observamos, como outro exemplo, diversos casos de cadeias de mudança correlatas em que vogais [tensas] periféricas se elevam, mas nenhuma na direção contaria. (WEINREICH; LABOV; HERZOG, 2006 [1968], p. 121)

O problema dos fatores condicionantes foi, contudo, abandonado por Labov (1982). Lucchesi (2004) argumenta que o problema dos fatores condicionantes formulados por Weinreich, Labov e Herzog (2006 [1968]) levaria a teoria a assumir princípios gerais ou até mesmo universais para a mudança, o que traria sérios problemas à natureza da própria teoria. Em face disso, o próprio Labov revisita mais tarde o problema dos fatores condicionantes e assume que a busca por restrições universais é uma busca por uma faculdade isolada da linguagem, e isso vai de encontro ao que se tem verificado quanto à natureza da linguagem.

O problema dos fatores condicionantes, como proposta de estabelecer direções gerais ou universais para a mudança linguística foi, nas palavras de Lucchesi (2004), um equívoco, e não por acaso foi abandonado mais tarde pelo próprio Labov.

2.1.1.2 O problema da transição

De acordo com o problema transição, Weinreich, Labov e Herzog (2006 [1968]) postulam que a mudança linguística não ocorre de forma discreta entre as gerações. Uma variante não surge e simplesmente substitui outra, mas ambas concorrem entre si até que uma delas se torne preferencial em detrimento da outra.

Faraco (2017, p. 45) afirma que um estado de língua qualquer, seja no passado ou no presente, “é sempre o resultado de um longo e contínuo processo histórico”, de modo que toda mudança prevê um estado de variação anterior, o que evidencia o caráter gradual e contínuo da mudança linguística. Weinreich, Labov e Herzog (2006 [1968]) postulam que a mudança ocorre em três momentos: no primeiro, o falante aprende uma forma alternativa; no segundo, as duas formas coexistem em contato dentro da competência do falante; no terceiro, por fim, uma das formas cai em desuso.

O problema da transição diz respeito, portanto, ao caminho percorrido pela mudança linguística. Nesse sentido, Lucchesi (2004, p. 174) afirma que “quanto mais acurada for a resposta ao problema da transição, maior será a compreensão dos linguistas sobre o processo através do qual a mudança linguística ocorre”. Esse caminho da mudança a ser traçado pelo pesquisador perpassa pelo encaixamento no sistema linguístico e social da comunidade de fala.

2.1.1.3 O problema do encaixamento

O problema do encaixamento busca explicar como determinada mudança linguística está encaixada nas estruturas interna e externa da língua. Ao conceber a língua como um sistema que não é estruturado apenas por fatores internos, como propunha visão estruturalista de sistema imanente e homogêneo, o pesquisador deverá também buscar explicações para o encaixamento da mudança linguística na estrutura social. Weinreich, Labov e Herzog (2006 [1968]) afirmam que

Haverá pouca discordância entre os linguistas de que as mudanças linguísticas sob investigação devem ser vistas como encaixadas no sistema linguístico como um todo. O problema de oferecer fundamentos empíricos sólidos para a teoria da mudança traz à tona diversas questões sobre a natureza e a extensão desse encaixamento. (WEINREICH; LABOV; HERZOG, 2006 [1968], p. 122)

Ao lidar com o problema do encaixamento, o linguista deve, além de demonstrar a motivação social da mudança em estudo “determinar o grau de correlação social que existe e mostrar como ela pesa sobre o sistema linguístico abstrato” (WEINREICH; LABOV; HERZOG, 2006 [1968], p. 123).

Segundo Lucchesi (2004), o problema do encaixamento representa um dos maiores avanços da Sociolinguística, na medida em que seu modelo de análise linguística vai para além da estrutura interna e procura compreender a mudança linguística a partir da covariação de fatores estruturais e sociais.

2.1.1.4 O problema da avaliação

O problema da avaliação busca explicar como a mudança observada pode ser avaliada em termos subjetivos. Estando a mudança linguística associada à mudança no comportamento social, a avaliação da comunidade de fala sobre o fenômeno variável pode reforçar ou inibir o grau de ocorrência das formas variantes.

Weinreich, Labov e Herzog (2006 [1968], p. 124) indicam que “o nível de consciência social é uma propriedade importante da mudança linguística que tem de ser determinada diretamente” e acrescentam que a investigação desses níveis de avaliação subjetiva ajuda a aprofundar a compreensão da implementação ou não da mudança.

Para resolver o problema da avaliação, o sociolinguista deve verificar se o status social da forma variante é positivo ou negativo na comunidade de fala. Dessa forma, Coelho (2010) destaca que

[...] são favoráveis as condições quando a forma é prestigiada na sociedade e desfavoráveis as condições quando a forma é estigmatizada. Ou seja, a avaliação da mudança linguística pelos membros da comunidade de fala se dá pelo seu significado social – prestígio ou não. (COELHO, 2010, p. 103)

O problema da avaliação lida diretamente com os sujeitos da comunidade de fala, como estes avaliam uma nova forma na língua. Desse modo, a mudança linguística adquire uma significação social que vai agir diretamente no processo de acordo com o valor lhe é atribuído, de modo que a mudança será favorecida se seu significado social for positivo ou desfavorecida se o significado atribuído for negativo.

2.1.1.5 O problema da implementação

O problema da implementação busca investigar por que determinada mudança ocorre em determinada comunidade de fala, em determinado lugar, em determinado tempo e não em outros. O processo de implementação da mudança linguística envolve fatores intrínsecos e extrínsecos às línguas – fatores condicionantes – e a compreensão de como esses fatores se encaixam na mudança – o problema do encaixamento.

Weinreich, Labov e Herzog (2006 [1968]) apontam ser provável que as explicações para o processo de implementação de determinada mudança sejam dadas *a posteriori*, posto que a mudança linguística vem sempre como resultado de uma mudança no comportamento social. Sobre esse ponto, Coelho (2010, p. 105) afirma que “à medida que identificamos os condicionamentos que agem sobre a mudança, damos uma explicação sobre a forma como a mudança vai se implementando nos diferentes contextos estruturais e nos diferentes estratos sociais”.

Lucchesi (2004) considera que o problema da implementação constitui uma questão epistemológica, pois diz respeito ao que significa explicar alguma coisa em Sociolinguística, bem como diz respeito à maneira como essa própria corrente da Linguística concebe o seu objeto de estudo. O autor afirma que a forma de explicar a implementação da mudança linguística, em termos de descrever os mecanismos de causa e efeito que constituem determinada mudança, vai de encontro com a própria visão de língua como um construto sócio-histórico e cultural ao qual a relação de causa e efeito não se adéqua.

2.1.1.6 O tempo real e o tempo aparente

Labov (1994) propõe que a mudança linguística pode ser observada a partir da comparação de duas sincronias. Assim, comparando um estado A da língua a um estado B, considerando um espaço de tempo mais ou menos longo entre A e B, pode-se ter indícios de como certo fenômeno tem se comportado ao longo desse tempo. Esse tipo de procedimento de estudo é chamado tempo real. O estudo em tempo real pode ser feito seguindo dois modelos, são eles o estudo de tendência e estudo de painel. No estudo de tendência, a amostra é coletada de maneira randomizada no período B, sendo que nesta ocasião apenas a comunidade de fala é a mesma, os informantes não. No estudo de painel, os informantes da amostra B devem ser exatamente aqueles contatados durante as gravações da amostra A.

O estudo em tempo real oferece algumas dificuldades para o pesquisador, a primeira é o lapso temporal necessário para que se possa perceber a mudança; a segunda é a disponibilidade de dados, pois nem sempre se encontram dados de fala de sincronias passadas para que se possa proceder ao estudo em tempo real. Descrever o status da variação e a possível mudança na realização das formas de gerúndio na fala de Feira de Santana-BA, por exemplo, ainda não seria possível, pois a amostra do projeto A Língua Portuguesa no Semiárido Baiano, a partir da qual se constituiu o *corpus* da pesquisa apresentada nesta dissertação, ainda é recente (as últimas gravações foram feitas no ano de 2010) e não há outra mais antiga que permita o estudo diacrônico do fenômeno. Para solucionar problemas metodológicos como esse, Labov (1994) propõe a noção de tempo aparente.

O estudo da variação/mudança no tempo aparente busca aporte teórico na hipótese clássica de aquisição da linguagem que, conforme Lenneberg (1967), considera que a aprendizagem da língua se concretiza até o final da puberdade, e consiste em coletar dados de uma amostra sincrônica X e comparar a frequência da variante em estudo nas diversas faixas etárias. Sabe-se que a mudança linguística não é discreta (WEINREICH; LABOV; HERZOG, 2006 [1968]). E, por isso mesmo, formas em variação convivem no dia a dia da comunidade de fala. Segundo Oliveira (2006),

A concepção e o alcance do modelo sociolinguístico são, ao mesmo tempo, sincrônicos e diacrônicos, uma vez que toda mudança implica uma variação. Nesse modelo, a estrutura de uma língua somente será totalmente entendida à medida que se compreendam efetivamente os processos históricos de sua configuração. (OLIVEIRA, 2006, p. 49)

Naro (2013) afirma que,

A mudança linguística não é absolutamente mecânica e regular a curto prazo. Em qualquer estado da língua costumam coexistir formas de diversos estágios de evolução [...] existem pessoas que, apesar de estarem em interação constante (do tipo pai/filho), costumam falar de maneira distinta. Entretanto isso não chega a comprometer a compreensão, já que ambos os lados são capazes de utilizar e entender todas as formas. (NARO, 2013, p. 43)

Para Naro (2013), o conceito de tempo aparente, sustentado na hipótese clássica de que a aquisição da língua se completa durante a puberdade, faz com que seja possível projetar, a partir de um recorte sincrônico, estados futuros da língua com base na faixa etária dos

falantes. Essa projeção diacrônica dentro da sincronia é o que se chama em sociolinguística de tempo aparente.

Segundo Oliveira (2006),

O tempo aparente serve como um procedimento através do qual se pode projetar o tempo real futuro, ou seja, estudando-se a transição e a implementação de variáveis – uma análise acurada de uma comunidade sociolinguística com seu componente de variação e de mudança –, pode-se dimensionar historicamente uma variável, não só em relação ao passado mas também em relação a um possível comportamento no futuro. (OLIVEIRA, 2006, p. 49)

A concepção de tempo aparente em Sociolinguística representa uma grande contribuição para a Linguística Histórica, uma vez que permite ao pesquisador, para além de estudar e descrever como as formas variantes estão em variação no sistema, identificar mudanças linguísticas em progresso. Sobre isso, Oliveira (2006) afirma que,

Analisando a correlação entre as variantes e os fatores internos e considerando o fator idade, pode-se observar a estratificação das mesmas. Se, por exemplo, a variante inovadora for mais frequente entre falantes mais jovens, decrescendo em relação à idade dos outros informantes, tem-se indícios de uma mudança em progresso. (OLIVEIRA, 2006, p. 49)

A noção de tempo aparente é um recurso metodológico que permite ao pesquisador inferir, com segurança, o status de uma variável linguística no sistema da comunidade de fala quando não se dispõe de dados diacrônicos. Contudo Labov (2008[1972]) afirma que, tendo captado indícios de mudança em tempo aparente, um estudo em tempo real deve ser feito para que se tenha a confirmação dos resultados alcançados em tempo aparente.

2.1.1.7. O estudo de Labov em Martha's Vineyard

Em 1963, William Labov defendeu a sua dissertação de Mestrado, resultado de uma pesquisa realizada na ilha de Martha's Vineyard, no município de Dukes, litoral de Massachusetts, nos Estados Unidos. Nesse trabalho, o autor realizou uma pesquisa em que relacionou condicionadores extralinguísticos como sexo, idade, profissão e atitudes dos falantes em relação ao espaço geográfico para a compreensão do fenômeno de centralização dos ditongos /ay/ e /aw/ na variedade da língua inglesa falada pelos vineyardenses.

Essa pesquisa em Martha's Vineyard lançou um novo modelo de análise linguística e tornou-se um clássico, sendo reconhecido como o marco do surgimento da Sociolinguística

Variacionista. Por ser de fundamental importância para a Sociolinguística, na medida em que o método criado pelo autor vem até hoje sendo adaptado e usado para a compreensão de fenômenos linguísticos, apresenta-se aqui um breve relato desse estudo clássico, com vistas ao aprofundamento no método de pesquisa em Sociolinguística.

Costuma-se referir à pesquisa feita por Labov sobre a centralização dos ditongos /ay/ e /aw/ em Martha's Vineyard como o primeiro estudo sociolinguístico de um dialeto. Em 1961, quando começou a pesquisa na ilha, William Labov era aluno da Columbia University, orientando do professor Uriel Weinreich. Embora naquela época os estudos dialetológicos estivessem inclinados a documentar e descrever diversidades dialetais, Labov estava mesmo interessado na relação entre as formas variantes da língua com fatores de ordem social.

De acordo com o censo de 1960, a ilha de Martha's Vineyard tinha pouco mais de 5.500 habitantes à época em que o autor desenvolveu sua pesquisa. Mas durante o verão, mais precisamente nos meses de junho e julho, a ilha se tornava um dos locais prediletos para turismo e a quantidade de turistas alcançava o número de 42.000. Esse número elevado de veranistas bem como a dependência que a ilha tinha da economia turística exerciam grande efeito sobre a mudança linguística que Labov abordou, como se descreve a seguir.

Em suas incursões na comunidade da ilha, chamou a atenção do pesquisador o fato de alguns moradores dali realizarem de maneiras diferentes as vogais nucleares nos ditongos /ay/, em palavras do tipo *ice* (gelo) e *time* (tempo), que em vez de [ay], como falavam as pessoas do continente, era realizado [əy], como um *shwa*. Embora essa singularidade já tivesse sido registrada no LANE (Atlas Linguístico da Nova Inglaterra) de 1951, uma vez que quatro moradores da ilha foram documentados na época, Labov notou que, diferentemente dos falantes do continente, as pessoas da ilha usavam as duas formas. Mais interessante ainda, ele notou que as formas variavam não só de pessoa para pessoa mas também na fala de uma mesma pessoa. O mesmo fenômeno ocorria com o ditongo /aw/, em palavras do tipo *South* (Sul) e *loud* (barulhento).

Diante daquela realidade, Labov procurou descobrir se a variação observada derivava de diferenças diatópicas ou se havia outros aspectos envolvidos. Para proceder à sua pesquisa, o autor desenvolveu uma metodologia inovadora em estudos dialetológicos. Em vez de documentar o que as pessoas diziam, escrevendo as respostas que elas davam às perguntas previamente formuladas, ele decidiu gravá-las; primeiro pedindo a elas que lessem um texto e depois em entrevistas em que os conduzia a falar de assuntos que os deixassem à vontade, a fim de chegar o mais próximo possível do vernáculo. Labov procurou, ainda, representar em

sua amostra de fala dados de natureza social, como a ascendência do informante (a ilha era habitada por descendentes de portugueses, ingleses e indígenas), a profissão, o sexo e a idade.

Labov conseguiu enquadrar o fenômeno de centralização dos ditongos na estrutura interna da língua. O ditongo /ay/, por exemplo, tinha como ambiente motivador as consoantes desvozeadas (/t, s, p, f/) e como ambiente desmotivador as consoantes vozeadas (/l, r, n, m/). Mas para além desses condicionadores linguísticos, ele notou que havia algumas relações entre a centralização dos ditongos /ay/ e /aw/ com alguns fatores sociais. Por exemplo, os moradores de uma região da ilha centralizavam mais que os de outra, os pescadores centralizavam mais que qualquer outra classe trabalhista, os falantes entre 31 e 45 anos centralizavam mais que os de outra faixa etária.

Como percebeu também durante suas entrevistas que os moradores tinham atitudes diferentes sobre a ilha – alguns eram positivos sobre continuar morando lá e outros não –, Labov decidiu observar se havia relação entre a atitude do falante para com a ilha e a centralização dos ditongos. Os resultados indicaram que, quando os falantes demonstravam atitudes positivas sobre a ilha, a ocorrência de centralização de /ay/ e /aw/ era maior, sendo a ocorrência baixa ou nenhuma entre os que demonstravam não gostar de viver ali.

Os achados de Labov para a centralização de dos ditongos /ay/ e /aw/ resumem-se nos seguintes pontos:

- Os falantes que centralizavam mais eram moradores de uma região específica da ilha;
- Por opção ou tradição, os falantes que mais centralizavam viviam da pesca, que era a maneira tradicional de vida da ilha;
- Os falantes que mais centralizavam tinham entre 31 e 45 anos;
- Os falantes que mais centralizavam gostavam de viver na ilha e tinham orgulho de serem nativos dali.

A partir desses dados e, é claro, de uma compreensão profunda das relações histórico-sociais, culturais e linguísticas da ilha, Labov elaborou um esquema para explicar a aparente mudança linguística que ora acontecia em Martha's Vineyard com relação aos ditongos /ay/ e /aw/. Levando em conta padrões linguísticos e sociais, o pesquisador mostrou que a variação não era livre e não-condicionada, mas que esta obedecia a uma regra interna que se relacionava com uma oposição social feita entre ser ilhéu e ser turista, entre ser ilhéu e gostar da vida na ilha e ser ilhéu e não se identificar com a vida levada no local.

Como se pôde ver a partir do relato da pesquisa realizada por Labov em Martha's Vineyard, em 1963, seria difícil chegar a alguma conclusão sobre o fenômeno linguístico estudado na ilha caso o pesquisador não tivesse considerado os fatores socioculturais e históricos da comunidade de fala. Labov apenas chegou a uma compreensão satisfatória sobre a centralização dos /ay/ e /aw/ encaixando a mudança linguística nas estruturas interna e externa ao sistema, mostrando que língua e sociedade são indissociáveis.

2.2 AS DIMENSÕES INTERNA E EXTERNA DA VARIAÇÃO LINGUÍSTICA

A concepção do modelo que concebe a língua como sistema heterogêneo e ordenado se propõe a lidar com a variação e a mudança linguística nas dimensões interna e externa ao sistema linguístico. Segundo Coelho (2007),

A variação pode ser influenciada por fatores linguísticos ou extralinguísticos, como origem geográfica, status socioeconômico, grau de escolarização, idade, sexo, mercado de trabalho, redes sociais do falante. Além, é claro, do estilo pessoal, porque cada indivíduo é único e possui uma maneira única de falar. (COELHO, 2007, p. 10)

Compreende-se que, em sua dimensão interna, a língua pode sofrer variação e mudança nos níveis lexical, fonológico, sintático e semântico. O fenômeno de variação pesquisado nesta dissertação, o apagamento da oclusiva /d/ no morfema de verbos no gerúndio, encontra-se na interface de dois níveis da gramática, a fonologia e a morfologia (CRISTÓFARO SILVA, 1996; FERREIRA e TENANI, 2009; FERREIRA, 2010), sendo classificado como um fenômeno de variação morfofonológica. Isso porque o apagamento da oclusiva /d/ no morfema *-ndo* é um fenômeno fonológico que ocorre num ambiente morfológico específico, a desinência verbal de gerúndio².

A compreensão satisfatória de formas em variação/mudança dentro de um modelo sociolinguístico de análise prevê não só descrever como o fenômeno variável se encaixa na estrutura interna do sistema, mas procura para, além daí, compreender como fatores externos condicionam o fenômeno variável.

Na dimensão externa, a variação pode ocorrer levando em conta aspectos regionais e sociais. Coelho (2010) afirma que

² Conforme estudos apresentados no capítulo 1, seção 1.2 deste texto, embora o apagamento seja mais recorrente no morfema de verbos no gerúndio, ele ocorre também em outras classes de palavras. Cf., por exemplo, Mollica e Mattos (1992).

O caráter heterogêneo do sistema linguístico é produto, portanto, de duas ou mais formas em variação – duas ou mais variantes – que se alternam de acordo com condicionadores linguísticos e extralinguísticos que motivam ou restringem a variação. (COELHO et al, 2010, p. 47)

Encaixar a variação na estrutura externa ao sistema linguístico significa relacionar o fenômeno em análise com fatores sociais que podem condicioná-lo. Os fatores externos ou sociais comumente controlados em pesquisas sociolinguísticas no Brasil são aqueles referentes à região, a características sociais como sexo, faixa etária, nível de escolarização e nível socioeconômico dos falantes. Algumas pesquisas controlam também os contextos de produção da fala ou os meios pelos quais a língua se manifesta, como pesquisas que estudam a variação entre as formas falada e escrita da língua bem como as influências de uma forma sobre a outra.

Os condicionadores externos da variação linguística podem ser representados em quatro grandes grupos: variação diatópica, variação diastrática, variação estilística e variação diamésica. Apresenta-se a seguir uma breve explanação de cada um desses quatro tipos de variação, uma vez que eles constituem conceitos chave para a compressão da variação linguística.

2.2.1 A variação diatópica

A variação diatópica se dá entre regiões geográficas e é também conhecida como variação regional. Pode ser notada entre países, como é o caso de Brasil e Portugal, entre regiões dentro do mesmo território linguístico, como ocorre entre o Nordeste e o Norte do Brasil; entre estados, como Bahia e Minas Gerais; ou cidades, como Feira de Santana e Salvador, por exemplo.

A variação diatópica é resultado de traços culturais característicos de cada comunidade de fala, traços esses que são construídos a partir da identificação dos falantes com a localização geográfica da comunidade. Os usuários da língua imprimem sua cultura no seu modo de falar como meio de afirmar sua identidade frente a outras comunidades linguísticas; como resultado disso, surgem variações de ordem fonética, lexical e até sintática que fazem com que surja a variação diatópica dentro da mesma língua. Segundo Marinho e Val (2006),

Os limites geográficos de uma comunidade, no entanto, não são necessariamente coincidentes com os limites políticos de um estado ou de

uma região. Eles também não são nítidos e precisos, mas graduais, na medida em que se encontram certas áreas de maior ou menor concentração de determinadas características pertencentes a determinados dialetos. (MARINHO; VAL, 2006, p. 25-26)

A variação diatópica constitui um tipo de variação linguística que é facilmente percebida pelo falante e por meio dela é possível, às vezes, inferir a origem de uma pessoa, seja por características fonéticas, pelo uso de itens lexicais específicos ou até estruturas sintáticas diferentes. A realização do gerúndio no Brasil e em Portugal constitui um bom exemplo de variação sintática: enquanto no PB usa-se a forma terminada em *-ndo*, os portugueses usam cada vez mais o infinitivo preposicionado, como foi demonstrado por Mothé (2006) ao comparar as variedades brasileira e europeia da língua portuguesa.

Dentro do território nacional brasileiro, vários são os tipos de variação fonológica marcados diatopicamente, um bom exemplo é a realização das vogais pretônicas /E/ e /O/ nos dialetos das regiões Sul e Nordeste do país, que são realizadas de maneira predominantemente fechada por sulistas e abertas por nordestinos, conforme afirma Chaves (2014).

A variação diatópica é usada muitas vezes para criar e reforçar preconceitos linguísticos e sociais. A história do desenvolvimento socioeconômico do Brasil se deu de maneira desigual e excludente, fazendo com que a maior concentração das riquezas bem como os grandes centros industriais e culturais fossem concentrados no eixo Sul-Sudeste do país. Esse modelo de desenvolvimento baseado no privilégio de uns e na marginalização de outros criou enormes gargalos entre as regiões do país. Essas diferenças socioeconômicas são projetadas na língua, fazendo com que algumas pessoas julguem ser melhores as variedades faladas por pessoas de regiões mais ricas do país e que têm acesso a bens culturalmente legitimados como de prestígio e, por outro lado, inferiorizem as variedades faladas por pessoas que moram em regiões desprivilegiadas do ponto de vista econômico e que não tenham acesso a bens culturais tidos como de prestígio.

É comum, por exemplo, em novelas brasileiras, ver a prosódia da fala de Salvador ser relacionada à preguiça. Parece ter lugar comum na produção novelística brasileira a visão de que o baiano fala “devagar” porque é preguiçoso. Esse tipo de julgamento não tem fundamento científico, encontrado base tão somente em atitudes preconceituosas e na falta de conhecimento científico sobre os fatos da língua.

A variação diatópica é uma característica natural das línguas, pois cada grupo regional tem seus costumes, suas crenças e modos de perceber e lidar com a realidade que lhe tornam único, definindo sua identidade em relação a outros grupos regionais. A forma de usar

a língua também faz parte da identidade regional de um grupo. Do ponto de vista científico, isso não faz com que essa forma de uso seja melhor ou pior em relação a qualquer outra, mas apenas diferente.

2.2.2 A variação diastrática

Também conhecida como variação social, a variação diastrática se dá entre os vários estratos sociais de uma comunidade. A variação diastrática está ligada a fatores como sexo, faixa etária, nível de escolaridade, nível socioeconômico, entre outros.

No tocante ao fator ‘sexo’, pesquisas têm mostrado que homens e mulheres apresentam atitudes diferentes quanto ao uso da língua no dia a dia, sendo que as mulheres, nas comunidades ocidentais, são, no geral, mais conservadoras diante de formas linguísticas inovadoras, sobretudo se se trata de formas linguísticas estigmatizadas, embora isso não seja regra geral. Por outro lado, os homens são mais propensos a formas novas na língua, e em muitas vezes eles são responsáveis por liderar processos de variação e mudanças, sobretudo quando a forma variante é avaliada como de baixo prestígio.

A língua varia também de acordo com a faixa etária dos falantes, sendo que as pessoas mais idosas conservam formas mais antigas enquanto as mais jovens são responsáveis por espalhar formas variantes novas na língua. Essa variação entre faixas etárias é também conhecida como variação diageracional.

O nível de escolaridade também está intimamente ligado à variação diastrática. Embora a Constituição Imperial 1824 já faça alusão ao direito à educação primária a todos os cidadãos, isso sem entrar no mérito da discussão de quem poderia ser cidadão naquela época, apenas em 1961 o Brasil teve aprovada uma lei que estabelecesse as diretrizes e bases da educação no território nacional. O acesso à educação formal no Brasil ainda hoje está longe de ser efetivamente um direito de todos, sendo que o acesso aos níveis mais elevados de formação esteve sempre restrito às classes sociais mais abastadas. O fato de o acesso à educação de qualidade ter sido historicamente restrito a uma pequena parcela da população faz com que os usos linguísticos feitos por esses falantes gozem de um certo prestígio por serem geralmente mais próximos da variedade padrão, prescrita na gramática normativa, distanciando-se dos usos comuns feitos pela grande maioria dos falantes.

Por fim, o nível socioeconômico dos falantes é também um motivador para a ocorrência da variação diastrática, visto que, quanto mais alta a renda de um indivíduo, maiores serão as possibilidades de ele acessar o capital cultural disponível na sociedade,

inclusive por meio do acesso aos mais elevados graus de escolarização, que lhe confere o domínio da variedade culta da língua, como já foi salientado acima. Essas pessoas são no geral mais conservadoras quanto a processos de variação e mudança linguística, já que fazem uso da variedade de maior prestígio, e veem, quase sempre, com maus olhos formas variantes que fogem do padrão gramatical normativo.

2.2.3 A variação estilística

A variação estilística está relacionada aos vários papéis sociais desempenhados pelo falante, ao contexto de comunicação e ao assunto tratado. Enquanto a variação diatópica e a variação diastrática acontecem no âmbito da comunidade de fala, a variação estilística é de caráter mais individual. Uma mesma pessoa pode usar formas linguísticas diferentes a depender de sua posição social no momento do enunciado. É possível, por exemplo, que um supervisor de uma certa empresa se dirija a um subordinado seu de uma maneira e a seu gerente superior de outra totalmente distinta, pois entram em jogo nesses contextos comunicacionais questões de poder. A fala de um professor na sala de aula com seus alunos é, com toda a certeza, diferente da fala desse mesmo professor com seus familiares em casa ou com seus amigos mais próximos, pois o ambiente escolar requer um estilo de fala mais formal e por isso mais monitorado.

O que se tem constatado em estudos sociolinguísticos é que em momentos de fala mais formais os falantes lançam mão de estilos de fala mais próximos do padrão, com alto grau de monitoramento; ao passo em que usam estilos mais coloquiais, menos monitorados, em contextos de fala mais descontraídos. Esses dois extremos do *continuum* da variação estilística são chamados, respectivamente, de registro formal e registro informal.

Esse tipo de variação de registro que ocorre entre situações de mais ou menos monitoramento da fala foi primeiramente explorado por Labov (1966) em um estudo sobre a estratificação social do /R/ na cidade de Nova York. Especificamente sobre o apagamento de /d/ no morfema de gerúndio, tema desta dissertação, Nascimento e Mota (2004) indicam que o percentual de apagamento da oclusiva /d/ no morfema de gerúndio está relacionado ao grau de monitoramento de fala, uma vez que as autoras constataram que, quanto mais o informante se aproximava do vernáculo, maior era o percentual de ocorrência de apagamento da oclusiva /d/ no morfema *-ndo*. As autoras destacaram, inclusive, que em uma das entrevistas que compuseram o *corpus* da pesquisa, certo informante apagou a oclusiva /d/ em todas as formas de gerúndio quando esteve relatando a morte do pai; no entanto esse mesmo informante, ao

recitar um poema que ele mesmo escreveu em homenagem ao pai morto, realizou a oclusiva /d/ em todas as formas de gerúndio.

Percebe-se, a partir do relato feito por Nascimento e Mota (2004) que o falante apresenta um nível de consciência de contextos de usos linguísticos mais ou menos formais e a partir dessa consciência escolhe qual registro se adéqua melhor ao evento discursivo, variando os estilos de fala.

A variação estilística constitui um campo fértil de pesquisa para a Sociolinguística Variacionista, mas convida também o pesquisador a lidar com o que Labov (2008 [1972]) definiu como o paradoxo do observador. Como o objetivo da Sociolinguística é observar e descrever o vernáculo de uma língua, há sempre a possibilidade de o informante optar por um estilo mais monitorado durante o momento de gravação das entrevistas. É nesse momento que o pesquisador precisa ser experiente e criativo o bastante para colher dados o mais próximo possível do vernáculo.

2.2.4 A variação diamésica

A variação diamésica está relacionada com as diferenças existentes entre fala e escrita. A fala é de natureza mais espontânea e improvisada enquanto a escrita é mais artificial e planejada. A fala difere da escrita mesmo nos momentos de alto grau de monitoramento, em que é guiada por um plano do que se vai falar, como ocorre numa conferência, por exemplo. Durante a produção da fala, o sujeito precisa às vezes refazer as sentenças no momento em que estão sendo proferidas, inclusive para adaptar suas ideias ao contexto, o que muitas das vezes pode redirecionar os caminhos previamente traçados pelo falante.

Na produção da escrita, o autor do texto não conta com os *feedbacks* instantâneos do seu interlocutor/leitor para redirecionar o texto, caso seja necessário, nem tem a possibilidade de corrigir ou explicar uma sentença mal interpretada no momento em que seu texto está sendo lido. Por isso, diferentemente da fala, a escrita é, no geral, planejada e pode sofrer diversas revisões antes de chegar ao leitor.

Por causa das diferenças nas condições de produção, a fala é geralmente mais livre e menos monitorada, enquanto a escrita é mais rígida e altamente monitorada, aproximando-se mais do padrão imposto pela gramática normativa. Por serem diferenciados os contextos de produção, essas duas maneiras de realização da língua – fala e escrita – exibem variações diversas.

Nessa seção dedicada ao referencial teórico que deu base à realização do estudo sobre o apagamento de /d/ no morfema de verbos no gerúndio na fala de Feira de Santana, foram apresentados os conceitos chave da Sociolinguística, antes fazendo-se um breve relato do seu surgimento como modelo teórico com vistas a ampliar a visão estruturalista do objeto de estudo da Linguística. Foram apresentados os postulados básicos que fulcram a Sociolinguística como modelo teórico, tais como o conceito de sistema heterogêneo e estruturado, o conceito de variação e mudança como fenômenos intrínsecos ao sistema linguístico, os conceitos de tempo real e tempo aparente, os cinco problemas empíricos postos à Sociolinguística como modelo teórico e as dimensões interna e externa à variação e à mudança linguística.

Os conceitos apresentados nesta seção deram norte aos passos que definiram o caminho metodológico da pesquisa apresentada nesta dissertação. Esse caminho e os passos que lhe percorreram são descritos no capítulo seguinte.

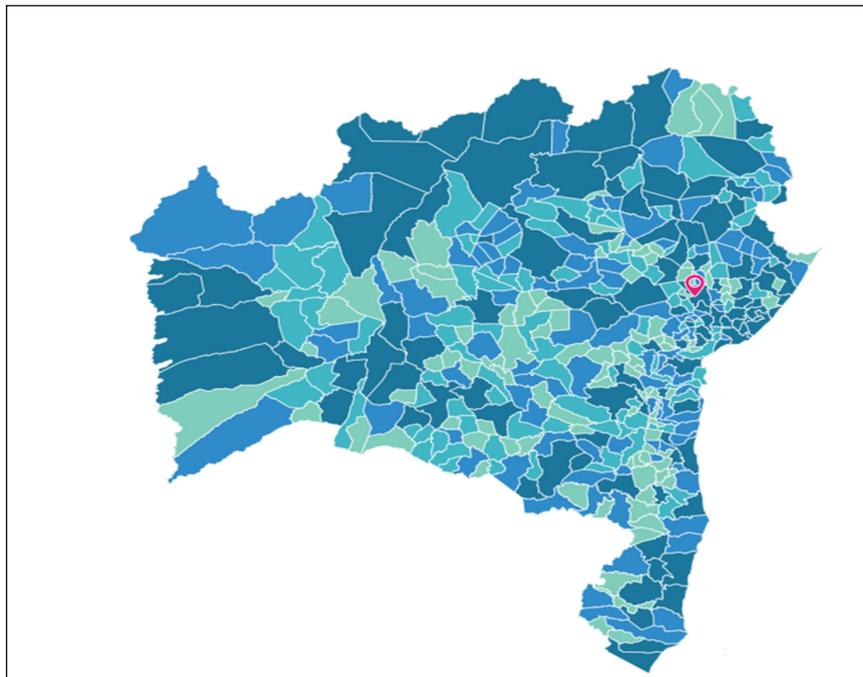
3 O CAMINHO METODOLÓGICO

Neste capítulo estão apresentados os procedimentos metodológicos realizados na fase de levantamento e análise dos dados sobre o apagamento de /d/ no morfema de verbos no gerúndio na fala de Feira de Santana. Apresentam-se nas subseções seguintes a comunidade de fala, o banco de dados de onde foram selecionadas as entrevistas, o *corpus* da pesquisa, a variável dependente, as variáveis independentes, o modelo estatístico e a ferramenta de análise computacional que subsidiou o tratamento quantitativo dos dados.

3.1 A COMUNIDADE DE FALA

Feira de Santana é uma cidade do Leste da Bahia localizada a 108 km da capital, via BR 324. A cidade está localizada no agreste baiano, entre a zona da mata e o sertão. O município tem 614.872 habitantes (IBGE, 2019), é o maior do interior do Estado do ponto de vista populacional, ficando atrás apenas da capital do estado, Salvador. O mapa seguinte situa a localização da cidade no cenário baiano.

Figura 4. Localização de Feira de Santana no Estado da Bahia



Fonte: IBGE, 2019. Disponível em: <https://cidades.ibge.gov.br/brasil/ba/feira-de-santana/panorama>. Acesso em: 18 ago. 2020.

A Figura 5 mostra uma vista aérea da sede da cidade de Feira de Santana, tendo como centro a Praça do Monumento ao Caminhoneiro, cortada pela Avenida Presidente Dutra.

Figura 5. Vista da cidade de Feira de Santana



Fonte: Secretaria Municipal de Planejamento de Feira de Santana.
Disponível em: <http://www.feiradesantana.ba.gov.br/>. Acesso em: 18 jun. 2020.

Feira de Santana tem população maior que a de algumas capitais do país, como Macapá-AP, Boa Vista-RR, Rio Branco-AC, Porto Velho-RO, Palmas-TO, Cuiabá-MT, Aracaju-SE, Vitória-ES e Florianópolis-SC. Além da sede, composta por 44 bairros, a cidade tem oito distritos: Bonfim de Feira, Governador João Durval Carneiro, Humildes, Jaquara, Jaíba, Maria Quitéria, Matinha e Tiquaruçu.

A origem de Feira de Santana é um dado ainda polêmico, sendo que há duas versões sobre como se deu o povoamento das terras que vieram a constituir o município. Segundo Andrade (1990) há uma versão, a que a autora chama tradicional, que atribui o povoamento

das terras que originaram o município ao casal de portugueses católicos Domingos Barbosa Brandão e Ana Brandão; e uma segunda versão, a que a autora chama polêmica, que põe em cheque a primeira a partir de pesquisas realizadas em documentos históricos que retomam ao século XVII.

A versão tradicional da origem do município é sustentada pelo historiador estadunidense Rollie Edwards Poppino, cuja tese de doutoramento, publicada em forma de livro em 1968, afirma que a vila que deu origem à cidade nasceu no século XVIII a partir da construção das capelas de Sant'Ana e de São Domingos, assentadas nas terras da fazenda Olhos d'Água, que pertenciam ao já mencionado casal Domingos Barbosa Brandão e Ana Brandão. Dentro dessa mesma linha histórica, que localiza as origens de Feira de Santana no século XVIII, Miranda, Silva e Oliveira (2013) afirmam que os arredores das capelas de Sant'Ana e de São Domingos serviam como ponto de apoio para viajantes nordestinos, principalmente os comerciantes tropeiros, que viajavam pelos interiores do Nordeste negociando seus produtos originários do sertão da Bahia.

Ainda de acordo com a versão que atribui o povoamento ao casal Brandão, Miranda, Silva e Oliveira (2013) afirmam que havia, ao lado das capelas de Sant'Ana e de São Domingos, uma feira que acontecia todas as segundas-feiras e que atraía os habitantes das fazendas vizinhas. Com o passar do tempo, a feira tomou grandes proporções, levando ao aumento do povoamento da vila, que viria a torna-se oficialmente município, em 1833.

Alguns historiadores como Galvão (1982) e Andrade (1990) polemizam essa versão tradicional dos fatos. Em estudos detalhados com base em documentos que retomam inclusive ao século XVI, esses autores afirmam que as origens de feira de Santana retomam, mais precisamente, ao século XVII, com as incursões do bandeirante João Peixoto Veigas no sertão baiano. Segundo Galvão (1990),

De 1619 a 1795, desde os irmãos João Peixoto Veigas e Felipe Peixoto, durante 150 longos anos, a região de Feira de Santana foi povoada, colonizada, pontilhada de currais ou de engenhos de açúcar, sobretudo com a cultura e exportação de fumo pelos Veigas, que ocuparam sempre as melhores terras, mantiveram hegemonia sobre a terra e o meio em regime quase feudal. (GALVÃO, 1982, p. 28)

Em uma larga e aprofundada pesquisa que resultou em sua dissertação de Mestrado, Andrade (1990) analisa o mito que gira em torno do casal Domingos e Ana Brandão e o silenciamento histórico acerca do bandeirante João Peixoto Veigas, a quem estaria muito mais relacionada a história do povoamento da região que deu origem ao município de Feira de

Santana. Galvão (1982), assim como Andrade (1990), atesta que o repovoamento da região da atual Feira de Santana, território originalmente habitado por índios paiaias, teve início com a fundação da vila de São José das Itaporocas, por volta de 1645, pelo sesmeiro João Peixoto Veigas, que era cristão novo, sendo que a fazenda Olhos d'Água foi fundada em 1820, pelo casal Domingos e Ana Brandão.

A partir da análise da historiografia do povoamento do município, Andrade (1990) procura esclarecer aquilo que ela chamou de controvérsia da origem da cidade, buscando desmistificar uma interpretação consagrada na historiografia tradicional que enaltece alguns personagens em detrimento de outros. Uma das causas para o silenciamento do pioneirismo dos Veigas no protagonismo das origens de Feira de Santana, indica a autora, é o fato de João Peixoto da Veiga ter sido um cristão novo, sendo o mérito atribuído ao casal de católicos.

Se há alguma controvérsia quanto à origem de Feira de Santana, não se pode dizer o mesmo com relação ao potencial socioeconômico do município. Feira de Santana possui o segundo maior entroncamento rodoviário do Brasil, o maior do Norte-Nordeste, perdendo apenas para São Paulo. Dentro do território do município encontram-se as rodovias federais BR 324, BR 116 e BR 101, além das rodovias estaduais BA 052 (Estrada do Feijão), BA 502 (Via São Gonçalo-BA), BA 503 (Via Coração de Maria-BA), BA 499 (que liga a cidade ao distrito de Bonfim de Feira), BA 504 (Via Alagoinhas-BA) e BA 513 (que liga a cidade ao distrito de Humildes). Feira de Santana é a única cidade brasileira cortada pelas duas maiores rodovias federais do país, a BR 101 e a BR 116.

Segundo Araújo (2014, p. 125), “a situação de entroncamento rodoviário faz com que Feira seja muito conhecida, de modo que é comum encontrar pessoas de outros estados brasileiros que conhecem ou já passaram por Feira de Santana”. A autora salienta ainda que, desde sua origem, a cidade já era rota comercial, por onde transportavam-se produtos regionais, gado e outras criações que eram comercializados nas cidades de Cachoeira, Santo Amaro e Salvador.

A localização privilegiada da cidade transformou-a em um motor de crescimento na região. Segundo informações disponíveis no site oficial de mobilidade do município, passam pelas rotas que cortam a cidade uma série de produtos agrícolas que saem da região Oeste. São produtos da fruticultura originários do Norte e produtos industriais provenientes das regiões Sul e Sudeste (FEIRA DE SANTANA, 2019). Ainda segundo informações disponíveis no mesmo endereço eletrônico oficial do município, o setor de serviços feirense corresponde a 2/3 (dois terços) da economia do estado da Bahia, sendo que a cidade funciona

como centro urbano, educacional, tecnológico, econômico, imobiliário, industrial, financeiro, cultural e comercial para mais de 80 municípios.

Outro fator importante para o desenvolvimento socioeconômico de Feira de Santana foi a criação do Centro Industrial Subaé (CIS), em 1970, pela prefeitura municipal da cidade. Treze anos depois, o CIS foi estadualizado por força da Lei nº 4.167/83. O CIS é uma autarquia estadual vinculada à Secretaria Estadual da Indústria, Comércio e Turismo que tem como finalidade fomentar e gerenciar o desenvolvimento da cidade do ponto de vista industrial. A área do CIS abrange desde o Sul do município de Feira de Santana até o Norte do município de São Gonçalo dos Campos. De acordo com o artigo 3º da Lei Estadual Ordinária nº 4167, de 7 de novembro de 1983, que criou o CIS, o mesmo tem como finalidade:

Prover a infraestrutura básica de localização industrial referente à urbanização, transporte, energia, abastecimento de água e esgotos, estimular a implantação de indústrias e promover a alienação de lotes de terrenos para fins industriais e correlatos, em harmonia com as diretrizes da política de desenvolvimento da União, do Estado e do Município de Feira de Santana [...]. (BAHIA, 1983)

O incentivo ao desenvolvimento industrial representado pelo CIS faz do parque industrial de Feira de Santana o terceiro maior do estado da Bahia, ficando atrás apenas do Polo Industrial da Cidade de Camaçari e do Centro Industrial de Aratu (CIA), localizados na região metropolitana de Salvador. Devido ao desenvolvimento no setor industrial e de serviços, a cidade tornou-se um local atrativo para migrantes provenientes de outros municípios baianos e de diversas outras regiões do país que chegam em busca de emprego.

Outro fator que atrai para Feira de Santana pessoas oriundas de outras regiões do estado e do país é a busca por formação e qualificação profissional. A cidade se destaca também como o segundo polo educacional do estado da Bahia. Nela encontram-se inúmeras instituições de ensino técnico profissional, como o Instituto Federal da Bahia (IFBA), entidades do sistema S (SENAI, SENAC, SENAT, SENAR) e o Centro Tecnológico do Estado da Bahia (CETEB). A cidade conta ainda com duas instituições públicas de ensino superior, a UEFS e a UFRB.

Vale salientar que, além dos 28 cursos de graduação, a UEFS oferece também 33 cursos de pós-graduação *stricto sensu*, sendo 24 deles em nível de Mestrado e nove em nível de Doutorado. De acordo com levantamento feito por Santos Jr. (2019), a rede privada de

educação superior na cidade conta com 16 unidades, sendo que nove delas ofertam cursos na modalidade presencial e sete na modalidade de ensino à distância (EAD).

Devido a esse potencial econômico e pela influência que exerce sobre os inúmeros municípios circunvizinhos, a cidade de Feira de Santana alcançou status de metrópole. O projeto para a metropolização da cidade foi protocolado em 2010 pelo deputado estadual Colbert Martins, e em 7 de julho do ano seguinte o governador Jacques Wagner sancionou a Lei Complementar nº 35, que criou a Região Metropolitana de Feira de Santana (RMFS).

A RMFS abrange os municípios de Amélia Rodrigues, Conceição da Feira, Conceição de Jacuípe, Tanquinho, São Gonçalo dos Campos e Feira de Santana. A relação dos outros cinco municípios da região com a metrópole está construída na base da dependência, uma vez que suas populações buscam na cidade de Feira de Santana serviços especializados como saúde e educação, entre outros. Muitos dos habitantes dessas cidades viajam a Feira de Santana diariamente para trabalhar e retornam às suas cidades para dormir, sendo que se tornou comum denominá-las cidades-dormitórios. O projeto de criação da RMFS prevê ainda uma segunda fase em que a região metropolitana será ampliada, com a inclusão de mais dez cidades, contudo não há notícias sobre a previsão para execução dessa segunda fase.

Todas essas características que foram expostas até aqui, típicas de cidades desenvolvidas, fazem com que Feira de Santana seja um local atrativo, conforme salientado anteriormente, sendo que muitas das pessoas que chegam à cidade em busca de emprego e formação decidem fixar moradia no município, de tal modo que não é difícil encontrar pessoas que não nasceram no município, mas que se integraram à sua população.

Essa característica da composição populacional por si só já dá indícios de haver em Feira de Santana grande diversidade em sua norma linguística. Araújo (2014, p. 127) afirma que a localização geográfica de Feira de Santana, no meio do caminho entre o recôncavo e o sertão, faz da cidade um portal para a região semiárida, o que possibilita haver em sua variedade linguística características das normas dos falares do interior e do litoral. Deve-se levar em conta ainda o processo histórico de formação da identidade do município, centrado na modernização, a exemplo da rápida substituição do seu passado rural com tradição na criação de gado, visando sempre a alcançar um status de metrópole, tendo como exemplo sempre a capital, Salvador, onde deveria sempre se espelhar.

A verdade é que Feira de Santana já não consegue dar conta da demanda do amplo número de habitantes que necessita de serviços urbanos especializados. Os serviços de saúde e saneamento, por exemplo, são bastante precários. Os bairros crescem desordenadamente,

fazendo aumentar as periferias próximas ao centro da cidade; somado a isso, grandes empreendimentos imobiliários inauguram, dia após dia, novos bairros planejados de classe média alta afastados do centro, polarizando cada vez mais as diferenças entre classes sociais.

Após essa breve apresentação sobre a cidade de Feira de Santana, comunidade de fala estudada na pesquisa apresentada nesta dissertação, descreve-se na seção seguinte o banco de dados que contém amostra representativa da fala feirense, de onde foram selecionadas as 24 entrevistas sociolinguísticas que compuseram o *corpus* da pesquisa.

3.2 O BANCO DE DADOS

O banco de dados do projeto “A Língua Portuguesa no Semiárido Baiano” (http://www2.uefs.br/nelp/sobre_projeto.htm) está sediado no Núcleo de Estudos da Língua Portuguesa (NELP), do Departamento de Letras e Artes (DLA) da Universidade Estadual de Feira de Santana (UEFS). O projeto foi implantado em 1998 pelas professoras Dra. Norma Almeida e Dra. Zenaide Carneiro e dividiu-se em três fases:

- Fase 1, em que coletaram-se amostras de fala das microrregiões da Chapada Diamantina, mais precisamente nas zonas rurais dos municípios de Anselino da Fonseca e Rio de Contas;
- Fase 2, em que coletaram-se amostras de fala de regiões rurais do norte do semiárido baiano, em comunidades de Jeremoabo (Casinhas, Lagoa do Inácio e Tapera) e região do Paraguaçu (São José das Itaporocas e Matinha);
- Fase 3, em que coletaram-se amostras de fala do município de Feira de Santana.

Na terceira e fase do projeto, realizada entre os anos de 2008 e 2010, o interesse voltou-se para a cidade de Feira de Santana, onde foram gravadas 72 entrevistas sociolinguísticas do tipo DID (diálogos entre documentador e informante) com intuito de representar a comunidade de fala feirense, somando mais de 70 horas de gravação.

Cada entrevista documentada pelos pesquisadores do projeto “A Língua Portuguesa no Semiárido Baiano”, professoras coordenadoras e alunos bolsistas, tem em média cinquenta minutos de gravação e os informantes estão estratificados por faixa etária, sexo, nível de escolarização e relação com a migração.

A amostra da fala feirense está dividida, conforme Araújo e Almeida (2014), da seguinte maneira:

- 48 informantes analfabetos ou pouco escolarizados, sendo que 12 são da zona rural (todos nascidos em Feira de Santana e também a maioria de seus pais) e 36 são da sede do município. Desses 36, 12 são feirenses filhos de feirenses, 12 são feirenses filhos de migrantes e 12 são migrantes;
- 12 informantes com ensino universitário completo, todos nascidos na cidade, filhos de feirenses;
- 12 informantes com ensino médio completo, todos feirenses filhos de feirenses.

Segundo Araújo e Almeida (2014),

A amostra possibilita uma análise contextualizada acerca do binômio variação/mudança, abrangendo aspectos marcantes da sócio-história da comunidade de fala, a exemplo dos contatos interdialetais e dos tardios processos de urbanização do município. (ARAÚJO; ALMEIDA, 2014, p. 36)

Ainda de acordo com Araújo e Almeida (2014), as entrevistas constituem um rico material que pode servir para estudos que visem a investigar, a partir do viés linguístico, “quais são as consequências dos estreitamentos de redes sociais no Brasil a partir da intensificação dos processos de urbanização bem como da democratização de acesso ao ensino” (ARAÚJO; ALMEIDA, 2014, p. 36). As autoras salientam que o cuidado na gravação das entrevistas para que se pudesse diminuir o paradoxo do observador (LABOV, 2008 [1972]) e obter os registros que mais se aproximassem do vernáculo foram uma tônica na coleta. Os temas das entrevistas versam sempre sobre memórias da cidade, da infância e da vida pessoal, inclusive momentos traumáticos vividos pelos informantes (ARAÚJO; ALMEIDA, 2014).

A subamostra selecionada dentre as entrevistas do banco de dados aqui descrito, que serviu para a composição do *corpus* da pesquisa sobre o apagamento da oclusiva /d/ no morfema de verbos no gerúndio na fala de Feira de Santana, está descrita na seção seguinte.

3.3 O *CORPUS*

O *corpus* para a realização da pesquisa apresentada nesta dissertação é composto de 24 entrevistas selecionadas dentre as gravações referentes à fase 3 do projeto “A Língua Portuguesa no Semiárido Baiano”. Os 24 informantes são feirenses filhos de feirenses e estão estratificados em escolaridade, sexo e faixa etária, conforme Quadro 2, a seguir.

Quadro 2. Estratificação social do o *corpus* da pesquisa

Sexo	Faixa etária	Escolaridade	Informante ³
Feminino	25 a 35 anos	4ª série	M1B31
		5ª série	M1B33
		Ensino Superior	M1S26
		Ensino Superior	M1S30
	45 a 55 anos	2ª série	M2B50
		4ª série	M2B54
		Ensino Superior	M2S45
		Ensino Superior	M2S48
	De 65 em diante	3ª série	M3B76
		3ª série	M3B70
		Ensino Superior	M3S68
		Ensino Superior	M3S67
Masculino	25 a 35 anos	4ª série	H1B35
		5ª série	H1B26
		Ensino Superior	H1S33
		Ensino Superior	H1S25
	45 a 55 anos	5ª série	H2B45
		2ª série	H2B50
		Ensino Superior	H2S53
		Ensino Superior	H2S56
	De 65 em diante	Analfabeto	H3B80
		2ª série	H3B72
		Ensino Superior	H3S67
		Ensino Superior	H3S69

Fonte: Elaborado pelo autor.

O processo de seleção dos dados foi feito primeiramente a partir da transcrição das entrevistas, etapa em que foram marcadas as ocorrências de verbos no gerúndio a partir da leitura dos inquéritos, depois os dados foram transcritos foneticamente a partir da escuta dos áudios das entrevistas, de oitiva.

3.4 A VARIÁVEL DEPENDENTE

Definiu-se como variável dependente manutenção (variante padrão) ou o apagamento (variante não-padrão) da consoante oclusiva coronal /d/ no morfema *-ndo*, indicador de gerúndio. A variável dependente é, portanto, de natureza binária, conforme se exemplifica a seguir, a partir de trechos de fala retirados do *corpus* da pesquisa.

³ A primeira letra do código refere-se ao sexo do informante, sendo M para mulher e H para homem; o número logo após a indicação do sexo refere-se à faixa etária, as letras B e S seguintes referem-se à escolaridade (baixa e superior) e o último número formado por dois algarismos refere-se à idade do informante. Por exemplo, o código H1S25 diz que o informante é do sexo masculino, da primeira faixa etária, de nível de escolaridade universitário e tem 25 anos de idade. A codificação foi criada pelo autor da dissertação.

Variante padrão:

(27) *Os programas que eu assisto é desenho animado [ri] com ele fazendo companhia, mas sempre que eu posso eu assisto normalmente o... Quando eu consigo assistir é o Jornal da Globo que é meia noite e depois o Programa do Jô. (M1S30)*

Variante não-padrão:

(28) *Então, e falano sobre o trote...é... como ritual, como ritual, como tradição, eu seria, eu estaria negano minha formação se eu falasse que eu era contra. (H1S25)*

3.5 AS VARIÁVEIS INDEPENDENTES SOCIAIS

Para compreender a variação e a mudança linguística, tomando como instrumental de análise a Sociolinguística Variacionista, o pesquisador deve controlar os fatores internos e externos que o ajudem a compreender como a variável dependente está sendo condicionada. Para descrever como se dá o apagamento de /d/ no morfema de verbos no gerúndio na variedade de fala feirense, as variáveis sociais controladas são aquelas apresentadas juntamente com a estratificação do *corpus* na seção anterior, ‘escolaridade’, ‘sexo’ e ‘faixa etária’.

3.5.1 Escolaridade

No que diz respeito à escolaridade, optou-se por controlar na pesquisa dois níveis de escolarização: Nível Fundamental I, representando a norma popular, e Nível Superior, representando a norma culta. O intuito foi verificar como a variável condiciona o uso do gerúndio a partir da atuação de dois níveis extremos de escolarização. Sabe-se que a escola funciona como mantenedora de formas linguísticas prestigiadas em detrimento daquelas tidas como estigmatizadas (VOTRE, 2013).

Estudos em Sociolinguística em geral mostram que indivíduos que tiveram acesso aos níveis mais elevados de escolarização tendem a utilizar formas tidas como de prestígio, enquanto aqueles que não tiveram a oportunidade de acesso à escola ou que o tiveram, mas de forma precária, a exemplo dos que cursaram apenas os anos iniciais da educação básica, tendem a usar mais as formas tidas como estigmatizadas. Quanto maior o contato com a cultura letrada, maior também é a chance de aproximação com a norma padrão, o que faz com

que os usos advindos dessa norma sejam mais comuns entre falantes da norma culta e, por isso, valoradas como sendo de prestígio.

A hipótese de pesquisa para a variável escolaridade é a de que o apagamento da oclusiva /d/ no morfema de verbos no gerúndio seja motivado por indivíduos que estudaram até o Ensino Fundamental I e desmotivado por aqueles que concluíram o Ensino Universitário, pois os resultados obtidos em estudos sobre o mesmo fenômeno indicam essa tendência de uso na variável dependente, mostrando que quanto mais alto o nível de escolarização, menor é a probabilidade de ocorrer o apagamento.

3.5.2 Sexo

A maioria dos estudos apresentados na seção 1.2 mostra, seguindo uma tendência geral em Sociolinguística, que homens e mulheres se comportam de maneiras diferentes frente ao fenômeno de apagamento de /d/ no morfema de verbos no gerúndio, sendo que as mulheres quase sempre se mostram mais conservadoras, e os homens quase sempre se mostram mais abertos à forma não-padrão. Isso, contudo, não é uma regra, visto que há estudos que mostram não haver diferenças no uso entre homens e mulheres (GONÇALVES, 2018) ou que as mulheres estão à frente do processo (NASCIMENTO; ARAÚJO; CARVALHO, 2013).

A hipótese aventada para a variável social ‘sexo’ na pesquisa sobre a fala de Feira de Santana é que o apagamento seja favorecido pelos homens e desfavorecido pelas mulheres, como tem sido constatado na maioria dos estudos sobre o mesmo tema.

3.5.3 Faixa etária

A partir da observação de como a variável ‘faixa etária’ se relaciona com a mudança/variação linguística no indivíduo e na comunidade, é possível inferir sobre uma mudança linguística em tempo aparente ou sobre uma gradação etária na comunidade estudada. Na pesquisa sobre o apagamento de /d/ no morfema de gerúndio na fala de Feira de Santana, foram controladas as três faixas etárias conforme a estratificação do banco de dados apresentado anteriormente (faixa 1: 25 a 35 anos, faixa 2: 35 a 45 e faixa 3: de 65 anos em diante). A hipótese aventada foi que os falantes da faixa 1 favoreçam o apagamento, uma vez que falantes mais jovens são mais receptivos a formas inovadoras na língua.

3.6 AS VARIÁVEIS INDEPENDENTES LINGUÍSTICAS

As variáveis sociais ‘sexo’, ‘escolaridade’ e ‘faixa etária’ são comumente controladas na maioria dos estudos sociolinguísticos, uma vez que são variáveis *default* nos bancos de dados existentes no Brasil (FREITAG, 2016), salvo as peculiaridades metodológicas de cada coleta. As variáveis linguísticas, por sua vez, têm a ver com a natureza do fenômeno pesquisado.

No que diz respeito ao apagamento de /d/ no morfema de verbos no gerúndio, as variáveis linguísticas comumente controladas são a ‘conjugação verbal’, a ‘extensão do verbo’ e o ‘contexto seguinte’. Há trabalhos que se controlaram também a ‘estrutura sintática da frase’, o ‘tipo de auxiliar na perífrase’, o ‘tamanho do auxiliar na perífrase’ e a ‘existência de material interveniente entre o auxiliar e o verbo no gerúndio’. Dentre essas quatro últimas variáveis citadas apenas a ‘estrutura sintática’ foi estatisticamente significativa na explicação do fenômeno.

Optou-se por controlar na pesquisa realizada sobre a fala de Feira de Santana apenas as variáveis internas que se mostraram significativas em pelo menos um dos trabalhos apresentados na seção 1.2 deste texto. São elas a ‘conjugação verbal’, a ‘extensão do verbo’, o ‘contexto seguinte’ e a ‘estrutura sintática’. Essas quatro variáveis, bem como as hipóteses de pesquisa para cada uma delas, estão apresentadas nas subseções seguintes.

3.6.1 Conjugação verbal

A variável ‘conjugação verbal’ revelou-se significativa para a explicação do apagamento da oclusiva /d/ no morfema de verbos no gerúndio na fala de Fortaleza-CE (NASCIMENTO; ARAÚJO; CARVALHO, 2013), sendo que verbos das duas primeiras conjugações favoreceram o apagamento. Resultados semelhantes a esses foram também observados por Gonçalves (2018) para a fala paulistana. Decidiu-se verificar também na pesquisa feita na fala de Feira de Santana se a conjugação verbal condiciona ou não o apagamento. A hipótese inicial foi de que verbos da primeira conjugação favorecessem a aplicação da regra de apagamento por serem os mais frequentes na língua, pois, segundo Bybee (2003), itens de alta frequência aderem rapidamente à mudança sonora. A variável ‘conjugação verbal’ está exemplificada a seguir.

1ª conjugação

(29) *O que é que tá por trás daquilo? Qual é o problema familiar? E conversano c'um ele em particular ou... ou levando... em diálogo ou... tentar compreendê-lo, tem resolvido. Nunca tive problema c'um aluno nesses é... vinte e oito anos que eu ensino aqui. Nunca tive.* (M2S53)

2ª conjugação

(30) *Da segunda vez retornei pros mesmos lugares, só que desta vez eu fui buscar uma filha que tava fazeno intercâmbio no norte dos Estados Unidos.* (H2S53)

3ª conjugação

(31) *É! E resgatar e entregar. A gente vai aos poucos saino do comando, que, aliás, o bando não tem comando. É todo fora de ordem!* (MS363)

3.6.2 Extensão do verbo

Estudos anteriores têm mostrado que a extensão do verbo condiciona a regra variável pesquisada, de modo que a probabilidade de apagamento é menor nos dissílabos e maior nos trissílabos e polissílabos (MOLLICA; MATTOS, 1992; LUCENA; VASCONCELOS, 2007; VIEIRA, 2011; NASCIMENTO; ARAÚJO; CARVALHO, 2013; ALMEIDA; OLIVEIRA, 2017; GONÇALVES, 2018; ARAÚJO, 2019). Assim, quanto maior a extensão do vocábulo, maior tendência ao apagamento de /d/ no morfema de gerúndio. A variável é caracterizada pela quantidade de sílabas do verbo, conforme está exemplificado a seguir.

Dissílabos

(32) *Minha mãe me dava as coisas, dava a meu irmão e dizia: 'Hoje eu estou lhe dando # porque eu tenho, se amanhã não puder, se amanhã não tiver, eu não vou poder lhe dar'.* (M1B30)

Trissílabos

(33) *Hoje não, hoje no sétimo semestre, é porque entrei no início do curso de enfermagem, foi a segunda turma de enfermagem, então veio mudano, veio mudano, veio mudano e hoje não, hoje tem vários professores da UEFES são de lá.* (H1S25)

Polissílabos

(34) *Curso noturno é muito difícil. Você acaba trabalhano o dia inteiro, ainda tem que vim pra faculdade, assim eu vim quase os meus seis anos de universidade eu vim todos os dias do trabalho.* (M2S42)

3.6.3 Contexto seguinte

A variável ‘contexto seguinte’ permite verificar se o fato de haver pausa, palavra iniciada por consoante ou palavra iniciada por vogal após o verbo no gerúndio condiciona e de que maneira condiciona o apagamento da oclusiva /d/ no morfema de gerúndio. Nascimento, Araújo e Carvalho (2013) observaram que o fato de haver segmento consonantal no contexto seguinte favorece o apagamento, enquanto que as ocorrências de segmento vocálico e de pausa inibem o apagamento.

Almeida e Oliveira (2017) sugerem que sejam controladas em separado, dentre as consoantes, aquelas que possibilitam a ocorrência de haplologia (no caso as consoantes /d/, /n/, /t/) no contexto seguinte a um verbo no gerúndio, pois a haplologia é um processo de natureza geral, sem motivação morfológica.

Optou-se por controlar na pesquisa aqui apresentada a variável ‘contexto seguinte’, uma vez que esta se mostrou significativa em vários estudos anteriores (MOLLICA; MATTOS, 1992; LUCENA; VASCONCELOS, 2007; VIEIRA, 2011; NASCIMENTO; ARAÚJO; CARVALHO, 2013; ALMEIDA; OLIVEIRA, 2017; GOLÇALVES, 2018; ALMEIDA; PEREIRA; ARAÚJO, 2020). Seguiu-se na pesquisa com a fala de Feira de Santana o modelo adotado por Nascimento, Araújo e Carvalho (2013), observando-se o que propõem Almeida e Oliveira (2017). Dessa forma, a variável ‘contexto seguinte’ na pesquisa é composta pelos fatores ‘contextos de pausa’, ‘contextos de vogais’, ‘contextos de consoantes’ e ‘contextos de haplologia’.

As 24 entrevistas selecionadas para compor o *corpus* da pesquisa já se encontram transcritas no banco de dados do projeto “A Língua Portuguesa no Semiárido Baiano”. O trabalho de transcrição das mesmas foi realizado pelas próprias professoras coordenadoras e pelos bolsistas do Projeto. Optou-se por partir desse material para buscar as ocorrências dos dados.

Como a transcrição se submete às normas de escrita formal da língua portuguesa, ela nem sempre sinaliza as pausas onde elas realmente acontecem na fala. De tal forma, há momentos em que há um ponto de segmento indicando o final do período, quando na verdade

a fala segue em fluxo contínuo. Outras vezes, o falante faz pausa quando, na verdade, não há possibilidade de marcação de pausa por pontuação conforme as normas da escrita. Diante disso, decidiu-se por indicar, na transcrição, a pausa por meio do símbolo #, mesmo onde a pontuação faça parecer óbvia a existência da pausa. Os fatores da variável linguística ‘contexto seguinte’ estão exemplificados a seguir.

Pausa

(35) *A gente estudar é bom, estudar a bíblia, ler a bíblia, e a gente.. e.. pá... ir... meditano #. A gen... tem que meditar.* (M3B76)

Vogal

(36) *Aí, assim, a gente acaba perdeno o contato, né, porque também a vida da gente cada um toma um rumo diferente, faculdade, casamento, filho, e aí cada um vai viver sua vida na verdade, né?* (M1S30)

(37) *Eu fico no MSN com meu marido que tá lá em Salvador ou fico assistino um filme, que é assim, é a hora também que, além de diversão, eu uso pra descansar, né, um pouquinho.* (M1S30)

Consoante

(38) *Hoje em dia, eu tava lendo uma entrevista um dia desses de uma psicóloga falano que o povo tá com a mania de dizer que criança não pode...* (M3B76)

(39) *O tempo todo, né? Assim, na verdade eu não tô teno muito tempo pra lazer, minha diversão é...* (M1S30)

Haplologia

(40) *A Bíblia não leio muito porque não estou enxergano direito.* (M3B79)

(41) *Trabalho de roça eu nunca gostei, aí ficava den’de casa, cuidano dos ôto menor.* (M2B50)

3.6.4 Estrutura sintática

Segundo Ferreira (2010), as formas de gerúndio ocorrem basicamente em três estruturas sintáticas: perífrase, justaposição e oração reduzida. O grupo de fatores ‘estrutura sintática’, proposto pela referida autora para o estudo da variedade de fala de São José do Rio

Preto-SP, revelou que os contextos de perífrase e justaposição favorecem a aplicação da regra de apagamento de /d/ no morfema de verbos no gerúndio naquela variedade do PB, ao passo que o contexto de oração reduzida é inibidor. De tal forma, decidiu-se por controlar a variável ‘estrutura sintática’ na pesquisa sobre o apagamento de /d/ no morfema de verbos no gerúndio na fala de Feira de Santana-BA sob a hipótese de que o fator ‘perífrase’ seja favorecedor do apagamento, de acordo com o que constatou Ferreira (2010). Os três fatores controlados na variável linguística estrutura sintática são exemplificados a seguir.

Perífrase

(42) *Hum, hum. Porque já foi... procurado... esses estudantes foram procurados e... por diversas vezes... é... tentou-se diálogo, tentou-se um... mudar as coisa, mas é a gen... estamos vendo que a cada ano a violência tem é... tem se estabelecido e se ampliado.* (H2S53)

Nos contextos de justaposição, o primeiro verbo no gerúndio está em uma estrutura perifrástica e é imediatamente seguido por um ou mais verbos também no gerúndio. Nesse tipo de estrutura, considera-se que cada verbo da construção sintática é um dado de justaposição, conforme está exemplificado a seguir.

Justaposição

(43) *Aí quando a barriga vai crescono, a família vai olhano de outro modo, vai gostano, gostano e tá aí hoje.* (M1B30)

Oração reduzida

(44) *Eu acho a família assim é... a família ela tem, ela tem uma importância na vida das pessoas, seja no momento que você nega ela e constrói a sua vida totalmente contrário ao que sua família... ou seja, você tentano imitar.* (H1S25)

Definidas as variáveis independentes, procedeu-se à etapa de codificação para que os dados fossem quantitativamente tratados por meio do programa computacional GoldVarb X, a fim de que se obtivessem resultados em taxa de aplicação, frequência percentual e pesos relativos⁴.

⁴ O peso relativo tem a ver com a probabilidade de a variável independente estar ou não condicionando a regra variável. Este conceito é explanado na seção seguinte.

Após a apresentação da comunidade de fala, do *corpus* da pesquisa, da variável dependente e das variáveis linguísticas controladas no estudo, apresentam-se, na seção seguinte, o modelo matemático de tratamento dos dados e a ferramenta computacional que serviu de apoio para a análise quantitativa dos dados.

3.7 O MODELO ESTATÍSTICO APLICADO NO TRATAMENTO DOS DADOS

De acordo com Guy e Zilles (2007), a pesquisa quantitativa se resume a três passos básicos, os quais estão apresentados em (a), (b) e (c) abaixo:

- (a) A coleta dos dados;
- (b) A redução e apresentação dos dados;
- (c) A interpretação e a explicação dos dados.

Como já foi salientado aqui nesta mesma seção, dedicada a explicar a metodologia da pesquisa, a fase a que se refere o item (a), citado acima, foi realizada no âmbito do Projeto “A Língua Portuguesa no Semiárido Baiano”, de onde foi selecionada a amostra que compôs o *corpus* da pesquisa. Coube ao pesquisador definir o *corpus*, identificar, selecionar e tratar os dados a fim de que se pudesse responder às perguntas de pesquisa. Conforme o esquema apresentado por Guy e Zilles (2007), acima, esta seção se dedica a explanar o segundo passo da pesquisa, o passo (b), com destaque mais precisamente para como foram tratados estatisticamente os dados. O passo (c) será desenvolvido no capítulo 4, referente à apresentação e análise dos dados.

Berlinck e Biazolli (2018) argumentam que a diversidade de caminhos viáveis existentes para a realização de investigações em linguística torna a escolha do modelo metodológico uma decisão de extrema importância. Em Sociolinguística Variacionista, conforme afirma Oliveira (2009), a escolha do modelo estatístico tem a ver com a pergunta que se procura responder e com a natureza das variáveis envolvidas no estudo. No caso da pesquisa sobre o apagamento de /d/ no morfema de verbos no gerúndio na fala feirense, a pergunta, do ponto de vista estatístico, seria basicamente a seguinte: como as variáveis independentes $X_1, X_2, X_3, \dots, X_n$ se relacionam com a variável dependente Y , sendo Y uma variável discreta? Esse tipo de problema pode ser resolvido por meio de uma técnica matemática denominada de modelo de regressão logística. Esse modelo matemático permite que se possa explicar um fenômeno variável (a variável dependente) relacionando-o a um

conjunto de fatores (variável independente) que podem ou não estar influenciando a variação (regra variável) no referido fenômeno.

O modelo de regressão logística é aplicado quando a regra variável é de natureza binária e discreta, ou seja, quando há duas possibilidades de ocorrências na variável dependente. No caso do apagamento de /d/ no morfema *-ndo* de verbos no gerúndio, temos uma variável dependente binária: realização e não realização da oclusiva /d/ no morfema *-ndo*. Além de binária, a variável em estudo é discreta, ou se realiza como *-ndo* ou como *-no*, podendo ser representada em termos de 0 e 1, sendo 0 a não aplicação e 1 a aplicação da regra de apagamento. Essa configuração da variável dependente faz com que o modelo de regressão logística seja ideal para análise do fenômeno variável de apagamento de /d/ no morfema de gerúndio.

Além de permitir demonstrar se há ou não relação entre a variável dependente e as variáveis independentes, o modelo de regressão logística permite responder, em termos de probabilidade, a chance de uma regra variável se aplicar ou não quando relacionada a cada fator das variáveis independentes. De tal forma, sendo o grupo de fatores ‘sexo’ selecionado como significativo para a aplicação da regra, pode-se obter a probabilidade de a oclusiva /d/ ser apagada caso o falante seja do sexo masculino, por exemplo. Tais resultados são fornecidos em termos de peso relativo.

O peso relativo é um resultado probabilístico representado por um número entre 0 e 1, em que 0 representa a probabilidade mínima de relação entre os fatores da variável independente e a variável dependente, e 1 representa a possibilidade máxima de relação entre eles. Em Sociolinguística, diz-se que um resultado inferior a 0,5 inibe a aplicação de certa regra, e que um resultado acima de 0,5 favorece a sua aplicação. Um resultado de exatamente 0,5 significa que a relação entre o fator e a aplicação da regra é inexistente.

Esse modelo matemático, além de outros aplicáveis na análise de fenômenos variáveis, vem sendo implementado, por diversos pesquisadores, em algoritmos computacionais para que se possam operar grandes quantidades de dados em estudos sobre a linguagem. Na pesquisa sobre a fala de Feira de Santana, foi usado o *software* GoldVarb X (SANKOFF; TAGLIAMONTE; SMITH, 2005), que é apresentado na subseção seguinte.

3.7.1 A ferramenta computacional para a análise estatística

Na pesquisa sobre o apagamento de /d/ no morfema *-ndo* de verbos no gerúndio, usou-se o programa computacional GoldVarb X (SANKOFF; TAGLIAMONTE; SMITH,

2005), desenvolvido especificamente para análise estatística de fenômenos variáveis em Sociolinguística. O GoldVarb X aplica o modelo de regressão logística, conforme descrito acima, aos dados fornecidos pelo pesquisador. Os dados devem ser previamente codificados, seguindo uma sintaxe específica e convertidos em um arquivo de extensão *.tkn* para que esse possa ser lido pelo programa.

O *software* GoldVarb X seleciona as variáveis independentes significativas para a explicação da regra variável, ou seja, os grupos cujos fatores condicionam a aplicação da regra em estudo, e fornece pesos relativos para cada fator do grupo em função da aplicação da regra. A seleção dos grupos de fatores se baseia numa medida estatística chamada de nível de significância, cujo resultado é fornecido em *p-valor*. O *p-valor* é a probabilidade de que a distribuição dos dados seja verdadeira, obedecendo a uma margem de erro de 5%. Assim, se o *p-valor* for igual ou menor a 0,05 então o grupo de fatores será selecionado pelo *software* como sendo significativo, caso o *p-valor* seja maior que 0,05 o grupo de fatores será descartado e a distribuição dos dados será atribuída ao acaso. Isso significa que a chance aceitável de a distribuição dos dados ser dada ao acaso deve ser de no máximo de 1 em 20.

Como se pode observar, a metodologia da Sociolinguística Variacionista conta com um instrumental de análise que permite ao pesquisador lidar com grandes quantidades de dados e obter resultados estatísticos de maneira computadorizada. Contudo deve-se ter em conta, conforme afirmam Guy e Zilles (2007), que

Toda estatística inferencial deve ser vista como ferramenta para auxiliar no processo de descoberta do pesquisador, na testagem de hipóteses e assim por diante, em vez de ser considerada como prova ou refutação definitiva de uma hipótese. (GUY; ZILLES, 2007, p. 33)

Cabe ao pesquisador, por conseguinte, interpretar qualitativamente, com base na teoria e no seu conhecimento sobre a comunidade estudada, os resultados estatísticos fornecidos pelo programa computacional, e propor explicações que ajudem a compreender a configuração de fenômenos variáveis na língua.

O capítulo dedicado à metodologia da pesquisa se encerra aqui, com a apresentação do modelo estatístico e da ferramenta de análise computacional que subsidiaram a pesquisa. Esses dois últimos pontos são de extrema importância em qualquer pesquisa que tome como aporte teórico-metodológico a Sociolinguística, visto que esse ramo da Linguística procura explicar os fatos da língua a partir observação de contextos reais de fala, e, portanto, lida com análise qualitativa de grande quantidade de dados, o que justifica a boa escolha de um modelo

matemático bem como o auxílio de uma ferramenta computacional capaz de operar cálculos complexos.

Os resultados do tratamento estatístico dos dados do *corpus* da pesquisa são apresentados juntamente com a análise qualitativa no capítulo seguinte.

4 APRESENTAÇÃO E ANÁLISE DOS DADOS

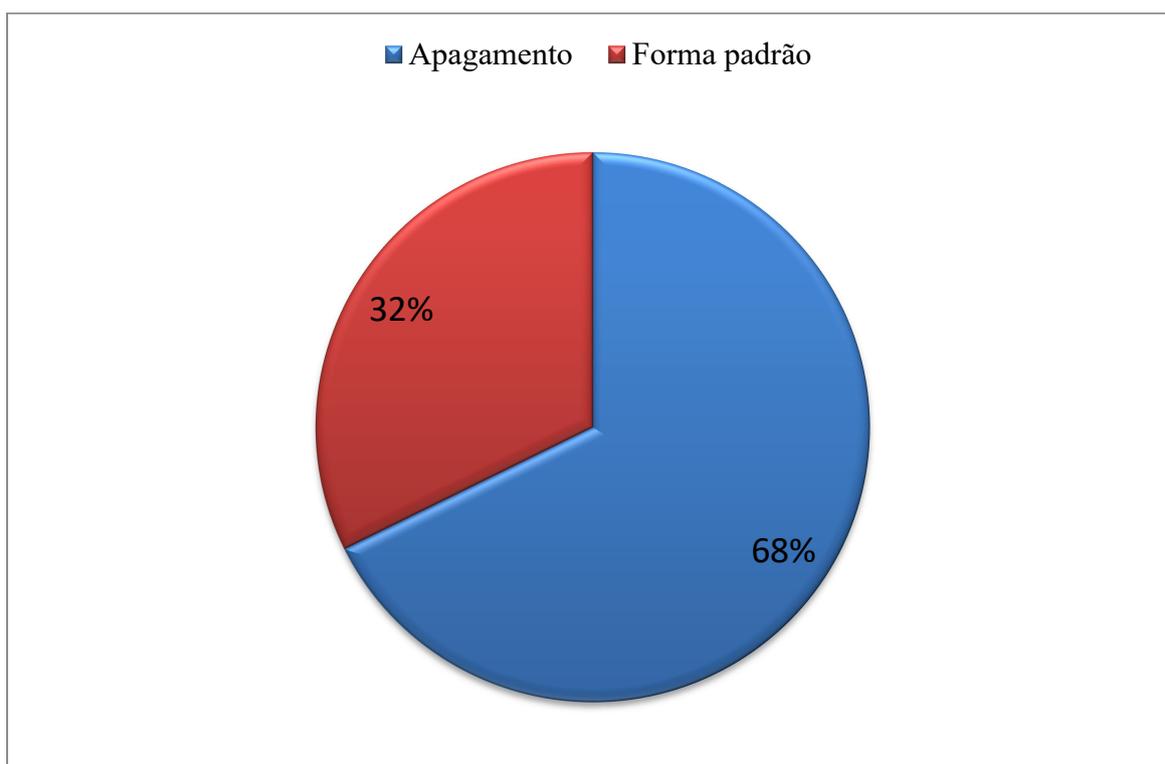
No capítulo anterior desta dissertação, foi apresentado o caminho metodológico percorrido para o estudo do fenômeno de apagamento de /d/ em morfema de verbos no gerúndio na variedade do português falada na cidade de Feira de Santana-BA. Agora são apresentadas a análise quantitativa dos dados e a interpretação dos mesmos à luz da Sociolinguística (WEINREICH; LABOV; HERZOG, 2006 [1968]; LABOV, 2008 [1972]).

4.1 DADOS GERAIS

Após a preparação do *corpus*, foi realizada uma primeira rodada no programa computacional GoldVarb X e constatou-se um total de 904 ocorrências de formas verbais de gerúndio, das quais 612 foram de apagamento (forma não-padrão) e 292 foram de manutenção (forma padrão) de /d/ no morfema *-ndo*. O percentual de apagamento no *corpus* é de 67,7%.

O resultado geral para a variável dependente aponta para um alto índice de ocorrência da variante não-padrão, revelando que o fenômeno é característico da fala feirense. O Gráfico 1 ilustra as variantes em competição na amostra.

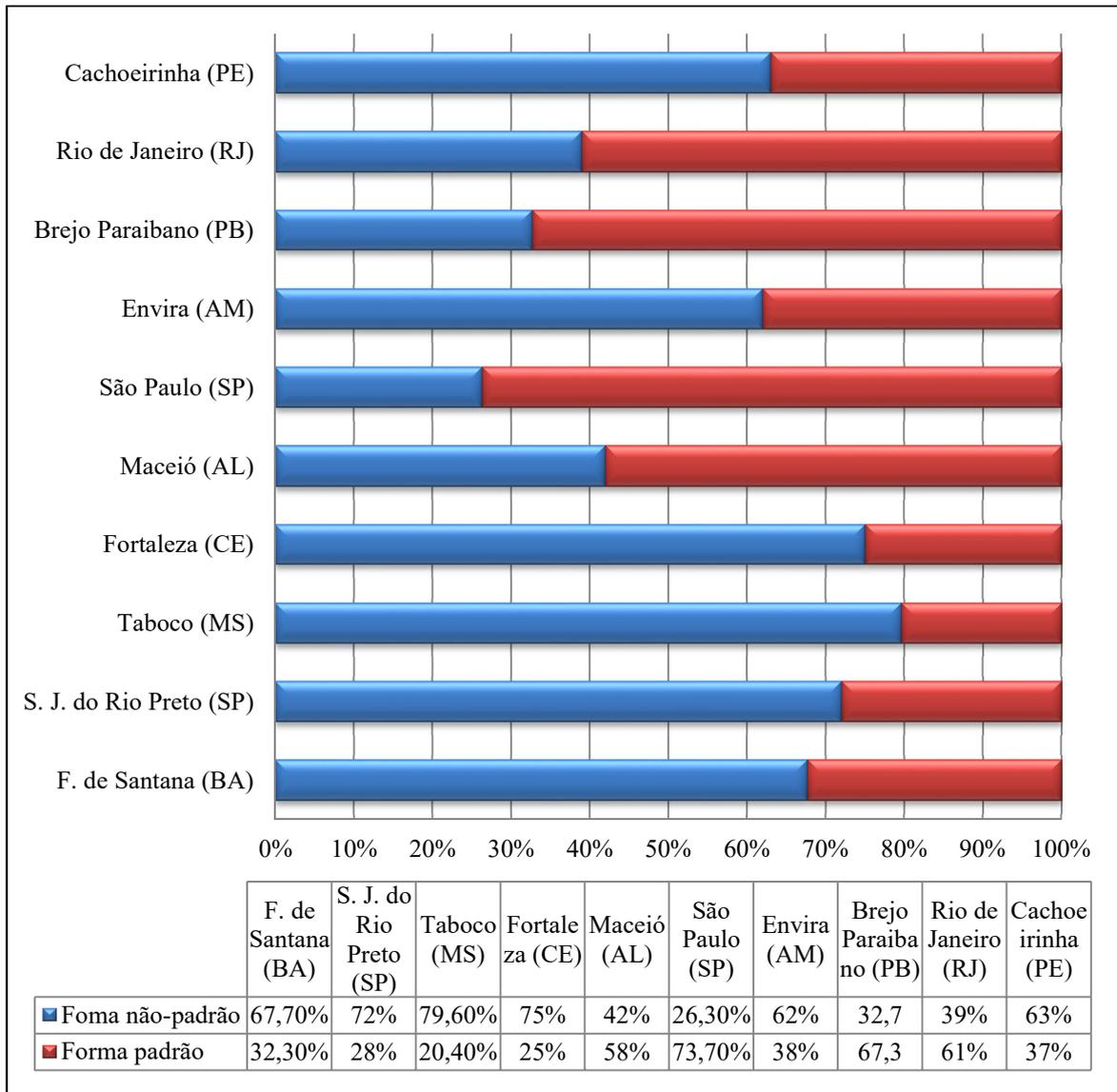
Gráfico 1. Apagamento de /d/ no morfema de gerúndio: Feira de Santana (BA)



Fonte: Elaborado pelo autor.

O Gráfico 2 mostra o avanço da forma inovadora sobre a forma padrão na fala de Feira de Santana em comparação com outras variedades do PB, conforme estudos apresentados no Capítulo 1, seção 1.2 desta dissertação.

Gráfico 2. As variantes em competição: Feira de Santana-BA e outras variedades do PB



Fonte: Elaborado pelo autor.

Percebe-se a partir da leitura do Gráfico 2 que a taxa de apagamento na fala de Feira de Santana aparece semelhante àquelas obtidas para outras variedades do PB. É notável que, com exceção de São Paulo-SP e Brejo Paraibano-PB, a frequência de uso da forma inovadora se mostra bastante avançada em relação à forma padrão, de maneira que os resultados para a

taxa percentual de uso em Feira de Santana se assemelha às encontradas na maior parte dos estudos apresentados na seção 1.2 deste texto.

4.2 ANÁLISE DAS VARIÁVEIS ESTATISTICAMENTE SIGNIFICATIVAS

Apresentam-se, a seguir, os resultados encontrados no *corpus* de acordo com os grupos de fatores linguísticos e extralinguísticos controlados na pesquisa. As variáveis selecionadas pelo *software* de análise estatística foram, nesta ordem de importância: ‘escolaridade’, ‘sexo’, ‘contexto seguinte’, ‘faixa etária’ e ‘extensão do verbo’, conforme Quadro 3. A melhor rodada teve *input* 0,734, *log likelihood* -449.983 e nível de significância 0,007. As variáveis ‘conjugação verbal’ e ‘estrutura sintática’ foram descartadas pelo programa computacional e serão apresentadas na seção 4.4.

Quadro 3. Ordem de seleção das variáveis independentes pelo *software* GoldVarb X

SELEÇÃO DAS VARIÁVEIS	
Variável	Ordem
Escolaridade	1 ^a
Sexo	2 ^a
Contexto seguinte	3 ^a
Faixa etária	4 ^a
Extensão do verbo	5 ^a
Conjugação verbal	Descartada (sem significância estatística)
Estrutura sintática	Descartada (sem significância estatística)

Fonte: Elaborado pelo autor.

A regra de aplicação escolhida foi o apagamento da oclusiva /d/ no morfema *-ndo*, como em “*Eu fiquei aqui trabalhano com ele e gostei*” (H3B72). Assim, quando se fala em aplicação da regra, neste capítulo, refere-se ao apagamento da oclusiva /d/ no morfema *-ndo*, indicador de gerúndio. Ainda que a aplicação do apagamento tenha índice de 67,7% no *corpus*, serão usadas as expressões “forma inovadora” e “forma não-padrão”, ao passo que a forma variante em que há a manutenção do segmento /d/ no morfema *-ndo* será tratada como “forma conservadora” e “forma padrão”, como acontece nos trabalhos apresentados no capítulo 1, seção 1.4 desta dissertação.

Ressalte-se que as três variáveis sociais foram selecionadas, duas delas como as primeiras em ordem de importância, como pode ser visto no Quadro 3, indicando que o fenômeno variável analisado é mais condicionado social do que estruturalmente. A ordem de apresentação dos resultados, contudo, inicia-se pelas variáveis linguísticas, seguindo-se das sociais, conforme se apresenta a seguir.

4.2.1 Análise das variáveis linguísticas

Dentre as variáveis linguísticas controladas na pesquisa, a saber, ‘conjugação verbal’, ‘contexto seguinte’, ‘extensão do verbo’ e ‘estrutura sintática’, foram selecionadas como relevantes para a aplicação da regra o ‘contexto seguinte’ e a ‘extensão do verbo’, em terceiro e quinto lugares na ordem de importância, respectivamente. Os resultados encontrados estão apresentados nas subseções a seguir.

4.2.1.1 Contexto seguinte

A variável ‘contexto seguinte’ foi controlada em várias pesquisas anteriores sobre o apagamento do /d/ no morfema de gerúndio (MOLLICA; MATTOS, 1992; LUCENA; VASCONCELOS, 2007; VIEIRA, 2011; NASCIMENTO; ARAÚJO; CARVALHO, 2013; ALMEIDA; OLIVEIRA, 2017; GOLÇALVES, 2018; ALMEIDA; PEREIRA; ARAÚJO, 2020). De tal forma, decidiu-se observar na pesquisa com dados de Feira de Santana se há relação entre o contexto fonético-fonológico seguinte e o fenômeno linguístico investigado. Os contextos de haplogogia, consoantes, vogal e pausa estão exemplificados a seguir.

Haplogogia (quando há /t, d, n/ no ambiente seguinte ao morfema *-ndo*)

(45) *E a outra foi que comi um negócio, fiquei vomitano, tomei soro.* (H2B50)

(46) *Me casei, minha vida era essa, era costurano direto.* (M3B70)

(47) *Eu vou ver, tô pensano niss: se eu volto a estudar esse ano.* (M1B33)

Consoantes

(48) *Hoje eu tô viajano muito pelo som pernambucano. Bombojó, Cordel do Fogo Encantado... esse tipo de coisa.* (H1S25)

Pausa

(49) *Ele começou a martelar aquilo na minha ideia e me deu confiança dizendo # que eu tinha toda confiança de ensinar e eu terminei indo # e depois eu abandonei todas as outras ideias.* (H2S56)

Vogal

(50) *Bota um pouquinho de água no caruru e vai mexeno até virá mingau.* (M1B33)

A hipótese inicial de pesquisa foi a de que contextos de consoantes favorecessem o apagamento. Os resultados podem ser vistos na Tabela 14.

Tabela 14. Apagamento de /d/ no morfema de gerúndio na fala de Feira de Santana: Contexto Seguinte

Contexto seguinte	Apl./Total	%	PR
Haplologia (/t/, /d/, /n/)	93/124	75%	0,59
Consoantes	217/308	71%	0,58
Vogal	207/324	63,9%	0,44
Pausa	95/148	64,2%	0,37

Input 0,734; Significância = 0,007

Fonte: Elaborada pelo autor.

Os dados da Tabela 14 mostram que no contexto seguinte os casos de haplologia (PR 0,59) e de consoantes (PR 0,58) favorecem a aplicação da regra, tendo, ambos, pesos relativos bastante próximos. Já os contextos seguintes de vogal (PR 0,44) e pausa (PR 0,37) inibem o apagamento.

Segundo Oliveira (2017), quando a motivação para o apagamento de um segmento é a existência de outro semelhante no contexto seguinte, tem-se um processo de natureza geral e sem motivação morfológica. Com base nessa afirmação, Almeida e Oliveira (2017) afirmam que “a presença de tais consoantes [d, n, t] no contexto seguinte provoca o apagamento devido a um contexto fonético favorecedor, e não somente por se tratar de gerúndio” (p. 207). A partir desse ponto de vista, os autores sugerem que esses contextos sejam tratados de maneira especial na análise, conforme fizeram em estudo realizado na fala de Maceió, cujo resultado pode ser visualizado na Tabela 15.

Tabela 15. Apagamento de /d/ no morfema de gerúndio na fala de Maceió-AL: Contexto Seguinte

Contexto seguinte	Apl./Total	%	PR
Haplogia (/t/, /d/, /n/)	77/88	87,5	0,91
Consoantes	43/144	29,9	0,37
Vogal	85/168	5,06	0,57
Pausa	14/124	11,3	0,12

Fonte: Almeida e Oliveira (2017, p. 207) (adaptada).

Embora os resultados para a fala de Maceió indiquem alta probabilidade de apagamento em contextos de haplogia (PR 0,91), Almeida e Oliveira (2017) desconsideraram tal fator e assumiram que contextos de vogais é que favorecem o apagamento de /d/ no morfema *-ndo* naquela variedade de fala.

Na pesquisa sobre a fala de Feira de Santana, os dados do *corpus* não indicaram haver diferença estatística significativa entre os fatores haplogia (PR 0,59) e consoantes em geral (PR 0,58), como pôde ser visto na Tabela 14. Diante desse resultado para a variável ‘contexto seguinte’, realizou-se uma segunda rodada dos dados no GoldVarb X, amalgamando-se os fatores ‘haplogia’ e ‘consoantes’. Os resultados estão na Tabela 16.

Tabela 16. Apagamento de /d/ no morfema de gerúndio na fala de Feira de Santana: Contexto Seguinte (segunda rodada)

Contexto seguinte	Apl./Total	%	PR
Consoantes + haplogia	310/432	71,8%	0,59
Vogal	207/324	63,9%	0,44
Pausa	95/148	64,2%	0,37

Input 0,734; Significância 0,007

Fonte: Elaborada pelo autor.

Os dados da Tabela 16 confirmam a hipótese inicial de que contextos de consoantes motivam a aplicação da regra, mostrando que o apagamento é levemente favorecido quando há segmento consonantal e desfavorecido quando há segmento vocálico ou pausa no contexto seguinte. Resultados semelhantes a esses foram verificados por Nascimento, Araújo e Carvalho (2013) na fala popular de Fortaleza, o que aproxima as duas variedades no tocante ao condicionamento da regra variável quanto ao contexto seguinte.

4.2.1.2 Extensão do verbo

A variável ‘extensão do verbo’ permite observar se o número de sílabas do vocábulo (dissílabos, trissílabos, polissílabos) está relacionado com a aplicação da regra. Os exemplos seguintes ilustram a variável ‘extensão do verbo’.

Dissílabo

(51) *Aí depois eu acabei vino ... vino fazer vestibular em Feira, na UEFS, aí passei, aí fiquei, aí meio que tinha desistido, aí fiz.* (M1S30)

Trissílabo

(52) *Hoje, a gente se tratano de feira, a gente vê que os exemplos do Clériston Andrade, onde as pessoas vão sã e saem doentes ou morrem lá dentro. Justamente por falta de conhecimento.* (H2S56)

Polissílabo

(53) *Eu acho que o trote deveria ser rediscutido. Também num concordo com trote entregano alimento como o pessoal [do curso] de Pedagogia.* (H1S25)

O grupo de fatores ‘extensão do verbo’ mostrou-se estatisticamente relevante no estudo sobre a fala de Feira de Santana, sendo o quinto a ser selecionado pelo programa GoldVarb X. Os resultados estão na Tabela 17, a seguir.

Tabela 17. Apagamento de /d/ no morfema de gerúndio na fala de Feira de Santana: Extensão do Verbo

Extensão do verbo	Apl./Total	%	PR
Dissílabos	50/92	54%	0,40
Trissílabos	378/509	74%	0,56
Polissílabos	184/303	60%	0,43

Input 0,734; Significância = 0,007

Fonte: Elaborada pelo autor.

Os dados da Tabela 17 mostram que os verbos trissílabos favorecem levemente a aplicação da regra, enquanto os dissílabos e polissílabos mostram-se inibidores. Esses

resultados contrariam em parte a hipótese inicial de que trissílabos e polissílabos favorecessem o apagamento e que dissílabos o inibissem.

A variável ‘extensão do verbo’ foi observada na maioria dos estudos apresentados na seção 1.2 desta dissertação; nos casos em que se mostrou estatisticamente significativa, os resultados, em geral, apontaram que verbos com duas sílabas inibem a aplicação da regra e que verbos com três ou mais sílabas favorecem-na (cf. MOLLICA; MATTOS 1992; ALMEIDA; OLIVEIRA, 2017; CARDOSO; PINHEIRO; SILVA, 2019).

A partir dos resultados apresentados para a variável ‘extensão do verbo’, constata-se que esse grupo de fatores condiciona a regra variável de maneira singular na fala feirense. Esse resultado suscita uma investigação mais apurada dos dados para verificar os efeitos de frequência de tipo e de dado⁵ no *corpus*, por exemplo. Pelo que se propõe nesta pesquisa, cabe, por hora, afirmar que na fala feirense os verbos trissílabos favorecem e que os dissílabos e polissílabos desfavorecem a aplicação da regra de apagamento de /d/ no morfema *-ndo*.

4.2.3 Análise das variáveis sociais

Até aqui foram apresentados os resultados gerais sobre a taxa de aplicação da regra de apagamento do /d/ no morfema *-ndo* na fala de Feira de Santana bem como foram apresentados e discutidos os resultados encontrados para as variáveis linguísticas estatisticamente significativas. A partir de agora serão apresentados os resultados encontrados para as variáveis sociais. Os três grupos de fatores extralinguísticos controlados mostraram-se importantes no estudo sobre o apagamento de /d/ em morfema de gerúndio no *corpus* estudado, sendo que a escolaridade, o sexo e a faixa etária foram selecionados, respectivamente, na primeira, segunda e quarta posição em ordem de relevância estatística. Os resultados são apresentados a seguir.

4.2.3.1 Escolaridade

A escolaridade foi a variável estatisticamente mais significativa na pesquisa, sendo a primeira a ser selecionada pelo GoldVarb X na análise computacional. Foram controlados os fatores ‘nível fundamental I’ e ‘nível universitário’ e os resultados estão resumidos na Tabela 18, a seguir.

⁵ Sobre os efeitos de frequência de tipo (*type frequency*) e frequência de dado (*token frequency*) em mudanças, ver Bybee (2020).

Tabela 18. Apagamento de /d/ no morfema de gerúndio na fala de Feira de Santana:
Escolaridade

Fatores	Apl./Tot.	%	PR
Nível Fundamental I	374/425	88%	0,77
Nível Universitário	238/479	50%	0,25

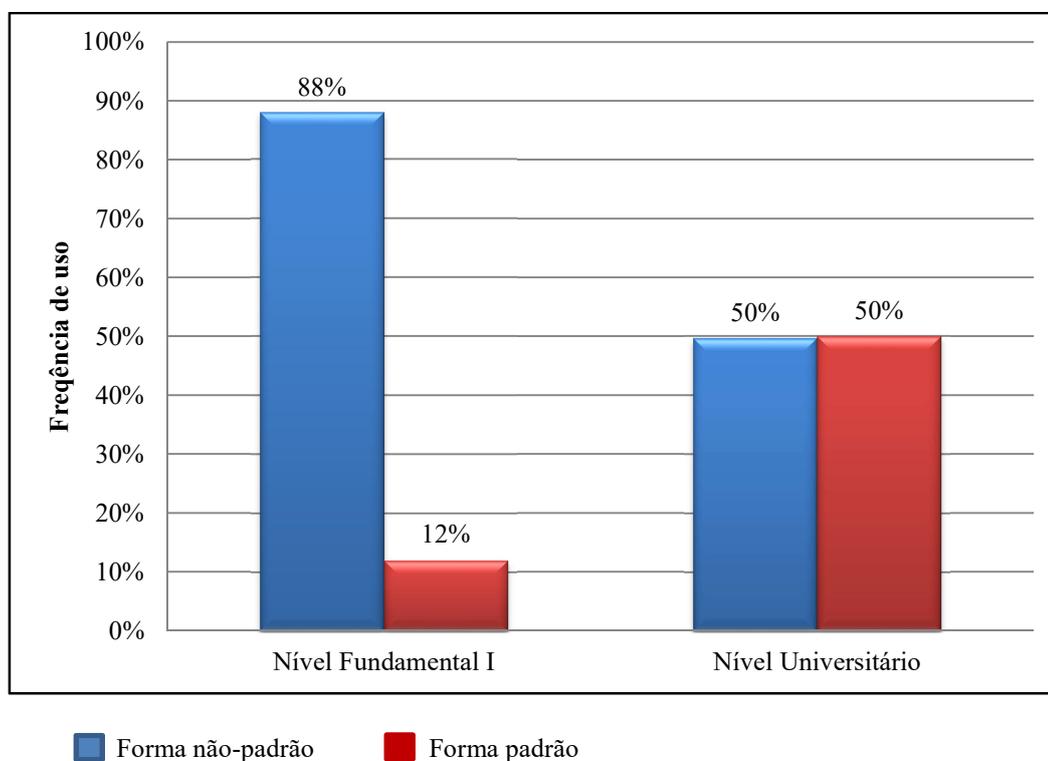
Input 0,734; Significância = 0,007

Fonte: Elaborada pelo autor.

Com base nos dados da Tabela 18, constata-se que a aplicação da regra é fortemente motivada por informantes que estudaram até o Nível Fundamental 1 (PR 0,77) e fortemente inibida por aqueles com Nível Universitário (PR 0,25). Esse resultado confirma a hipótese inicial de que o apagamento de /d/ no morfema de verbos no gerúndio é mais comum na fala dos menos escolarizados.

O Gráfico 3 ilustra o percentual de aplicação da regra de acordo com a variável ‘nível de escolaridade’ no *corpus* estudado.

Gráfico 3. Frequência de apagamento de /d/ em morfema de gerúndio: Escolaridade



Fonte: Elaborado pelo autor.

O Gráfico 3 indica que as formas variantes do tipo *cantando* x *cantano* competem de maneiras diferentes nas variedades popular e culta de Feira de Santana, sendo que a forma inovadora parece, na verdade, ser a variante *default* na norma popular e que a competição entre as formas variantes se mostra equilibrada na norma culta. O fato de os falantes de nível universitário terem mais contato com a língua escrita provavelmente estimula o uso da forma padrão, como em *cantando* ao lado da forma inovadora, como em *cantano*.

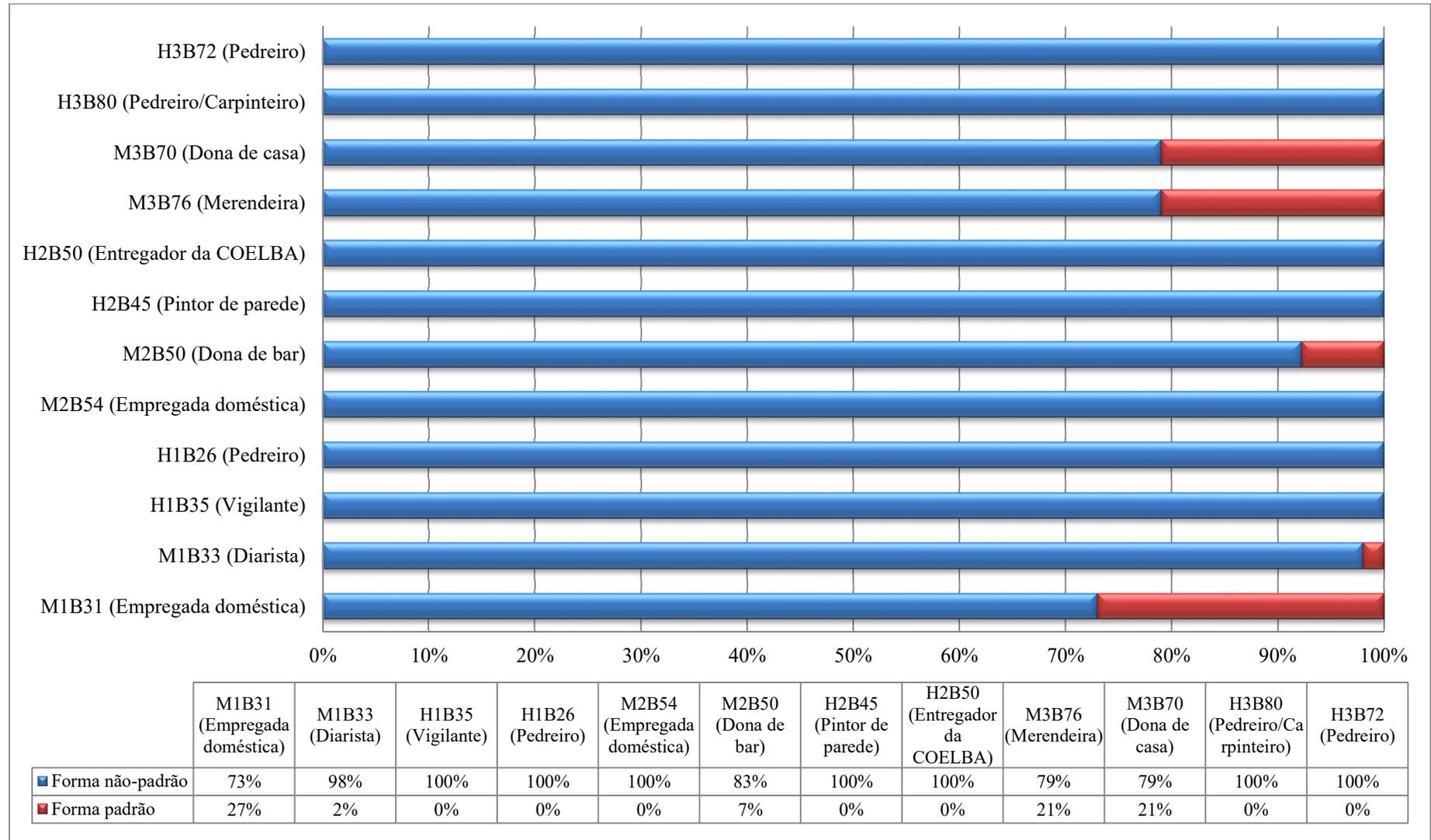
Araújo e Mota (2004) observaram que o falante faz juízo de valor entre as formas variantes de realização do gerúndio, atribuindo maior valor à forma padrão e menor valor à forma não-padrão, sendo que, no estudo realizado pelas autoras, a taxa de aplicação decresceu sempre em estilos mais monitorados de fala.

As frequências de uso e pesos relativos encontrados no *corpus* de Feira de Santana (nível fundamental I, PR 0,77 e nível universitário, PR 0,25) apontam para a existência de prestígio da forma padrão em relação à forma inovadora na fala feirense.

Bortoni-Ricardo (2004, p. 48) argumenta que “os anos de escolarização de um indivíduo e a qualidade das escolas que frequentou também têm influência em seu repertório sociolinguístico”. Nessa mesma linha, Votre (2013) chama a atenção para o papel que exerce a escola como mantenedora de formas tidas como de prestígio. Segundo o autor, como “veículo de familiarização com a literatura nacional, a escola incute gostos, normas, padrões estéticos e morais em face da conformidade de dizer e escrever” (VOTRE, 2013, p. 51). A interpretação dos dados de Feira de Santana para a variável ‘escolaridade’ ilustra bem essa relação estreita entre o grau de instrução de um indivíduo e seu repertório sociolinguístico.

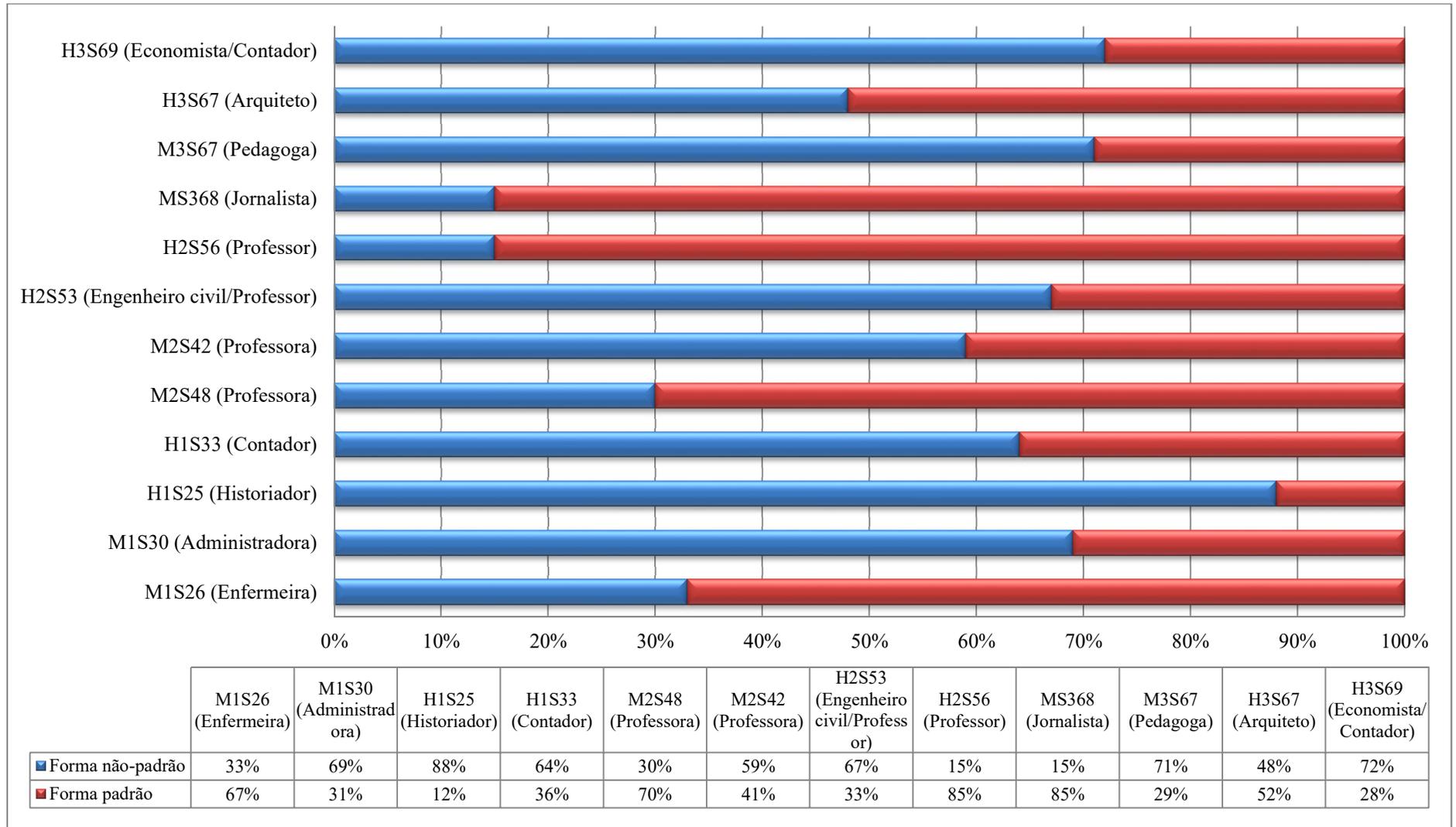
Os dados da variável ‘escolaridade’ indicam que a forma padrão está relacionada ao alto nível de escolarização. Contudo o valor percentual de aplicação da forma não-padrão na norma culta (50%) instiga a busca por dados que permitam a melhor compreensão de como as formas variantes estão distribuídas na amostra. De tal forma, buscou-se verificar, sobretudo entre os falantes cultos, se as ocorrências poderiam estar relacionadas com a profissão, visto que ambientes de trabalho que envolvem eventos de letramento e que requerem alto grau de monitoramento da fala tendem a motivar o uso de formas linguísticas pertencentes à variedade padrão. Esses dados podem ser visualizados nos gráficos 4 e 5, a seguir.

Gráfico 4. Frequência de apagamento de /d/ em morfema de gerúndio por informante: norma popular



Fonte: Elaborado pelo autor.

Gráfico 5. Frequência de apagamento de /d/ em morfema de gerúndio por informante: norma culta



Fonte: Elaborado pelo autor.

Como se pode observar no Gráfico 4, a forma conservadora foi documentada na fala de apenas cinco informantes da norma popular, o que confirma que a variante não-padrão é *default* na fala de indivíduos que tiveram pouco contato com a escola. Buscou-se, a partir da reescuta das entrevistas sociolinguísticas, levantar dados que pudessem dar pistas de como a forma padrão teria entrado no repertório das três informantes que apresentaram maior índice de manutenção de /d/ no morfema *-ndo* (a dona de casa, a merendeira e a empregada doméstica).

Infere-se que a presença da forma variante padrão na fala das três informantes populares resulta do contato delas com falantes cultos em seu ambiente de trabalho ou resulta de sua rede de relacionamento social. É possível que a empregada doméstica (M1B31) e a merendeira (M3B76) tenham adquirido a forma padrão nos seus repertórios através do contato em seus ambientes de trabalho, onde teriam convivido com falantes cultos, pois a primeira trabalhou como representante comercial e a segunda conviveu por muito tempo em um ambiente escolar, além de ir à igreja e gostar de ler a Bíblia. Quanto à dona de casa (M3B70), trata-se de uma senhora ativa que frequenta atividades formativas voltadas ao público da terceira idade, além de ir à igreja e também ler bastante a Bíblia.

Já o Gráfico 5 mostra que, na norma culta, ao contrário do que acontece na norma popular, todos os informantes fazem uso das duas formas variantes. Esperava-se que informantes cujas profissões estivessem diretamente relacionadas com ambientes de letramento, como professor, pedagogo, historiador e jornalista, fossem mais conservadores. Mas isso não se confirmou, sobretudo para a pedagoga e para o historiador. Sobre a profissão de professor, embora o professor (H2S56) tenha apresentado taxa de apenas 15% de apagamento, os professores (M2S42) e (H2S53) apresentaram taxas de 59% e 65%, respectivamente, indicando não haver relação direta entre ser professor fazer mais ou menos uso de uma das duas formas em variação. A jornalista mostrou-se conservadora, mas, diferentemente do que se pôde observar entre os professores, não havia outro informante jornalista na amostra para que se pudesse verificar se esse comportamento se repetiria, o que traria mais detalhe à análise.

Essa discussão sobre a profissão dos informantes procurou trazer mais detalhes sobre os dados com relação à variável ‘escolaridade’. Como se disse anteriormente, os pesos relativos polarizados entre os fatores ‘nível fundamental 1’ (PR 0,77) e ‘nível universitário’ (PR 0,25) parecem indicar prestígio da forma padrão sobre a forma não-padrão, mas as elevadas taxas de aplicação entre falantes que lidam com a língua escrita no dia a dia das suas profissões sugerem que a forma inovadora não é estigmatizada na comunidade feirense.

Cabe ainda dizer que os resultados obtidos para a variável ‘escolaridade’ na fala de Feira de Santana se coadunam com outros mostrados na maioria dos estudos apresentados anteriormente nesta dissertação, nos quais essa variável foi controlada (FEREIRA, 2010; NASCIMENTO; ARAÚJO; CARVALHO, 2013; GONÇALVES, 2018; MACEDO et al., 2018; CARDOSO; PINHEIRO; SILVA, 2019; ALMEIDA; PEREIRA; ARAÚJO, 2020). À semelhança do que ocorreu na pesquisa com dados de Feira de Santana, a variável ‘escolaridade’ foi também a primeira a ser selecionada em Nascimento, Araújo e Carvalho (2013) ao investigarem o mesmo fenômeno na fala de Fortaleza, o que aproxima as duas variedades no tocante à maneira como a variável dependente é condicionada pelo nível de escolarização dos falantes.

4.2.3.2 Sexo

A variável ‘sexo’ foi a segunda selecionada pelo programa estatístico GoldVarb X. A hipótese inicial para esse grupo de fatores era a de que o apagamento fosse motivado por indivíduos do sexo masculino, conforme foi identificado na maioria dos estudos revisitados nesta dissertação. Constata-se, a partir dos pesos relativos constantes da Tabela 19, que o apagamento é favorecido pelos homens (PR 0,61) e inibido pelas mulheres (PR 0,40), o que confirma a hipótese inicial de pesquisa.

Tabela 19. Apagamento de /d/ no morfema de gerúndio na fala de Feira de Santana: Sexo

Sexo	Apl./Total	%	PR
Mulher	305/488	63	0,41
Homem	307/416	74	0,61

Input 0,734; Significância = 0,007

Fonte: Elaborada pelo autor.

O sexo aparece como sendo uma variável significativa na explicação do fenômeno de apagamento de /d/ em morfema gerúndio na maioria dos trabalhos apresentados na seção 1.2 desta dissertação (ARAÚJO; MOTA, 2004; LUCENA; VASCONCELOS, 2007; FERREIRA, 2010; VIEIRA, 2011; ARAÚJO; NASCIMENTO; CARVALHO, 2013; ALMEIDA; OLIVEIRA, 2017; GONÇALVES, 2018; ARAÚJO, 2019; ALMEIDA; PEREIRA; ARAÚJO, 2020), sendo que os homens aparecem quase sempre como favorecedores do apagamento. A única exceção dentre as pesquisas resenhadas no primeiro capítulo deste texto é encontrada em Araújo, Nascimento e Carvalho (2013), para a fala de Fortaleza-CE, em que

as mulheres mostraram-se favorecedoras do apagamento e os homens foram desfavorecidos.

Estudos em Sociolinguística têm mostrado que homens e mulheres se comportam de maneiras diferentes no que concerne à variação e à mudança linguística, o que não deve ser de grande admiração, já que, sendo a língua um fenômeno inerentemente social, espera-se que ela venha a refletir também as relações sociais existentes entre homens e mulheres. Paiva (2013) argumenta que

A análise da dimensão social da variação e da mudança linguística não pode ignorar [...] que a maior ou menor ocorrência de certas variantes, principalmente daquelas que envolvem o binômio forma padrão/forma não padrão e o processo de implementação de mudanças estejam associadas ao gênero/sexo do falante e à forma de construção social dos papéis feminino e masculino. (PAIVA, 2013, p. 35)

Nesse mesmo sentido, Oliveira (2006, p. 46) afirma que “do ponto de vista social, homens e mulheres têm papéis diferentes e estão, portanto, expostos a situações diversas. Assim, esse fator pode influenciar a escolha de uma ou outra forma linguística”.

A análise dos dados extraídos da amostra de fala de Feira de Santana indica maior sensibilidade à forma padrão por parte das mulheres, enquanto os homens parecem ser mais abertos à forma inovadora.

4.2.3.3 Faixa etária

A variável ‘faixa etária’ foi a quarta variável independente a ser selecionada pelo programa estatístico GoldVarb X. No estudo sobre o apagamento de /d/ no morfema de gerúndio na fala de Feira de Santana, foram controladas três faixas etárias, faixa 1 (de 25 a 35 anos), faixa 2 (de 45 a 55 anos) e faixa 3 (acima de 65 anos). A hipótese inicial foi a de que falantes da faixa 1 apagassem mais do que aqueles das faixas 2 e 3, visto que os mais jovens são mais sensíveis a formas novas na língua (LABOV, 2008 [1972]). Os resultados para esse grupo de fatores estão na Tabela 20.

Tabela 20. Apagamento de /d/ no morfema de gerúndio na fala de Feira de Santana: Faixa Etária

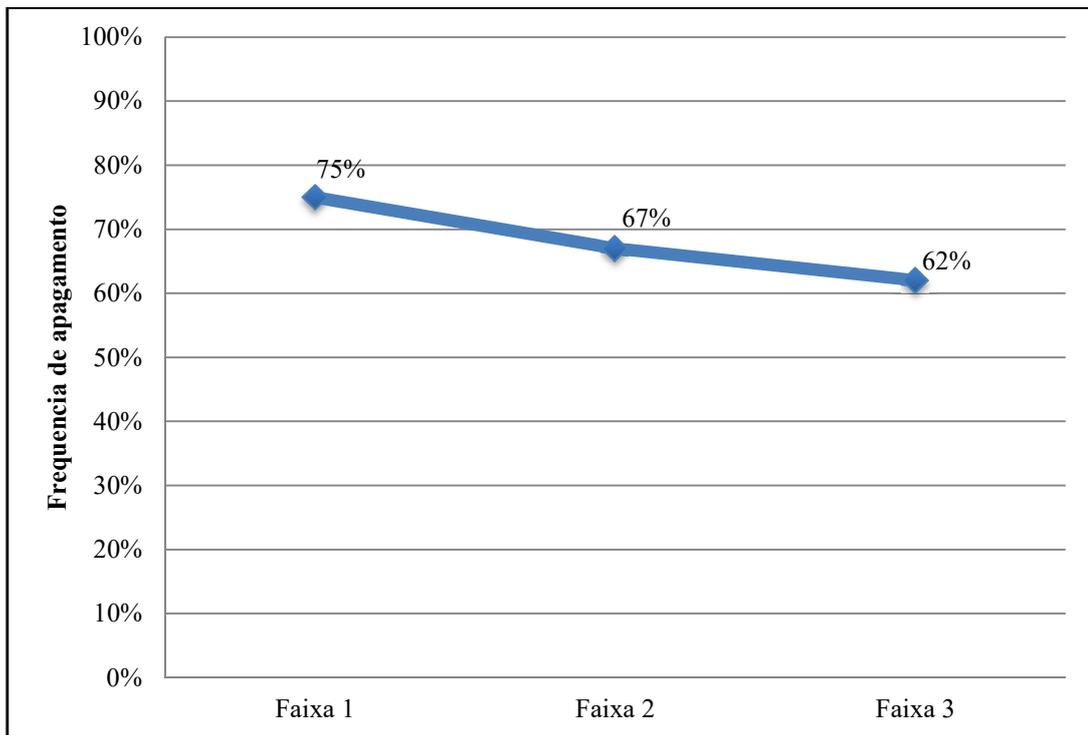
Faixa etária	Apl./Total	%	PR
25 a 35 anos	286/382	75	0,60
45 a 55 anos	161/256	67	0,44
65 anos ou mais	165/266	62	0,41

Input 0,734; Significância = 0,007

Fonte: Elaborada pelo autor.

Os pesos relativos atribuídos a cada fator do grupo (faixa 1, PR 0,60; faixa 2, PR 0,44; faixa 3, PR 0,41) mostram que os mais jovens favorecem a aplicação da regra e que os mais velhos a inibem, o que confirma a hipótese inicial de pesquisa. O Gráfico 4, a seguir, ilustra como o uso da forma não-padrão diminui à medida que aumenta a idade dos informantes.

Gráfico 6. Frequência percentual de apagamento por Faixa Etária



Fonte: Elaborado pelo autor.

De acordo com os estudos apresentados na seção 1.2 desta dissertação, a variável ‘faixa etária’ também condiciona a regra variável na fala de São José do Rio Preto (FERREIRA, 2010) de maneira semelhante a esta apresentada na Tabela 20. Nas duas

variedades, o uso da forma inovadora é liderado pelos mais jovens e decresce até a faixa dos mais idosos, como se pode observar na Tabela 21.

Tabela 21. Apagamento de /d/ no morfema de gerúndio: comparação entre Feira de Santana-BA e São José do Rio Preto-SP

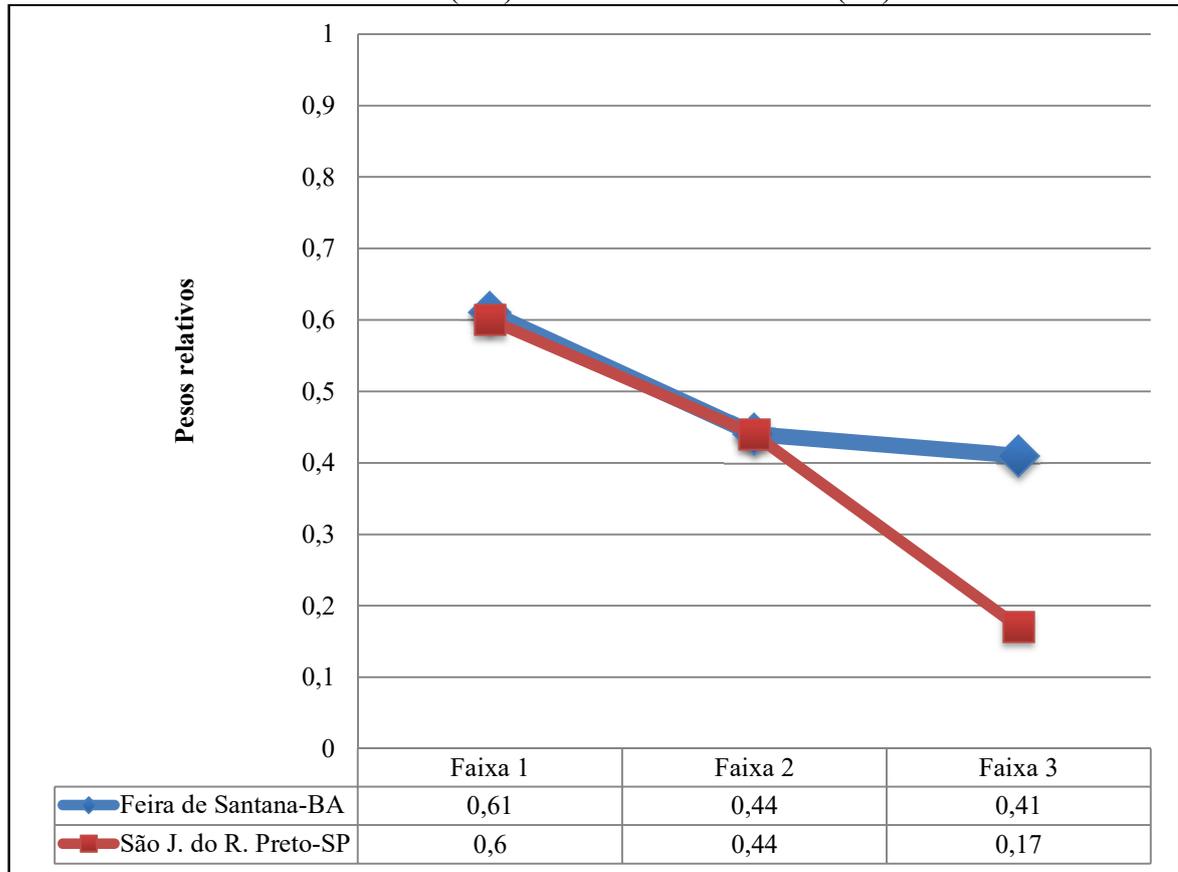
Faixa Etária	SJRP-SP			Feira de Santana-BA		
	Apl./Tot.	%	PR	Apl./Tot.	%	PR
1	138/164	79	0,64	286/382	75	0,60
2	130/202	63	0,44	161/256	67	0,44
3	51/143	35	0,17	165/266	62	0,41

Fonte: Elaborada pelo autor partir dos dados de Ferreira (2010, p. 103).

A comparação entre estudos realizados com diferentes *corpora* demanda sempre a dificuldade de lidar com diferenças metodológicas relacionadas às estratificações de bancos de dados diversos. No estudo de Ferreira (2010), a amostra foi estratificada em 5 faixas etárias (7 a 15 anos, 16 a 25 anos, 26 a 35 anos, 36 a 55 anos, mais de 55 anos). A fim de possibilitar a comparação entre as duas variedades, as duas primeiras faixas em Ferreira (2010) foram descartadas.

A Tabela 21 mostra que nas duas variedades as taxas de aplicação da regra de apagamento apresentam valores percentuais e pesos relativos inversamente proporcionais à idade dos informantes. Ainda que os pesos sejam menos acentuados entre a faixa intermediária e a dos mais velhos na fala feirense, pode-se afirmar que, no tocante à variável ‘faixa etária’, os resultados encontrados nos dados de Feira de Santana se coadunam com aqueles encontrados por Ferreira (2010), a partir dos quais a autora concluiu tratar-se de mudança em curso na variedade falada em São José do Rio Preto. O Gráfico 7 mostra a comparação entre as duas variedades.

Gráfico 7. Probabilidade de apagamento em função da Faixa Etária: comparação entre Feira de Santana (BA) e São José do Rio Preto (SP)



Fonte: Elaborado pelo autor.

De acordo com a representação gráfica acima, pode-se perceber que nas duas variedades os falantes da primeira e da segunda faixa etária se comportam de maneira semelhante, sendo que na variedade feirense a probabilidade de uso da forma inovadora na faixa 3 é bem maior do que na fala riopretense.

Na ausência de dados diacrônicos, a variável 'faixa etária' permite ao pesquisador inferir por meio de estudo em tempo aparente se formas em variação configuram uma mudança linguística em curso ou um caso de variação estável (cf. LABOV, 1972). A interpretação dos dados de Feira de Santana aponta para uma mudança em curso em tempo aparente, pois falantes mais jovens tendem a usar mais a forma inovadora e os mais idosos tendem a ser mais conservadores.

4.3 CRUZAMENTOS DAS VARIÁVEIS SOCIAIS

Após as análises multivariadas dos dados do *corpus* da pesquisa, foi realizado o cruzamento das variáveis sociais para maior aprofundamento sobre como elas interagem entre si nos dados de fala de Feira de Santana. Foram realizados os cruzamentos ‘faixa etária’ x ‘escolaridade’, ‘faixa etária’ x ‘sexo’ e ‘escolaridade’ x ‘sexo’. Os resultados estão apresentados a seguir, iniciando-se pelos dados de ‘faixa etária’ x ‘escolaridade’, dispostos na Tabela 22.

Tabela 22. Apagamento de /d/ no morfema de gerúndio na fala de Feira de Santana: cruzamento Faixa Etária e Escolaridade

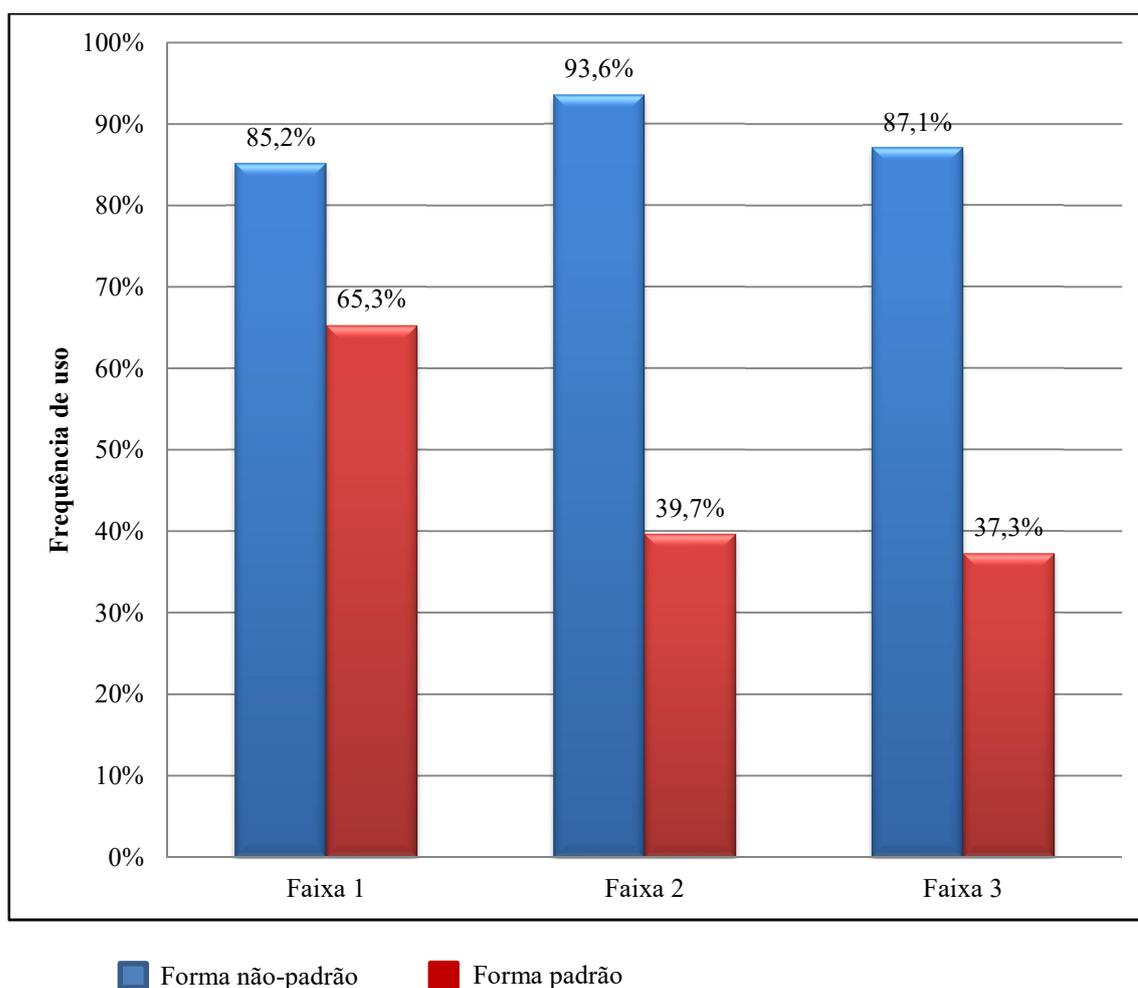
Escolaridade x Faixa Etária	Nível Fundamental I			Nível Universitário		
	Apl./Tot.	%	PR	Apl./Tot.	%	PR
25 a 35 anos	156/183	85,2	0,78	130/199	65,3	0,40
45 a 55 anos	103/110	93,6	0,87	58/146	39,7	0,17
65 anos ou mais	115/132	87,1	0,75	50/134	37,3	0,17

Input 0,737; Significância 0,006

Fonte: Elaborada pelo autor.

Os dados do cruzamento entre as variáveis ‘faixa etária’ e ‘escolaridade’ mostram frequências acima de 85% nas três faixas de idade entre os informantes que estudaram somente até o Ensino Fundamental I (faixa 1: 85%, faixa 2: 94% e faixa 3: 87%), sendo que a faixa intermediária apresenta taxa um pouco mais alta em relação às outras duas. Os pesos relativos exibem essa mesma configuração (faixa 1: PR 0,78, faixa 2: PR 0,87, faixa 3: PR 0,75), apresentando também valor levemente mais alto a faixa intermediária.

Para os falantes de nível universitário, os números percentuais se apresentam inversamente proporcionais à idade dos informantes (faixa 1: 65%, faixa 2: 40%, faixa 3: 37%), exibindo uma queda relativamente acentuada entre a faixa 1 e as demais e uma queda muito pequena, apenas 2,4 pontos percentuais, da faixa intermediária para a faixa dos mais velhos. Entre os informantes de nível universitário, os pesos relativos foram 0,40 para a faixa 1, indicando que, dentre os mais escolarizados, os mais jovens inibem levemente a aplicação da regra; e 0,17 para as faixas 2 e 3, indicando que esses falantes inibem fortemente a aplicação da regra. Os resultados do cruzamento entre as variáveis ‘faixa etária’ e ‘escolaridade’ podem ser visualizados no Gráfico 8, a seguir.

Gráfico 8. Cruzamento das variáveis Faixa etária e Escolaridade

Fonte: Elaborado pelo autor.

A interpretação dos dados do cruzamento entre a faixa etária e a escolaridade dos informantes indica que os menos escolarizados, no geral, motivam fortemente o apagamento e que os mais escolarizados o inibem, sendo que, dentre estes últimos, os mais jovens inibem levemente e os mais velhos inibem fortemente o apagamento. Se os dados da variável 'faixa etária' apontam para uma mudança em curso, o cruzamento desta com a variável 'escolaridade' confirma que o alto nível de escolarização atua como um mecanismo refreador do processo de mudança, independentemente da idade dos informantes, ou seja, se realmente se trata de uma mudança, esta é de baixo para cima (cf. LABOV, 1982), localizando-se entre os menos escolarizados.

Após a verificação de como se dá a interação entre a faixa etária e a escolaridade dos informantes no tocante ao apagamento de /d/ no morfema de verbos no gerúndio na fala de

Feira de Santana, são apresentados, a seguir, os dados do cruzamento entre as variáveis 'faixa etária' e 'sexo'. Os dados estão na Tabela 23.

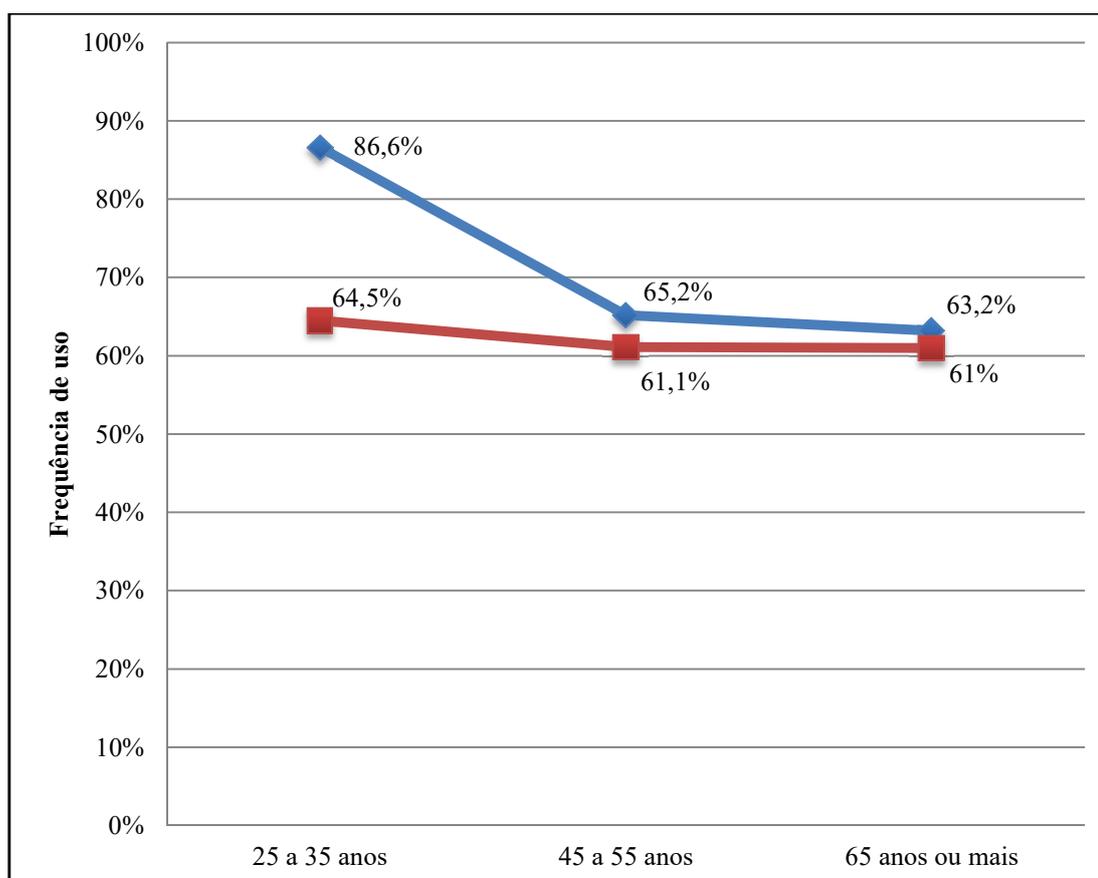
Tabela 23. Apagamento de /d/ no morfema de gerúndio na fala de Feira de Santana
cruzamento Faixa Etária e Sexo

Faixa etária	Homens			Mulheres		
	Apl./Tot.	%	PR	Apl./Tot.	%	PR
25 a 35 anos	155/179	86,6	0,80	131/203	64,5	0,41
45 a 55 anos	73/112	65,2	0,47	88/144	61,1	0,39
65 anos ou mais	79/125	63,2	0,46	86/141	61	0,37

Input 0,737; Significância 0,006

Fonte: Elaborada pelo autor.

Os números da Tabela 23 revelam que a taxa de aplicação entre os homens, em função da idade (faixa 1: 86,6%, faixa 2: 65,2%, faixa 3: 63,2%), assim como os pesos relativos (faixa 1: PR 0,80, faixa 2: PR 0,47, faixa 3: PR 0,46), apresentam valores polarizados entre a primeira faixa e as faixas intermediária e dos mais velhos, sendo que não há diferença relevante entre as duas últimas faixas nem em termos percentuais nem em pesos relativos. Quanto aos dados para o sexo feminino em função da idade, verifica-se que há quase constância no uso da forma inovadora nas três faixas etárias em termos percentuais, embora os pesos relativos apontem para a não aplicação da regra de apagamento entre as mulheres de todas as idades. Pode-se observar no Gráfico 9, a seguir, uma linha praticamente horizontal, mas indicando um leve decréscimo entre a primeira e as demais faixas etárias.

Gráfico 9. Cruzamento das variáveis Faixa etária e Sexo

Fonte: Elaborado pelo autor.

A interpretação dos dados do cruzamento entre a faixa etária e o sexo dos informantes confirma tratar-se do início de um processo de mudança que está sendo puxado por falantes mais jovens e do sexo masculino, ratificando o que se observou anteriormente na análise multivariada para essas duas variáveis sociais. Finalizando a análise dos cruzamentos entre as variáveis sociais, apresentam-se os dados do cruzamento entre ‘escolaridade’ e ‘sexo’, na Tabela 24, a seguir.

Tabela 24. Apagamento de /d/ no morfema de gerúndio na fala de Feira de Santana: cruzamento Escolaridade e Sexo

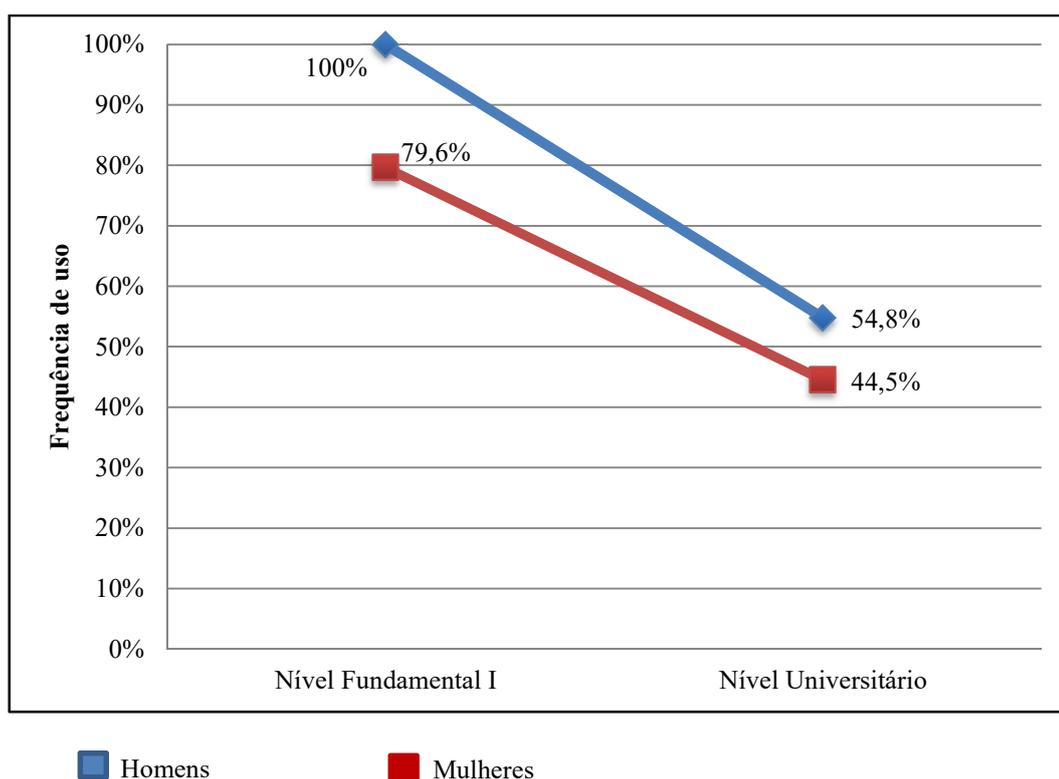
Faixa etária	Homens			Mulheres		PR
	Apl./Tot.	%	PR	Apl./Tot.	%	
Nível Fundamental I	175/175	100	1	199/250	79,6	0,73
Nível Universitário	132/241	54,8	0,42	106/238	44,5	0,34

Input 0,647; Significância 0,005

Fonte: Elaborada pelo autor.

Os dados da Tabela 24 mostram valores categóricos para os informantes do sexo masculino que estudaram até Ensino Fundamental I, isso significa que não houve nenhuma ocorrência da forma padrão de realização do gerúndio na fala desse grupo de informantes. Os dados indicam percentual de 54,8% de aplicação e peso relativo 0,42 para os homens de nível universitário, percentual de 79,6% e PR 0,73 para mulheres com nível fundamental I e percentual de 44,5% e PR 0,34 para mulheres de nível universitário. Esses valores percentuais e probabilísticos mostram a importância da variável ‘escolaridade’ sobre a variável ‘sexo’ na compreensão do fenômeno variável estudado, pois, se excluídos os informantes categóricos da rodada, as mulheres com nível fundamental I foram o fator que mais favoreceu a aplicação da regra. O Gráfico 10 ilustra o cruzamento entre as variáveis ‘sexo’ e ‘escolaridade’ para a amostra de fala de Feira de Santana.

Gráfico 10. Cruzamento das variáveis Escolaridade e Sexo – Apagamento



Fonte: Elaborado pelo autor.

Percebe-se a partir da leitura do Gráfico 10 que homens e mulheres comportam-se de maneira semelhante com relação ao fenômeno observado, uma vez que as linhas são paralelas. Entretanto os homens, independentemente do nível de escolaridade, favorecem mais o apagamento do que as mulheres. As duas linhas são descendentes, indicando que ambos os

sexos apresentam maior frequência de uso da forma não-padrão quando associados ao nível mais baixo de instrução e que essas frequências caem quando aumenta o grau de instrução.

4.4 VARIÁVEIS DESCARTADAS NA ANÁLISE ESTATÍSTICA

A apresentação de variáveis não significativas estatisticamente é importante na pesquisa variacionista na medida em que permite que se façam posteriores comparações entre estudos anteriores e adicionais sobre o mesmo fenômeno. Além disso, é preciso considerar que a não seleção decorre, às vezes, da má distribuição dos dados na amostra analisada. Segundo Guy e Zilles (2007), o fato de uma variável controlada não ser estatisticamente relevante implica que ela não condiciona o fenômeno no mundo real, e essa é uma informação relevante para o pesquisador. De tal modo, apresentam-se, nesta seção, as variáveis linguísticas descartadas na investigação sobre o apagamento de /d/ em morfema de gerúndio na fala de Feira de Santana, a saber, ‘conjugação verbal’ e ‘estrutura sintática’.

4.4.1 Conjugação verbal

A variável ‘conjugação verbal’ foi controlada com o intuito de observar se a aplicação da regra de apagamento estaria sendo condicionada pela vogal temática que antecede o morfema *-ndo*, caracterizando as 1ª, 2ª e 3ª conjugações, conforme está exemplificado a seguir.

1ª conjugação

(54) *E aí a gente percebe que não houve um avanço, mesmo trocano de partido, né?* (M2S42)

2ª conjugação

(55) *Eu acho que todo mundo deveria estudar, mas estudar de verdade, não é o que a gente vê acontecendo hoje nesses cursos que o estado implementou, e empurra goela abaixo do povo que são os cursos de fluxo e aceleração.* (H2S56)

3ª conjugação

(56) *A gente tá produzindo cinema, o cinema hoje, vamos dizer, é algo que é, faz parte do meu dia a dia.* (H1S25)

A variável ‘conjugação verbal’ foi controlada por Lucena e Vasconcelos (2007), Vieira (2011), Nascimento, Araújo e Carvalho (2013), Almeida e Oliveira (2017), Macedo et al. (2018), Gonçalves (2018), Araújo (2019) e Cardoso e Pinheiro (2019). Em todas essas pesquisas, com exceção da realizada por Vieira (2011), os dados foram submetidos a análise probabilística computacional, sendo que a variável ‘conjugação verbal’ foi significativa na fala do Brejo Paraibano-PB, com a 2ª conjugação favorecendo o apagamento (cf. LUCENA; VASCONCELOS, 2007), na fala de Fortaleza-CE, em que as 1ª e a 2ª conjugações foram favorecedoras (cf. NASCIMENTO; ARAÚJO; CARVALHO, 2013), e na fala de Aracaju-SE, também com a 1ª e a 2ª conjugações favorecendo o apagamento (cf. CARDOSO; PINHEIRO; SILVA, 2019).

Como se pode notar a partir das informações acima, dos oito trabalhos em que a conjugação verbal foi controlada, três levaram à conclusão de que esse grupo de fatores influencia a regra variável de apagamento da oclusiva /d/ no morfema de gerúndio nos *corpora* estudados, com a primeira e a segunda conjugações favorecendo o apagamento e a terceira conjugação o inibindo. Foram esses trabalhos que deram base para assumir hipótese a inicial de que verbos da 1ª e 2ª conjugações fossem favorecedores da regra de apagamento na fala de Feira de Santana, hipótese esta que não se confirmou na pesquisa, visto que a conjugação verbal não se mostrou estatisticamente relevante e as frequências de aplicação da regra variável são bastante próximas, como se pode observar na Tabela 25, a seguir.

Tabela 25. Apagamento de /d/ no morfema de gerúndio na fala de Feira de Santana:
Conjugação Verbal

Fatores	Apl./Tot.	%
1ª conjugação	380/554	67%
2ª conjugação	176/272	65%
3ª conjugação	56/78	71%

Fonte: Elaborada pelo autor.

Os valores das taxas de aplicação são de 74%, 66% e 75% para os fatores 1ª, 2ª e 3ª conjugação, respectivamente. Esses são valores percentuais muito próximos e indicam, como se constatou por meio da análise estatística computacional, que a variável não tem relação com o apagamento de /d/ no morfema *-ndo* no *corpus*. Esse resultado é semelhante àqueles alcançados maioria dos estudos sobre o mesmo fenômeno, apresentados na seção 1.2 desta dissertação. Vale ressaltar o evidente desequilíbrio na distribuição dos dados pelos três

fatores, o que certamente não decorre da especificidade da amostra analisada, pois é notória a prevalência na língua portuguesa de verbos da 1ª conjugação em detrimento das demais.

4.4.2 Estrutura sintática

Ao controlar a variável ‘estrutura sintática’ na fala de São José do Rio Preto-SP, Ferreira (2010) constatou que contextos de perífrase favorecem o apagamento de /d/ no morfema de gerúndio. Segundo a autora, as formas de gerúndio ocorrem basicamente em três estruturas sintáticas: justaposição, orações reduzidas e perífrases, conforme se exemplificam a seguir:

Justaposição:

(57) *E foi se debilitano, debilitano, quando chegou um dia que a, a irmã dele chegou em casa pra... que tomava conta da casa, ele tava caído.* (M3S69)

Oração reduzida:

(58) *Minha vida foi assim, apertada, muita luta, cuidano de filho, cuidano de casa.* (M3B79)

Perífrase:

(59) *Oxe, mar menino num tô nem ligano #, saio mermo.* (M2B50)

Com base no estudo de Ferreira (2010), decidiu-se controlar a variável ‘estrutura sintática’ no estudo realizado sobre a fala de Feira de Santana, sob a hipótese de que contextos sintáticos de perífrase favorecessem o apagamento de /d/ no morfema *-ndo*. A hipótese inicial não se confirmou, uma vez que a variável foi descartada pelo GoldVarb X durante a análise estatística. A Tabela 26 apresenta os valores percentuais encontrados no *corpus* pesquisado.

Tabela 26. Apagamento de /d/ no morfema de gerúndio na fala de Feira de Santana: Estrutura Sintática

Estrutura sintática	Apl./Total	%
Justaposição	25/30	80%
Perífrase	362/532	68%
Oração reduzida	225/342	66%

Fonte: Elaborada pelo autor.

As taxas percentuais de aplicação mostram que o fator ‘justaposição’ se destaca em relação aos fatores ‘perífrase’ e ‘oração reduzida’, o que poderia levar à conclusão de que esse fosse o contexto sintático mais propício ao apagamento. No entanto é preciso levar em conta que a quantidade de dados de justaposição é muito pequena diante dos números encontrados para os outros dois fatores do grupo, correspondendo a apenas 3,3% do total de ocorrências no *corpus*, e por isso, muito provavelmente, a variável foi descartada na análise estatística pelo programa GoldVarb X.

A partir da interpretação dos resultados estatísticos apresentados neste capítulo, pode-se dizer então que atuam na aplicação da regra variável de apagamento do /d/ em formas de gerúndio na fala de Feira de Santana fatores de ordem social (escolaridade, sexo e faixa etária dos informantes) e de ordem estrutural (contexto seguinte e extensão do verbo). Trata-se, pois, de um fenômeno condicionado social e estruturalmente. Embora tenham sido descartadas as outras duas variáveis controladas na pesquisa (conjugação verbal e estrutura sintática), não se pode abandonar as hipóteses de pesquisa levantadas, mas apenas não confirmá-las por duas razões: no caso da conjugação verbal, é preciso considerar a distribuição dos verbos pelas três conjugações na própria língua; no caso da estrutura sintática, talvez uma amostra ampliada modifique os resultados aqui encontrados. Finda a apresentação da análise dos dados da pesquisa, passa-se às considerações finais deste trabalho.

CONSIDERAÇÕES FINAIS

Esta pesquisa de Mestrado teve como objetivo descrever o fenômeno variável do apagamento da consoante oclusiva /d/ no morfema de verbos no gerúndio na fala de Feira de Santana-BA. O referencial teórico e metodológico que embasou a pesquisa foi o da Sociolinguística Variacionista, dentro das proposições de Weinreich, Labov e Herzog (2006 [1968]) e Labov (2008 [1972]), que considera que a língua é um sistema heterogêneo e ordenado, que a língua varia e muda constantemente, que variação e mudança são sistemáticas e que a explicação da variação e da mudança linguística se faz possível apenas por meio do encaixamento destas nas estruturas interna e externa à língua.

De acordo com esse referencial teórico e seguindo a metodologia apresentada no terceiro capítulo deste texto, compôs-se um *corpus* a partir de 24 entrevistas sociolinguísticas e procedeu-se à seleção dos dados e posterior análise. Estipulou-se como variável dependente binária o apagamento (forma variante não-padrão) *versus* a manutenção (variante padrão) da oclusiva /d/ no morfema *-ndo*, indicador de gerúndio. Controlaram-se as variáveis sociais ‘escolaridade’, ‘sexo’ e ‘faixa etária’ e as variáveis linguísticas ‘conjugação verbal’, ‘extensão do verbo’, ‘contexto seguinte’ e ‘estrutura sintática’.

Após a codificação dos dados e feita a primeira rodada no *software* GoldVarb X, verificou-se um total de 904 ocorrências de formas verbais no gerúndio, das quais 612 foram da forma não-padrão e 292 foram da forma padrão. Assim, o percentual de apagamento corresponde a 67,7% dos dados da amostra. A partir desses números, pôde-se observar que a regra de apagamento está bastante avançada na fala da comunidade de Feira de Santana.

Os resultados da primeira rodada no GoldVarb X revelaram que os fatores sociais incidiram mais sobre a regra variável do que os fatores de ordem estrutural, sendo que foram selecionados, durante as rodadas realizadas pelo programa computacional, nesta ordem de significância, as variáveis ‘escolaridade’, ‘sexo’, ‘contexto seguinte’, ‘faixa etária’ e ‘extensão do verbo’. As variáveis linguísticas ‘conjugação verbal’ e ‘estrutura sintática’ foram descartadas na análise estatística computacional. A melhor rodada teve *input* final de 0,734 e nível de significância 0,007.

Os resultados para a variável mais significativa, a escolaridade, mostraram que, embora a forma inovadora esteja presente nas normas culta e popular da variedade de fala feirense, a regra de apagamento é muito mais produtiva na fala de indivíduos que cursaram apenas o nível fundamental I do que na fala daqueles que concluíram o nível universitário. Os pesos relativos atribuídos a cada fator da variável ‘escolaridade’ apontaram que o fator ‘nível

universitário’, com peso relativo de 0,25, inibe fortemente a aplicação da regra e que o fator ‘nível fundamental I’, com peso relativo de 0,77, favorece o apagamento. Tais resultados comprovam a hipótese inicial de que o apagamento fosse mais comum na fala dos menos escolarizados.

A variável ‘sexo’ foi a segunda selecionada pelo programa estatístico em ordem de relevância. Os resultados para esse grupo de fatores mostraram que informantes do sexo masculino (peso relativo de 0,61) favorecem a aplicação da regra, enquanto informantes do sexo feminino (peso relativo de 0,41) a inibem. Os resultados para a variável ‘sexo’ confirmam as hipóteses iniciais da pesquisa e se coadunam com uma assertiva clássica, embora não categórica, em Sociolinguística, de que homens e mulheres se comportam de maneiras diferentes no que concerne a fenômenos de variação e mudança linguística.

A variável ‘faixa etária’ foi a quarta selecionada pelo *software* de análise quantitativa em ordem de significância. Os resultados para esse grupo de fatores indicam que a regra de apagamento é mais produtiva na faixa mais jovem (de 25 a 35 anos) e que a aplicação diminui da primeira faixa para a segunda (de 45 a 55 anos) e desta para a terceira (a partir de 65 anos). Os valores percentuais de aplicação para as faixas 1, 2 e 3 foram, respectivamente, 75%, 67% e 62%, e os pesos relativos foram, respectivamente PR 0,60, PR 0,44 e PR 0,41. Esses valores percentuais indicam que a regra de apagamento está bastante avançada em todas as faixas etárias, contudo os pesos relativos indicam que a faixa 1 motiva a aplicação da regra de apagamento e que as faixas 2 e 3 a inibem, o que configura um quadro de mudança linguística em curso, considerado o tempo aparente.

Dentre as quatro variáveis linguísticas controladas na pesquisa, foram selecionadas como relevantes o ‘contexto seguinte’ e a ‘extensão do verbo’. Já as variáveis ‘conjugação verbal’ e ‘estrutura sintática’ foram descartadas durante a análise estatística computacional, certamente pela má distribuição dos dados da amostra.

O ‘contexto seguinte’ foi a primeira variável linguística selecionada pelo programa computacional de análise estatística, sendo a terceira selecionada no conjunto total das variáveis controladas. Ao controlar a variável ‘contexto seguinte’, pôde-se observar como os fatores ‘pausa’, ‘vogal’, ‘consoante’ e ‘haplologia’ condicionam a regra de apagamento. Os resultados para o grupo de fatores ‘contexto seguinte’ revelaram que os fatores ‘haplologia’ (PR 0,59) e ‘consoante’ (PR 0,58) favorecem a regra de apagamento e que os fatores ‘vogal’ (PR 0,44) e ‘pausa’ (PR 0,37) desfavoreceram a regra. Por não haver diferença estatística entre os dois fatores favorecedores, estes foram amalgamados e procedeu-se a uma segunda rodada dos dados no GoldVarb X, chegando ao resultado de que o fato de haver segmento

consonantal no contexto seguinte de fato favorece o apagamento do /d/ em morfema de gerúndio.

A ‘extensão do verbo’ foi a quinta variável em ordem de significância, sendo a segunda variável linguística a ser selecionada. Dentro desse grupo de fatores, observou-se se o fato de o verbo ser dissílabo, trissílabo ou polissílabo influenciaria a aplicação da regra de apagamento. A hipótese inicial era a de que quanto maior fosse o vocábulo maior também fosse a probabilidade de apagamento de /d/ no morfema *-ndo*, no entanto os resultados surpreenderam ao mostrarem que, ainda que levemente, os verbos trissílabos (PR 0,56) é que favorecem a aplicação da regra e que dissílabos (PR 0,40) e polissílabos (PR 0,43) desfavorecem-na.

Uma nova rodada dos dados foi feita no GoldVarb X a fim de verificar se haveria significância estatística entre dissílabos e verbos com três ou mais sílabas, como fez Mollica (1992), contudo o novo grupo de fatores não foi estatisticamente relevante, confirmando que, no *corpus* estudado, o fator ‘trissílabos’ é que favorece a regra, o que singulariza a variedade feirense em comparação com outras que foram apresentadas nesta dissertação, muito embora não pode ser desconsiderado o fato de os dados estarem mal distribuídos pelos fatores controlados.

As variáveis ‘conjugação verbal’ e ‘estrutura sintática’ como se disse, não se mostraram estatisticamente relevantes para a compreensão do condicionamento da regra de apagamento no *corpus*, pois as duas variáveis foram descartadas durante a análise estatística. Os valores de frequência percentual foram de 67%, 65% e 71% para os fatores 1^a, 2^a e 3^a conjugações, respectivamente. Esses valores percentuais são bastante próximos, corroborando o resultado da análise estatística.

Para a variável ‘estrutura sintática’, as frequências percentuais foram de 66%, 68% e 80% para os fatores ‘oração reduzida’, ‘perífrase’ e ‘justaposição’, respectivamente. Esses valores percentuais permitem inferir que a regra de apagamento de /d/ em *-ndo* é mais produtiva no contexto de justaposição, no entanto os dados de justaposição equivalem apenas a pouco mais de 3% do total de dados analisados, o que compromete a análise estatística e requer uma ampliação da amostra em estudo posterior.

Com vistas a aprofundar a compreensão dos resultados apresentados até aqui, foram realizadas mais algumas rodadas em que se cruzaram os dados das variáveis sociais. Assim, cruzaram-se as variáveis ‘faixa etária’ e ‘escolaridade’, ‘faixa etária’ e ‘sexo’ e, por fim, ‘escolaridade’ e ‘sexo’.

Os resultados do cruzamento entre as variáveis ‘faixa etária’ e ‘escolaridade’ revelaram que informantes mais jovens, no geral, apagam mais que os mais velhos, independentemente do nível de escolaridade. Contudo o nível mais elevado de escolaridade funciona como refreador em todas as faixas etárias, o que aponta para uma mudança de baixo para cima, admitindo-se a relação entre nível de escolaridade e nível socioeconômico na comunidade investigada.

O cruzamento das variáveis ‘faixa etária’ e ‘sexo’ evidenciou que o apagamento é favorecido pelos homens da primeira faixa etária e que os homens das faixas etárias intermediária e avançada, bem como as mulheres das três faixas etárias, inibem levemente o apagamento, revelando um processo de mudança que é liderado por indivíduos do sexo masculino e da faixa etária mais jovem.

Por fim, o cruzamento entre ‘escolaridade’ e ‘sexo’ revelou que homens e mulheres de ambos os sexos favorecem o apagamento quando relacionados ao menor nível de escolaridade e que, relacionando ambos os sexos ao alto nível de escolarização, ambos desfavorecem o apagamento. Em outras palavras, parece claro que a escolaridade atua mais fortemente na aplicação da regra de apagamento do /d/ nas formas de gerúndio do que o sexo, fato que explica a seleção da variável ‘escolaridade’ em primeiro lugar pelo programa estatístico utilizado na pesquisa.

Os resultados apresentados aqui são representativos da amostra de fala de Feira de Santana-BA constante do arquivo sonoro do projeto “A Língua Portuguesa no Semiárido Baiano” e, claro, embora possam ser generalizados para a comunidade de fala representada pela amostra, não são conclusivos e não encerram o tema pesquisado. Tais resultados representam muito bem o *corpus*, mas o aumento do tamanho deste em futuras pesquisas pode oferecer conclusões mais precisas sobre a variação na pronúncia do morfema de gerúndio na variedade de fala feirense.

A pesquisa aqui apresentada se faz inovadora e importante na medida em que não havia, até o presente momento, estudo sobre o tema na fala de Feira de Santana. Os dados analisados nesta dissertação permitem, portanto, a comparação entre a fala feirense e outras variedades de fala para a construção do mapeamento do apagamento do /d/ em morfema de gerúndio na língua portuguesa falada no território brasileiro.

Como toda pesquisa, nem tudo se esgota neste trabalho. Embora tenha sido abordada a Teoria da Fonologia Lexical, esta dissertação concentrou-se na descrição sociolinguística do fenômeno estudado na comunidade de Feira de Santana. A abordagem dessa teoria fonológica se impôs, sobretudo, para uma melhor compreensão do trabalho de Ferreira (2010). Todavia

houve o intuito também de instigar a continuação do trabalho em estudos posteriores que se concentrem numa interpretação linguística mais teórica do ponto de vista fonológico.

Outra direção possível para o aprofundamento do tema seria um estudo de crenças e atitudes para se averiguar o grau de aceitação da redução do gerúndio na comunidade feirense. Assim, seria possível verificar se o apagamento do /d/ no grupo *-ndo* é estigmatizado ou não.

Daí a necessidade de continuação da pesquisa para refinamento da análise aqui apresentada, cujos limites se impõem por se tratar de uma dissertação de Mestrado, curso que tem uma duração de apenas dois anos e, no caso específico deste trabalho, desenvolvido durante um ano de isolamento social devido à pandemia do Coronavírus que assola o mundo neste momento.

Por fim, espera-se que esta pesquisa possa contribuir para a descrição da variedade falada em Feira de Santana e do português brasileiro em geral, pois que quanto mais descrições com rigor metodológico, mais se divulga e conhece a realidade da língua falada num país de dimensões continentais.

REFERÊNCIAS

- ALMEIDA, A. N. S.; OLIVEIRA, A. J. Você fala cantano? Uma análise sistemática no morfema *-ndo* formador de gerúndio no português. *Letrônica: Revista Digital do Programa de Pós-Graduação em Letras da PUCRS*. Porto Alegre, v. 10, n. 1, p. 200-209, jan.-jun./2017.
- ALMEIDA, B. K. M.; PEREIRA, M. L. S.; ARAÚJO, A. A. O Apagamento de /d/ no morfema de gerúndio no falar culto de Fortaleza-CE. *Revista Educação e Linguagens*. Campo Mourão, v. 17, n. 9, p. 511-538, jul.-dez./2020.
- AMARAL, A. *O dialeto caipira: gramática, vocabulário*. 4. ed. São Paulo: INL, 1982.
- ANDRADE, C. M. P. *Origens do povoamento de Feira de Santana: um estudo de história colonial*. 1990. 166f. Dissertação (Mestrado em Ciências Sociais) – Universidade Federal da Bahia, Salvador, 1990.
- ARAÚJO, R. C. *Apagamento da oclusiva dental /d/ no morfema {-ndo} formador de gerúndio na fala envirence*. 2019. 143f. Dissertação (Mestrado em Letras) – Universidade Federal do Amazonas, Manaus, 2019.
- ARAÚJO, S. S. F. *A concordância verbal no português falado em Feira de Santana-BA: sociolinguística e sócio-história do português brasileiro*. 2014. 342f. Tese (Doutorado em Língua e Cultura) – Universidade Federal da Bahia, Salvador, 2014.
- ARAÚJO, S. S. F.; ALMEIDA, N. L. F. O projeto A língua portuguesa no Semiárido Baiano – fase 3: critérios de constituição e da amostragem de dados. In: FREITAG, R. (Org.). *Metodologia de coleta e de manipulação de dados em sociolinguística*. São Paulo: Blucher, 2014. p. 27-48.
- BAHIA. *Lei ordinária n 4167 de 7 de novembro de 1983*. Autoriza o Poder Executivo a criar, sob a forma de autarquia, o Centro Industrial do Subaé – CIS e dá outras providências. Salvador, 1983.
- BERLINK, R. A.; BIAZOLLI, C. C. Ferramentas metodológicas para análises (socio)linguísticas. *Estudos Linguísticos*. São Paulo, v. 1, n. 47, p. 260-273, 2018.
- BISOL, L. Fonologia lexical. In: HORA, D.; MATZENAUER, C. L. *Fonologia, fonologias: uma introdução*. São Paulo: Contexto, 2017. p. 81-92.
- BORTONI-RICARDO, S. M. *Educação em língua maternal: a sociolinguística na sala de aula*. São Paulo: Parábola Editorial, 2004.
- BRAGA, M. L.; CORIOLANO, J. Construções do gerúndio no português do Brasil. *Alfa*. São Paulo, v. 51, n. 1, p. 175-187, 2007.
- BYBEE, J. *Mudança linguística*. Trad. Marcos Bagno. Petrópolis: Vozes, 2020.

- BYBEE, J. Mechanisms of change in grammaticization: the role of frequency. In: JOSEPH, B. D.; JANDA, R. D. (Eds.). *The handbook of historical linguistics*. Oxford: Blackwell, 2003. p. 602-623.
- CAMACHO, R. G. *Da linguística formal à linguística social*. São Paulo: Parábola Editorial, 2013.
- CAMPOS, O. G. L. A. S. *O gerúndio no português (estudo histórico-descritivo)*. 1972. Tese (Doutorado em Letras) – Universidade de São Paulo, Araraquara, 1972.
- CARVALHO, M. C. M. *Pensando no gerúndio, acabei fazendo uma tese: análise dos usos do gerúndio não perifrástico em gêneros argumentativos escritos do Português Brasileiro*. 2018. 389 f. Dissertação (Doutorado em Línguas Clássicas) – Universidade de Brasília, Brasília, 2018.
- CASTILHO, A. T. *Nova gramática do português brasileiro*. São Paulo: Contexto, 2019.
- CHAVES, I. O. Panorama dos estudos das vogais pretônicas no português do Brasil. *Revista Científica Vozes do Vale*. Diamantina, v. 6, n. 1, p. 1-25, 2014.
- COELHO, I. *Sociolinguística*. Santa Catarina: UFSC, 2010.
- COELHO, P. M. C. R. *O tratamento da variação linguística nos livros didáticos de português*. 2007. 162f. Dissertação (Mestrado em Linguística) – Universidade de Brasília, Brasília, 2007.
- CRISTÓFARO SILVA, T. Fonologia: por uma análise integrada à morfologia e à sintaxe. *Cadernos do Departamento de Letras Vernáculas* (Anais da 2ª Semana de Estudos Portugueses). Belo Horizonte, v. 2, p. 60-70, 1996.
- FEIRA DE SANTANA. *Maior entroncamento rodoviário do norte-nordeste*. 2019. Disponível em: <https://www.mobilidade.feira.br/publicação/feira-de-santana-maior-entroncamento-rodoviario-do-norte-e-nordeste>. Acesso em: 4 jun., 2020.
- FERREIRA, J. S. *O apagamento do /d/ em morfema de gerúndio no dialeto de São José do Rio Preto*. 2010. 148f. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Estadual Paulista, São José do Rio Preto, 2010.
- FERREIRA, J. S.; TENANI, L. A redução do gerúndio à luz da Fonologia Lexical. *Estudos Linguísticos*. São Paulo, v. 38, p. 59-68, 2009.
- FREITAG, R. M. K. Idade, uma variável sociolinguística complexa. *Revista Línguas & Letras*. Cascavel, v. 6, n. 11, p. 105-121, 2005.
- FREITAG, R. M. K.; CARDOSO, P. B.; PINHEIRO, B. F. M. Saliência na conservação de /d/ no segmento /ndo/: efeitos sociais e estilísticos. *Gragotá*. Niterói, v. 23, n. 46, p. 654-677, 2018.

- GALVÃO, R. A. Os povoadores da região de Feira de Santana. *Sitientibus*. Feira de Santana, v.1, n. 1, p. 25-31, 1982.
- GONÇALVES, D. T. *Pronúncia variável de (ndo) na fala paulistana*. 2018. 131f. Dissertação (Mestrado em Semiótica e Linguística Geral) – Universidade de São Paulo, São Paulo, 2018.
- GUY, G. R.; ZILLES, A. *Sociolinguística quantitativa: instrumental de análise*. São Paulo: Parábola Editorial, 2007.
- IBGE. Instituto Brasileiro de Geografia e Estatística. *Cidades e estados: Feira de Santana*. 2019. Disponível em: <https://www.ibge.gov.br/cidades-e-estados/ba/feira-de-santana.html>. Acesso em: 27 jun. 2019.
- LABOV, W. *Padrões sociolinguísticos*. Trad. de Marcos Bagno, Maria Marta Scherre e Caroline R. Cardoso. São Paulo: Parábola Editorial, 2008 [1972].
- LABOV, W. Building on empirical foundations. In: LEHMANN, W.; MALKIEL, Y. (Orgs.). *Perspectives on historical linguistics*. Amsterdam/Philadelphia: John Benjamins Publishing, 1982. p. 79-92.
- LABOV, W. *Principles of linguistic change: internal factors*. Oxford: Blackwell, 1994.
- LABOV, W. *The social stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics, 1966.
- LEE, S. H. *Morfologia e fonologia lexical do português do Brasil*. 1995. 201f. Tese (Doutorado em Linguística) – Universidade Estadual de Campinas, Campinas, 1995.
- LEE, S. H. Fonologia lexical: modelos e princípios. *Letras de hoje*. Porto Alegre, v. 31, n. 2, p. 129-137, 1996.
- LENNEBERG, E. H. *Biological foundations of language*. New York: Wiley, 1967.
- LUCCHESI, D. *Sistema, mudança e linguagem: um percurso pela história da linguística moderna*. São Paulo: Parábola Editorial, 2004.
- LUCENA, R. M.; VASCONCELOS, D. C. Apagamento da oclusiva dental no dialeto do Brejo Paraibano: uma regra variável. *A Cor das Letras*. Feira de Santana, v. 8, n. 1, p. 231-239, 2007.
- MACEDO, L. S. A. et al. Estudo descritivo sobre o apagamento do /d/ no morfema de gerúndio nos verbos do português falado em Cachoeirinha-PE. *Web-Revista Socioleto*. Campo Grande, v. 9, n. 26, p. 235-264, 2018. Disponível em: <http://sociodialeto.com.br/index.php/sociodialeto/article/view/151>. Acesso em: 5 nov. 2020.
- MARINHO, J. H. C.; VAL, M. G. C. *Variação linguística e ensino: caderno do professor*. Belo Horizonte: Ceale, 2006.

- MARROQUIM, M. *A língua do Nordeste: Alagoas e Pernambuco*. São Paulo: Nacional, 1934.
- MENDES, R. B. Língua e variação. In: FIORIN, J. L. et al. (Orgs.). *Linguística? Que é isso?* São Paulo: Contexto, 2013. p. 111-135.
- MIRANDA, E. O.; SILVA, H. M. S.; OLIVEIRA, L. P. Implicações da urbanização em Feira de Santana: o atual desenho da Rua Barão do Rio Branco a partir de uma análise social, econômica e territorial. *Seminário de Urbanização e Cidades da Bahia*. Salvador, v. 12, p. 1-10, 2013.
- MOLLICA, M. C.; MATTOS, P. B. Pela Conjugação das abordagens variacionista e difusionista. *Revista de Estudos da Linguagem*. Belo Horizonte, v. 1, p. 53-64, 1992.
- MOLLICA, M. C. Fundamentação teórica: conceituação e delimitação. In: MOLLICA, M. C.; BRAGA, M. L. (Orgs.). *Introdução à sociolinguística: o tratamento da variação*. São Paulo: Contexto, 2004. p. 9-14.
- MOTHÉ, N. G. M. Gerúndio *versus* infinitivo gerundivo: Brasil e Portugal em contraste nos séculos XIX e XX. *Estudos Linguísticos*. Belo Horizonte, n. 35, p. 1554-1563, 2006.
- NARO, A. J. O dinamismo das línguas. In: MOLLICA, M. C.; BRAGA, M. L. (Orgs.). *Introdução à sociolinguística: o tratamento da variação*. São Paulo: Contexto, 2013. p. 43-50.
- NASCIMENTO, K. R. S.; ARAÚJO, A. A.; CARVALHO, W. J. A. A redução do gerúndio no falar popular de Fortaleza: um olhar variacionista. *Veredas*. Juiz de Fora, v. 17, n. 2, p. 398-413, 2013.
- NASCIMENTO, L.; MOTA, J. A ausência de /d/ no gerúndio com base nos dados experimentais do Projeto ALiB. *Hyperion Letras*. Salvador, n. 7, s./p., 2004.
- OLIVEIRA, A. J. Análise quantitativa no estudo da variação linguística: noções de estatística e análise comparativa entre Varbrul e SPSS. *Revista de Estudos da Linguagem*. Belo Horizonte, v. 17, n. 2, p. 93-119, 2009.
- OLIVEIRA, J. M. *O futuro da língua portuguesa ontem e hoje: variação e mudança*. 2006. 254f. Tese (Doutorado em Letras Vernáculas) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2006.
- PAIVA, M. C. A variável sexo/gênero. In: MOLLICA, M. C.; BRAGA, M. L. (Orgs.). *Introdução à sociolinguística: o tratamento da variação*. 4. ed. São Paulo: Contexto, 2013. p. 33-42.
- SANKOFF, D.; TAGLIAMONTE, S.; SMITH, E. *GoldVarb X: a multivariate analysis application*. Toronto: Department of Linguistics; Ottawa: Department of Mathematics, 2005. Disponível em: <https://tinyurl.com/yxcuyops>. Acesso em: 20 jun. 2020.

- SANTOS JR., A. C. B. *Região metropolitana de Feira de Santana: uma proposta de gestão*. 2019. 151f. Dissertação (Mestrado em Planejamento Territorial) – Universidade Estadual de Feira de Santana, Feira de Santana, 2019.
- SAUSSURE, F. *Curso de linguística geral*. Trad. Antônio Chelini, José Paulo Paes e Izidoro Blikstein. São Paulo: Cultrix, 1995 [1916].
- SOUZA, M. M. Formas verbo-nominais latinas: ressonâncias em português. *Cadernos do CNLF*. Série VII, n.11 – Morfossintaxe (VII Congresso Nacional de Linguística e Filologia). Rio de Janeiro, 2003. Disponível em: <http://www.filologia.org.br/viicnlf/anais/caderno11-08.html>. Acesso em: 17 fev., 2021.
- VIEIRA, M. S. Apagamento do /d/: abordagem sociolinguística sobre a perspectiva do gênero sexual. *Web-Revista Socioleto*. Campo Grande, v. 1, n. 4, p. 1-27, 2011. Disponível em: <http://www.sociodialeto.com.br/edicoes/9/28092011063729.pdf>. Acesso em: 30 jan., 2018.
- VOTRE, S. J. Relevância da variável escolaridade. In: MOLLICA, M. C.; BRAGA, M. L. (Orgs.). *Introdução à sociolinguística: o tratamento da variação*. 4. ed. São Paulo: Contexto, 2013. p. 51-56.
- WEINREICH, U.; LABOV, W.; HERZOG, M. *Fundamentos empíricos para uma teoria da mudança linguística*. Trad. de Marcos Bagno. São Paulo: Parábola Editorial, 2006 [1968].