



UNIVERSIDADE ESTADUAL DO CEARÁ
CENTRO DE HUMANIDADES
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA APLICADA
MESTRADO EM LINGUÍSTICA APLICADA

FRANCISCO FERREIRA DE SOUZA

TEM CHANCE DE HAVER AINDA EXISTIR NO FALAR POPULAR?:
A VARIAÇÃO DOS VERBOS EXISTENCIAIS EM AMOSTRA DO NORPOFOR

FORTALEZA – CEARÁ

2015

FRANCISCO FERREIRA DE SOUZA

*TEM CHANCE DE HAVER AINDA EXISTIR NO FALAR POPULAR?:
A VARIAÇÃO DOS VERBOS EXISTENCIAIS EM AMOSTRA DO NORPOFOR*

Dissertação apresentada ao Programa de Pós-Graduação em Linguística Aplicada (PosLA) da Universidade Estadual do Ceará, como requisito parcial à obtenção do título de Mestre em Linguística Aplicada.

Área de Concentração: Linguagem e Interação

Orientadora: Prof.^a Dr.^a Aluiza Alves de Araújo.

FORTALEZA – CEARÁ

2015

Dados Internacionais de Catalogação na Publicação

Universidade Estadual do Ceará

Sistema de Bibliotecas

Souza, Francisco Ferreira de.

Tem chance de haver ainda existir no falar popular?: a variação dos verbos existenciais em amostra do NORPOFOR [recurso eletrônico] / Francisco Ferreira de Souza. - 2015.

1 CD-ROM: il.; 4 ¾ pol.

CD-ROM contendo o arquivo no formato PDF do trabalho acadêmico com 105 folhas, acondicionado em caixa de DVD Slim (19 x 14 cm x 7 mm).

Dissertação (mestrado acadêmico) - Universidade Estadual do Ceará, Centro de Humanidades, Programa de Pós-Graduação em Linguística Aplicada, Fortaleza, 2015.

Área de concentração: Linguagem e Interação.

Orientação: Prof.^a Dra. Aluiza Alves de Araújo.

1. Verbos haver/existir/ter. 2. Sociolinguística Variacionista. 3. Norma popular. 4. Fortaleza. I. Título.

FRANCISCO FERREIRA DE SOUZA

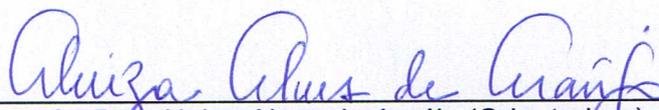
TEM CHANCE DE HAVER AINDA EXISTIR NO FALAR POPULAR? A
VARIAÇÃO DOS VERBOS EXISTENCIAIS EM AMOSTRA DO NORPOFOR

Dissertação apresentada ao Programa de Pós-Graduação em Linguística Aplicada do Centro de Humanidades da Universidade Estadual do Ceará, como requisito parcial para a obtenção do grau de Mestre em Linguística Aplicada.

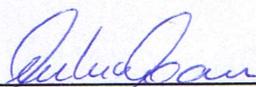
Área de Concentração: Linguagem e Interação

Aprovada em: 07 / 05 / 2015.

BANCA EXAMINADORA



Profa. Dra. Aluiza Alves de Araújo (Orientadora)
Universidade Estadual do Ceará – UECE



Profa. Dra. Márluce Coan
Universidade Federal do Ceará – UFC



Prof. Dr. Wilson Júnior de Araújo Carvalho
Universidade Estadual do Ceará – UECE

A Augustus Watters, que chegou em nosso lar para me ajudar a finalizar este trabalho e para trazer para mim e para minha esposa a maior inspiração que pode *ter/existir/haver* em um lar: nosso filho Davi, a quem também dedico este trabalho. Augustus viveu como uma brisa leve, que toca nossos rostos e vai embora, repentinamente, deixando um perfume das mais belas flores do campo, cujo odor recordaremos para sempre.

AGRADECIMENTOS

A **Deus**, por ter me dado a paz necessária e a saúde suficiente para concluir mais esta etapa da minha vida.

A minha mãe, **Guiomar**, e a meu pai, **Valdemar**, que apesar do pouco estudo que possuem, conseguiram orientar bem os filhos, ensinando-nos princípios que muitas vezes não aprendemos no mundo acadêmico, como amor ao próximo, humildade e, principalmente, honestidade.

A minha adorável esposa, **Wendla Souza**, por ser meu porto seguro não só durante esses dois anos de Mestrado, mas durante os quatro anos que estamos juntos; companheira em todos os momentos, ela me faz ir sempre mais longe, com seu jeito insistente de ser. E também por ter sido paciente nos momentos em que ficou em segundo plano devido às atividades me impostas pelo Mestrado. Obrigado, amor!

A **meu filho** Davi que está por vir e que, ainda do tamanho de uma semente de gergelim, foi o combustível que faltava para a conclusão deste trabalho.

A meu irmão **Vilmar**, a quem Deus selecionou para ser um desses desbravadores que existem no mundo. Por ter esse dom, tirou a família da miséria e atuou como meu segundo pai, sendo o grande responsável pela minha formação educacional.

As minhas amadas irmãs, **Quinha** e **Vilma**. A primeira por ser uma “parceria forte” e pelas longas conversas ao telefone. E a segunda pela meiguice e pelos “chás” que me fizeram relaxar e dormir bem quando se fez necessário.

As minhas adoradas sobrinhas, **Thayná** e **Thayane**, que foram bênçãos divinas, responsáveis por meu pai ter largado o álcool, trazendo paz e muitas alegrias para nossa família. A primeira, por ser a mais velha, ainda foi “explorada” em alguns trabalhos do Mestrado.

A minha querida orientadora, profa. Dra. **Aluiza Alves de Araújo**, por ter sido mais que uma orientadora, cuja paciência e verdadeiras sessões de terapia possibilitaram a realização desta pesquisa. Obrigado, professora, por saber aliar competência profissional com toques de compreensão, que me foram muito caros e sem os quais não teria conseguido concluir esse Mestrado, cheio de obstáculos.

Aos profs. Dr. **Wilson Júnior de Araújo Carvalho** e Dr. **Kilpatrik Müller Campelo**, pelas valiosas contribuições em meu projeto de qualificação e, especialmente ao **prof. Wilson**, por também ter feito parte da minha defesa de dissertação, brindando-me com outras pertinentes observações.

A profa. Dra. **Márluce Coan**, pelas questões levantadas em minha defesa e pelas sugestões relevantes para a conclusão deste trabalho.

Ao **PosLA**, por ter me concedido mais dois meses de prorrogação, sem os quais teria sido impossível a conclusão do prazo para a entrega do texto.

A **Keiliane Dantas**, que sempre estava pronta para me ajudar, mesmo quando a tarefa não era de sua responsabilidade, agindo de forma humana sem deixar de ser profissional.

A **Germana**, pela colaboração do texto de William Bright, “As dimensões da Sociolinguística”.

A minha mãe crateuense, **Cicinha**, por ter me escutado em todos os momentos de angústia durante esse mestrado.

A meu amigo desde a faculdade, **João Paulo**, por nunca ter deixado de torcer por mim em todas as minhas conquistas.

A minha amiga **Lucy Lanna**, por ter sido parceira de todos os momentos e irmãzona de coração.

A meu amigo **Nonato Furtado**, pelas conversas, apoio e pela disponibilidade de sua moradia para que eu pudesse fazer o exame de seleção do Mestrado.

Ao meus amigos **Israel, Warley e Simone, Sabino e Cássia, Márcio e Rita de Cássia** por me aportarem em suas casas quando eu precisava dormir em Fortaleza.

A **todos os meus alunos do IFCE – campus Crateús** (cursos técnicos e superiores), por terem sido muito compreensíveis em momentos que precisei de apoio.

Aos meus bolsistas **Julianne, Tálisson e Iara**, que ajudaram a constituir o *corpus* de língua escrita, que serviu de base para o primeiro embrião deste trabalho: apresentação de duas comunicações no II SIC, em 2011, que foi realizado em Redenção-CE.

A **dona Graça** (a tia da Cantina) e sua filha **Narjane**, por terem sido tão gentis comigo durante os dois anos de Mestrado.

Aos professores de todas as 7 disciplinas que cursei: Dra. **Rozania Maria Alves de Moraes**, Dr. **Expedito Eloísio Ximenes**, Dra. **Claudiana Nogueira de Alencar**, Dr. **Antônio Luciano Pontes**, Dra. **Cibele Gadelha Bernardino**, Dra. **Paula Lenz Costa Lima** e minha adorada orientadora, Dra. **Aluiza Alves de Araújo**.

Ao meu amigo **Gilberto Amorim**, simplesmente por encher meu saco e se preocupar com esse trabalho, sempre querendo saber como estava seu andamento.

A **Elyne Vitória**, por ter me cedido valiosos textos, entre os quais seu relatório de pós-doutoramento.

A **Marcos Bagno**, por ter sido sempre solícito em todos os e-mails que compartilhamos. A todos os **colegas de trabalho no IFCE** que, direta ou indiretamente, contribuíram para a realização deste trabalho.

Ao **IFCE *campus* Crateús**, por ter me dado condições estudar durante esses dois anos difíceis.

“No meio do caminho tinha uma pedra
tinha uma pedra no meio do caminho
tinha uma pedra
no meio do caminho tinha uma pedra.

Nunca me esquecerei desse
acontecimento
na vida de minhas retinas tão fatigadas.
Nunca me esquecerei que no meio do
caminho
tinha uma pedra
tinha uma pedra no meio do caminho
no meio do caminho tinha uma pedra.”

(Carlos Drummond de Andrade)

RESUMO

Este estudo tem como objetivo geral investigar o lugar dos verbos *haver* e *existir* nas sentenças existenciais do português popular de Fortaleza, uma vez que esse tipo de estrutura é predominantemente ocupado pelo verbo *ter* em pesquisas sobre o falar culto. Partindo da hipótese de que no falar popular essa predominância será ainda maior, traçamos dois objetivos específicos: o primeiro é identificar os fatores sociais e/ou linguísticos responsáveis pela manutenção de *haver* e *existir* no falar popular de Fortaleza; o segundo é investigar, a partir de análises no tempo aparente (faixa etária), se há indícios de esses verbos estarem passando por um processo de mudança em curso no sentido de serem substituídos pelo verbo *ter*. Para tal empreendimento, utilizamos os pressupostos teórico-metodológicos da Teoria da Variação e Mudança Linguística (WEINREINCH; LABOV; HERZOG, 1968; LABOV, 1994; 2001; 2008 [1972]; 2010) e as análises estatísticas, realizadas pelo programa computacional GoldVarb X (2005), de 931 dados de fala de 53 informantes, das 29 entrevistas do tipo D2 (Diálogo entre Dois Informantes), selecionadas a partir do banco de dados NORPOFOR (Norma Oral do Português Popular de Fortaleza). As variáveis independentes controladas nesta pesquisa foram: sexo, faixa etária, escolaridade (sociais) e forma verbal, animacidade do SN, peso do SN, posição do SN em relação ao verbo, concordância entre o verbo e o SN, presença x ausência de elementos à esquerda do verbo e repetição do verbo no mesmo enunciado (linguísticas). Inicialmente, realizamos a rodada ternária para sabermos as frequências de uso das três variantes analisadas e constatamos nossa hipótese inicial de que *ter* apresentaria um índice muito alto (94,1%) em relação a *haver* (2,9%) e *existir* (3,0%). Depois, controlamos os 10 grupos de fatores em três análises binárias (*haver/ter*, *existir/ter* e *haver/existir*). Os resultados dessas análises mostraram que o verbo *haver* ainda encontra espaço nas estruturas existenciais tanto em contraste com *ter*, condicionado pelos fatores *infinitivo*, *pretérito perfeito do indicativo* e *falantes a partir de 50 anos*, quanto em contraste com *existir*, condicionado pelas mesmas formas verbais da análise *haver/ter* (*infinitivo*, *pretérito perfeito do indicativo*) e pelos falantes do *sexo feminino*. O verbo *existir* também encontra espaço nas estruturas existenciais, tanto na análise com *ter*, favorecido pelo fator *SN à esquerda*, quanto na análise com *haver*, favorecido pelos fatores *presente do indicativo*, *pretérito imperfeito do indicativo* e *sexo masculino*. A análise em tempo aparente, realizada somente na variação *haver/ter*, indicou que existem indícios de mudança em progresso na direção

da total implementação de *ter* no falar popular em detrimento de *haver*. Concluímos que nossos resultados fornecem informações importantes acerca dos verbos existenciais no português popular de Fortaleza, apesar de nosso tema não representar uma marca específica dessa comunidade de fala. A variação *haver/existir/ter* parece ser um fenômeno mais geral do que específico, indicando, talvez, uma tendência da expansão de *ter* no português tanto culto quanto popular falado em outras capitais brasileiras.

Palavras-chave: Verbos *haver/existir/ter*. Sociolinguística Variacionista. Norma popular. Fortaleza.

ABSTRACT

This study aims at investigating the place of the verbs *there to be* and *to exist* in existential sentences of nonstandard Portuguese spoken in Fortaleza, since this type of structure is predominantly occupied by the verb *to have* on researches about the standard language. Assuming that this predominance will be even greater in nonstandard Portuguese of Fortaleza, we draw two specific objectives: the first is to identify the social and/or linguistic factors responsible for maintaining *there to be* and *to exist* in this type of language in the speech community; the second is to investigate, from analysis in apparent time (age), if there is evidence of these verbs are going through an ongoing process of change in order to be replaced by the verb *to have*. For these purposes, we used the theoretical and methodological assumptions of the Theory of Language Variation and Change (WEINREINCH; Labov; HERZOG, 1968; Labov, 1994; 2001; 2008 [1972]; 2010) and statistical analyzes, made by the software GoldVarb X (2005), of 931 speech data of 53 respondents, in 29 interviews of the D2 type (Dialogue between Two informants), selected from the NORPOFOR database (Oral norm of Nonstandard Portuguese of Fortaleza). The independent variables controlled in this research were: sex, age, level of education (social variables) and verb form, the animacy of SN, SN weight, SN position relative to the verb, agreement between the verb and the SN, presence x absence of elements to the left of the verb and repetition of the verb in the same statement (linguistic variables). Initially, we conducted the ternary round to know the use of frequencies of the three variants analyzed and found that our initial hypothesis that the verb *to have* would present a very high rate (94.1%) over the verbs *there to be* (2.9%) and *to exist* (3, 0%). Then, we controlled the 10 groups of factors in three binary analysis (*there to be/to have*, *to exist /to have* and *there to be/to exist*). The results of these analyzes showed that the verb *there to be* still finds room in the existential structures in contrast to the verb *to have*, conditioned primarily by factors *infinitive*, *imperfect past form* and *speakers from 50 years*. The verb *to exist* also finds room in the existential structures contrasted with *to have*, and *the sentences with SN on the left* is the most favorable factor. In the analyses of this verb with *there to be*, the factor which favors it the most are: *infinitive*, *imperfect past form* and *male speakers*. The analysis in apparent time, performed only on the variation *there to be/to have* indicated that there are signs of change in progress towards the full implementation of the verb *to have* in the speech community of Fortaleza. We

conclude that our results provide important information about the existential verbs in the popular Fortaleza Portuguese despite the status of our theme: it does not represent a specific feature of that speech community. The variation phenomenon *there to be/to exist/to have* seems to be more general than specific, indicating perhaps a trend of expansion of *to have* in both standard and nonstandard Portuguese spoken in other capitals of the country.

Keywords: Variation *there to be/to exist/to have*. Variationist Sociolinguistics. Nonstandard norm. Fortaleza.

LISTA DE TABELAS

Tabela 1 –	Distribuição das ocorrências no grupo peso do SN nos três verbos	67
Tabela 2 –	Distribuição das ocorrências no grupo forma verbal nos três verbos	68
Tabela 3 –	Atuação da variável forma verbal sobre o verbo <i>haver</i>	75
Tabela 4 –	Atuação da variável faixa etária sobre o verbo <i>haver</i>	76
Tabela 5 –	Atuação do cruzamento sexo x faixa etária sobre o verbo <i>haver</i>	77
Tabela 6 –	Atuação da variável posição do SN sobre o verbo <i>existir</i>	79
Tabela 7 –	Atuação da variável forma verbal sobre o verbo <i>haver</i>	86
Tabela 8 –	Atuação da variável sexo sobre o verbo <i>haver</i>	86

LISTA DE QUADROS

Quadro 1 –	Distribuição dos bairros de Fortaleza por Secretarias Executivas Regionais (SERs)	43
Quadro 2 –	Distribuição dos informantes do NORPOFOR por tipo de registro, sexo, faixa etária e escolaridade.....	46
Quadro 3 –	Distribuição dos informantes por sexo, escolaridade e faixa etária em nossa amostra.....	47
Quadro 4 –	Distribuição das ocorrências por sexo, escolaridade e faixa etária em nossa amostra.....	47
Quadro 5 –	Distribuição dos informantes da amostra por nº do inquérito/informante, sexo, idade, escolaridade, bairro, Secretaria Executiva Regional (SER) e atividade exercida	49
Quadro 6 –	Síntese dos resultados das três análises	87

LISTA DE GRÁFICOS

Gráfico 1 – Resultados gerais da frequência de uso para <i>ter</i> e <i>haver</i> em estudos variacionistas	33
Gráfico 2 – Frequência de uso das variantes <i>haver</i> , <i>existir</i> e <i>ter</i>	70
Gráfico 3 – Comparação das frequências de uso de <i>haver</i> , <i>existir</i> e <i>ter</i> em quatro estudos	71
Gráfico 4 – Frequência de uso das variantes <i>haver</i> e <i>ter</i>	72
Gráfico 5 – Comparação dos resultados gerais de <i>haver/ter</i> de Silva (2001), nosso estudo (2015) e Vitória (2012)	73
Gráfico 6 – Frequências de uso de <i>ter</i> nos dados do projeto NURC.....	74
Gráfico 7 – Frequência de uso das variantes <i>existir</i> e <i>ter</i>	78
Gráfico 8 – Atuação do cruzamento posição do SN x sexo sobre o verbo <i>existir</i>	80
Gráfico 9 – Atuação do cruzamento posição do SN x escolaridade sobre o verbo <i>existir</i>	81
Gráfico 10 – Atuação do cruzamento sexo x escolaridade sobre o verbo <i>existir</i>	82
Gráfico 11 – Frequência de uso das análises <i>ter/haver</i> e <i>ter/existir</i> em três estudos e no nosso	84
Gráfico 12 – Frequência de uso das variantes <i>haver</i> e <i>existir</i>	85
Gráfico 13 – Atuação da variável faixa etária sobre <i>ter</i>	90
Gráfico 14 – Atuação do cruzamento faixa etária x sexo sobre a variante <i>ter</i>	91
Gráfico 15 – Probabilidade de ocorrência das variantes nas três análises realizadas....	93

SUMÁRIO

1	INTRODUÇÃO	18
2	A VARIAÇÃO DOS VERBOS EXISTENCIAIS <i>Haver, existir e ter</i> NO PORTUGUÊS BRASILEIRO	23
2.1	BREVE INCURSÃO HISTÓRICA	23
2.2	ESTUDOS VARIACIONISTAS RELEVANTES NO PORTUGUÊS BRASILEIRO.....	25
3	A TEORIA DA VARIAÇÃO E DA MUDANÇA LINGUÍSTICA	34
4	METODOLOGIA	42
4.1	TIPO DE PESQUISA	42
4.2	A COMUNIDADE DE FALA: A CIDADE DE FORTALEZA	42
4.3	O <i>CORPUS</i> E A AMOSTRA	44
4.4	AS ENTREVISTAS D2 DO NORPOFOR.	48
4.5	OS INFORMANTES DE NOSSA AMOSTRA.....	49
4.6	LEVANTAMENTO DOS DADOS.	51
4.7	VARIÁVEIS.....	51
4.7.1	Variável dependente	51
4.7.2	Variáveis independentes	53
4.7.2.1	Variáveis sociais	53
4.7.2.1.1	<i>Sexo</i>	53
4.7.2.1.2	<i>Faixa etária</i>	54
4.7.2.1.3	<i>Escolaridade</i>	55
4.7.2.2	Variáveis linguísticas.....	55
4.7.2.2.1	<i>Forma verbal</i>	56
4.7.2.2.2	<i>Animacidade do SN</i>	57
4.7.2.2.3	<i>Peso do SN</i>	58
4.7.2.2.4	<i>Posição do SN em relação ao verbo</i>	58
4.7.2.2.5	<i>Concordância entre o verbo e o SN</i>	59
4.7.2.2.6	<i>Presença x ausência de elementos à esquerda do verbo</i>	59
4.7.2.2.7	<i>Repetição do verbo no mesmo enunciado</i>	60
4.8	CODIFICAÇÃO DOS DADOS	60
4.9	GOLDVARB X	61
5	DESCRIÇÃO DOS DADOS E ANÁLISE DOS RESULTADOS	66

5.1	RODADA EXPERIMENTAL	65
5.2	RODADA TERNÁRIA: VISÃO GERAL DAS VARIANTES	69
5.3	ANÁLISES DE <i>HAVER/TER</i>	71
5.3.1	Primeira análise: visão geral da regra	72
5.3.2	Segunda análise: visão mais detalhada da regra	74
5.3.2.1	Forma verbal.....	75
5.3.2.2	Faixa etária	75
5.3.2.3	Faixa etária x sexo	76
5.3.3	Considerações sobre as análises de <i>haver/ter</i>	77
5.4	ANÁLISES DE <i>EXISTIR/TER</i>	78
5.4.1	Primeira análise: visão geral da regra	78
5.4.2	Segunda análise: visão mais detalhada da regra	78
5.4.2.1	Posição do SN.....	79
5.4.2.2	Posição do SN x sexo	80
5.4.2.3	Posição do SN x escolaridade.....	81
5.4.2.4	Sexo x escolaridade	82
5.4.3	Considerações sobre as análises de <i>existir/ter</i>	83
5.5	ANÁLISES DE <i>HAVER/EXISTIR</i>	83
5.5.1	Primeira análise: visão geral da regra	83
5.5.2	Segunda análise: visão mais detalhada da regra	85
5.5.2.1	Forma verbal.....	86
5.5.2.2	Sexo	86
5.5.3	Considerações sobre as análises de <i>haver/existir</i>	87
5.6	ANÁLISE DO TEMPO APARENTE.....	88
5.6.1	O tempo aparente na variação <i>haver/ter</i>	89
6	CONCLUSÃO	92
	REFERÊNCIAS	96
	APÊNDICES	101
	APÊNDICE A – Mapa dos estudos variacionistas sobre <i>haver/ter</i> E sobre <i>haver/ter/existir</i>	102
	APÊNDICE B – Distribuição dos falantes por bairros (os triângulos representam os homens e os círculos representam as mulheres).....	103
	APÊNDICE C – Chave de codificação	104

1 INTRODUÇÃO

Atualmente, o ensino da Língua Portuguesa no Brasil passa por um processo de profundas mudanças. Dentre elas, está o modo como o ensino de gramática é abordado em sala de aula. Durante muitos séculos, a gramática tradicional, também conhecida por gramática normativa ou prescritiva, era o único guia que os professores e os alunos tinham a sua disposição para orientá-los quanto ao uso da língua. No entanto, com o surgimento da Linguística, a tradição de estudos gramaticais começou a ser criticada e suas regras começaram a ser refutadas por muitos estudiosos. Esse embate já pode ser constatado em Saussure (2006 [1916]):

Começou-se por fazer o que se chamava de “Gramática”. Esse estudo, inaugurado pelos gregos, e continuado principalmente pelos franceses, é baseado na lógica e está desprovido de qualquer visão científica e desinteressada da própria língua; visa unicamente a formular regras para distinguir as formas corretas das incorretas; é uma disciplina normativa, muito afastada da pura observação e cujo ponto de vista é forçosamente estreito. (SAUSSURE, 2006 [1916], p. 7, aspas do autor)

Após o aparecimento da Sociolinguística, na década de 1960, muitos estudiosos intensificaram os estudos sobre as descrições das línguas naturais a partir de dados empíricos de fala e, dessa forma, surgiu um novo modo de descrever as regras das línguas, diferente do modelo normativo. No Brasil, já são inúmeros os volumes na literatura linguística (artigos, dissertações, teses) que contrastam análises prescritivas com análises de cunho variacionista. Há também uma gama considerável de gramáticas descritivas que, levando em conta as pesquisas das inúmeras vertentes linguísticas (sociolinguística, linguística sistêmico-funcional e outras correntes funcionalistas, análise crítica do discurso, linguística textual, entre outras), oferecem outras formas de se analisar uma língua. Só para citar alguns, seguem cinco gramáticos que se dedicaram a essa tarefa: Perini (1996, 2008, 2010), Azeredo (2010), Neves (2011), Bagno (2011) e Castilho (2012). Essas obras surgem para, além das gramáticas normativas e muitas vezes em oposição a elas, serem mais uma referência na compreensão e explicação desses sistemas tão complexos, que são as línguas naturais. Consequentemente, quem mais se beneficia dessa disputa teórico-metodológica é a escola, pois pode tomar decisões relativas ao ensino de gramática a partir de uma grande quantidade de materiais didáticos.

No que concerne ao estudo da fala, Perini (2010) justifica seu trabalho a partir da necessidade de se descrever essa modalidade de língua. Essa descrição já está sendo feita em

todo o Brasil, pois além das gramáticas citadas acima, existem vários projetos de coleta de dados da língua oral, tanto culta quanto popular. Esses projetos têm como principal objetivo traçar o perfil das comunidades de fala no Brasil, descrevendo assim a língua falada pelo povo brasileiro. Para citar alguns, elencamos os seguintes: NURC (Norma Urbana Culta do Brasil), VARSUL (Variação Linguística do Sul do Brasil), VALPB (Variação Linguística da Paraíba), PORCUFORT (Português Oral Culto de Fortaleza), NORPOFOR (Norma Oral do Português Popular de Fortaleza), LUAL (Língua Usada em Alagoas), entre outros.

As pesquisas oriundas desses projetos têm a função de conhecer a realidade linguística brasileira, o que permite fornecer ao ensino de Língua Portuguesa nas escolas uma educação linguística mais heterogênea, que abranja tanto a tradição gramatical quanto as descrições baseadas no uso da língua por seus falantes. Para a Sociolinguística Quantitativa, quanto mais estudos existirem sobre um determinado fenômeno, mais generalizações podem ser feitas e, conseqüentemente, mais abrangente será a descrição de uma língua.

O objeto de estudo deste trabalho são os verbos *existir*, *haver* e *ter* em seu sentido existencial no falar popular de Fortaleza e, para isso, utiliza uma amostra do banco de dados NORPOFOR, coordenado pela professora Dra. Aluiza Alves de Araújo (ver metodologia), e que já está sendo utilizado em vários trabalhos que têm o objetivo de descrever diferentes fenômenos do falar popular de Fortaleza. Muitos destes trabalhos já defendidos sob a orientação da referida professora estão vinculados ao seu projeto *Fotografias sociolinguísticas de Fortaleza*, desenvolvido no Programa de Pós-graduação em Linguística Aplicada (PosLA) da Universidade Estadual do Ceará (UECE). Só para citar duas pesquisas, destacamos o estudo do enfraquecimento da fricativa /v/ no falar de Fortaleza, realizado por Rodrigues (2013) e o estudo das formas de tratamento (*tu*, *você*, *cê*, *o senhor/a senhora*, além das formas nominais *macho*, *rapaz*, *mulher*, *minha filha*, *cara* e *meu amigo*), empreendido por Guimarães (2014).

Nosso objeto de estudo (*existir*, *haver* e *ter*) tem sido muito estudado pela Sociolinguística Quantitativa no Brasil (SILVA, 2001; MARTINS; CALLOU, 2003; VITÓRIO, 2011, 2012, 2013; BATISTA, 2012; RAMOS; SANTOS, 2012; RIBEIRO; SOARES; LACERDA, 2013). De uma forma geral, essas pesquisas apontam para uma tendência das comunidades de fala estudadas em privilegiar o uso do verbo *ter existencial* em detrimento de *haver existencial* no que diz respeito à língua falada tanto por pessoas de alto grau de escolarização (fala culta) quanto por pessoas de baixo nível de escolarização (fala popular).

Os termos culto e popular podem ser entendidos nos projetos de língua falada espalhados pelo Brasil (e.g. NURC, NORPOFOR) como sendo antônimos. O critério utilizado para distingui-los é o fato de os falantes possuírem nível superior completo ou não. O projeto NURC (Norma Urbana Culta do Brasil), que trabalha com norma culta, utilizou esse critério para a realização de todos os seus inquéritos. Dessa forma, todos os falantes que compõem o acervo linguístico do NURC têm nível superior completo. Já, no projeto NORPOFOR (Norma Oral do Português Popular de Fortaleza), nenhum de seus informantes possui nível superior completo, sendo o número máximo de escolaridade de 11 anos, o que equivaleria ao ensino médio completo.

Para analisar os verbos *haver*, *existir* e *ter* no falar popular de Fortaleza, definimos como objetivo geral investigar qual o lugar dos verbos *haver* e *existir* nas construções existenciais, tendo em vista uma possível mudança em curso que aponta para a supremacia do verbo *ter* em frequência de uso na língua falada em contextos existenciais, mesmo nos trabalhos que abrangem o falar culto (MARTINS; CALLOU, 2003; BATISTA, 2012). Em nossa pesquisa, partimos da hipótese de que essa supremacia será ainda maior em nossa amostra, uma vez que trabalhamos com o falar popular. Portanto, para atingir esse objetivo geral, traçamos os seguintes objetivos específicos: a) Identificar os fatores sociais e/ou linguísticos responsáveis pela manutenção de *haver* e *existir* no falar popular de Fortaleza; b) Investigar, a partir de análises no tempo aparente (faixa etária), se há indícios de esses verbos estarem passando por um processo de mudança em curso no sentido de serem substituídos pelo verbo *ter*.

Este trabalho é relevante para as pesquisas sobre os verbos existenciais porque tenta preencher uma lacuna: dos três trabalhos que incluem o verbo *existir* (FRANCHI; NEGRÃO; VIOTTI, 1998; VITÓRIO, 2013; RIBEIRO; SOARES; LACERDA, 2013) que encontramos, nenhum deles investiga os fatores favorecedores desse verbo em relação a *haver* e a *ter* como fazemos aqui. O primeiro não é variacionista e apresenta apenas as frequências de uso de cada verbo, pois tem como objetivo descrever as estruturas existenciais. O segundo estudo é variacionista, mas também apresenta apenas as frequências de uso, não informando ao leitor quais os fatores selecionados pelo GoldVarb X como relevantes. E, por último, o terceiro estudo também é variacionista e, seguindo os outros dois, só trabalha com a frequência de uso. As autoras justificam essa escolha da seguinte forma: “Na análise quantitativa, não consideramos os resultados dos pesos relativos, em razão da dificuldade em calculá-los a partir de uma variável dependente terciária via GoldVarb 2001 e do excesso de

knockouts.” (RIBEIRO; SOARES; LACERDA, 2013, p. 559). Nesta análise, elas também não apresentam os fatores selecionados pelo programa como favorecedores dos verbos; apenas fazem inferências sobre a preferência dos falantes a partir das frequências de uso.

No que diz respeito ao panorama da comunidade de fala estudada, este estudo também se mostra relevante, uma vez que amplia o conhecimento do português popular falado em Fortaleza, juntando-se aos já citados trabalhos de Rodrigues (2013) e Guimarães (2014), realizados com dados dessa mesma variedade de língua. Além de ampliar a descrição do português popular de Fortaleza, nosso trabalho também serve para ampliar geograficamente a representação dos estados brasileiros sobre os verbos existenciais. Com este estudo, passam a ser 8 estados da federação representados, englobando 3 Regiões do país: Alagoas, Bahia, Ceará, Maranhão e Pernambuco (Nordeste); Minas Gerais e Rio de Janeiro (Sudeste) e Rio Grande do Sul (Sul). (Ver APÊNDICE A)

Vale ressaltar outra relevância de nossa pesquisa, que diz respeito ao ensino de língua portuguesa. Nosso tema é muito abordado no ensino, uma vez que o aluno apresenta muitas dúvidas em relação ao uso dos verbos *haver* e *existir*, principalmente no que diz respeito à concordância desses verbos: o verbo *haver* é impessoal e, portanto, não varia em sua concordância porque seu complemento é objeto direto (e.g. *há vários alunos* de recuperação) enquanto que o verbo *existir* não é impessoal e seu complemento é o sujeito da oração, portanto, concorda em número com esse complemento (e.g. *existem vários alunos* de recuperação). Em relação ao verbo *ter* com acepção de *existência*, normalmente a escola reprime o uso desse verbo e os alunos são orientados a escolher ou *haver* ou *existir*. Isso explica o porquê de esses dois verbos serem tão frequentes na língua escrita, diferente do que ocorre na língua falada. Vitória (2013), por exemplo, mostra esse contraste ao apresentar seus dados: retirando a variante *ter pessoal*, a autora analisou 424 ocorrências de *haver*, *ter* e *existir* na língua escrita, dos quais, 298 (70%) são do verbo *haver*, 105 (25%) são do verbo *existir* e apenas 21 (5%) são do verbo *ter*.

Portanto, os resultados deste trabalho colaboram para uma melhor compreensão das variantes *haver*, *existir* e *ter* na variedade popular e conseqüentemente explica as dificuldades que os alunos têm no aprendizado formal desses verbos.

Este trabalho está dividido em 6 seções: nesta introdução, apresentamos nosso objeto de estudo e nossos objetivos de pesquisa; a seção 2 fornece uma visão geral da variação *haver/ter/existir*, fazendo um breve percurso histórico desses verbos, mostrando como eles são abordados em algumas gramáticas e, por fim, descrevendo os principais

estudos variacionistas sobre o tema; a seção 3 apresenta os procedimentos metodológicos que utilizamos no trabalho; na seção 4, discutimos os principais pressupostos teórico-metodológicos da Teoria da Variação e Mudança Linguística; na seção 5, realizamos a descrição dos dados e a discussão dos resultados e, por fim, na seção 6, apresentamos algumas considerações finais acerca do trabalho, expondo claramente suas limitações e indicando possíveis maneiras de superá-las.

2 A VARIAÇÃO DOS VERBOS EXISTENCIAIS *Haver/Existir/ter* NO PORTUGUÊS BRASILEIRO

2.1 BREVE INCURSÃO HISTÓRICA

Antes de o verbo *ter* começar a ser usado nos contextos existenciais no lugar de *haver*, ele já havia disputado espaço com esse verbo nas estruturas de posse. Batista (2012, p. 21) afirma que “atribui-se ao século XIII o primeiro registro de *ter* com o valor de posse, contexto até então restrito a *haver*.” Dessa forma, ao passo que *haver* foi perdendo espaço nas estruturas de posse, passou então a disputar com o verbo *ser* um lugar nas estruturas existenciais. “Aos poucos, *ter*, que já predominava como verbo de posse, começa a dividir com *haver* a significação existencial. Os dois verbos permanecem em variação até hoje, com predominância deste na escrita e daquele na fala, principalmente dos indivíduos mais jovens” Batista (2012, p. 22).

Apesar de a variação *haver/ser existenciais* ter ocorrido até o século XVI, com o domínio de *haver*, este verbo só aparece nas gramáticas de língua portuguesa no século XVIII. Argote (1725) já registra esse verbo no sentido existencial em sua obra *Regras da Lingua Portuguesa* e afirma que o complemento de *haver* é um nominativo e, portanto, deve haver concordância entre os dois elementos (verbo + nominativo).

Lobato (1770) tem a mesma visão de Argote (1725) no que diz respeito à concordância de *haver* com seu *nominativo (sujeito)*, ou seja, deve haver a concordância. A diferença está entre o ponto de vista em relação às realizações sem essa concordância: enquanto que este vê esse fenômeno como um idiotismo, aquele o explica através de uma elipse:

Parece que se não observa a dita regra no verbo *Haver*, quando pelo costume da língua lhe ajuntamos na terceira pessoa do singular nominativo plural, como v.g. quando dizemos: *Ha muitos homens, que amam as ciências*: onde parece que ao verbo *Ha* na terceira pessoa do singular do presente do indicativo, lhe serve de nominativo *Homens* no numero plural: porém não he assim, por faltarem palavras, que se sobentendem, e quer dizer: *Ha numero de pessoas, que são muitos homens, que amam as ciências*: onde claramente se vê, que ao verbo *Ha* só lhe serve de nominativo o substantivo occulto *numero*, com quem concorda em estar no numero singular e na terceira pessoa. (LOBATO, 1770, p. 197-198)

Percebemos em Lobato (1770) uma pequena abordagem voltada ao uso da língua, uma vez que ele menciona o termo “costume da língua” e, por isso, oferece-nos uma

explicação ao não cumprimento da regra de concordância. O autor não menciona o verbo *ter existencial*. O mesmo faz Barbosa (1822, p. 196), que só fala de *ter* como auxiliar. Em relação a *haver*, ele reconhece que, além de sua função de auxiliar, “como impessoal, significa existência, como quando digo: *Ha muitos homens, Havia muita gente.*”.

Vemos a partir dos exemplos que esse gramático já não aborda a concordância de *haver* com seu complemento. Pode-se pensar que nessa época já se havia registrado nas gramáticas o que acontecia na fala: como impessoal, o complemento do verbo *haver* era acusativo e, portanto, não havia concordância. Mas Freire (1842, p. 69), assim como Argote (1725), já volta a mencionar a não concordância desses elementos como um *idiotismo*, afirmando que “Na concordância do verbo com o seu nominativo temos também um particular idiotismo no verbo *Haver*: porque nas terceiras pessoas do numero singular não concorda em numero com o seu nominativo.”. Aqui também não temos o reconhecimento do verbo *ter* como existencial.

Nas gramáticas normativas mais atuais, o verbo *haver*, como impessoal, já não concorda com seu complemento, sendo a forma padrão para representar existência em comparação com o verbo *ter*, ainda não aceito pela tradição gramatical como um verbo existencial. Só para citar dois exemplos de gramáticas normativas atuais, abordaremos Cunha e Cintra (2008) e Bechara (2009).

Cunha e Cintra (2008, p. 144-145), ao tratarem das orações sem sujeito, reconhecem o uso de *ter existencial* na “linguagem coloquial do Brasil”. Os autores também afirmam que “escritores modernos – e alguns dos maiores – não têm duvidado em alçar a construção à língua literária”, citando exemplos de Carlos Drummond de Andrade e Manuel Bandeira, no Brasil e Luandino Vieira, em Angola.

Bechara (2009, p. 562) não menciona o verbo *ter existencial* nem como um desvio de norma, apenas comenta que a concordância dos verbos impessoais (entre ele o *haver*) deve ser feita sempre na 3ª pessoa do singular. No que diz respeito aos usos desses verbos na 3ª pessoa do plural, o autor afirma que “os exemplos literários que se encontram de tais verbos no plural não ganharam fotos de cidade”; e exemplifica com uma ocorrência da obra de Camilo Castelo Branco e outra da obra “*Vidas e feitos Del-Rei D. Manuel*” do historiador português Jerónimo Osório.

Como podemos ver, o percurso histórico dos verbos *haver* e *ter* mostra que seus usos sempre se cruzaram e que isso não é uma característica de seus significados existenciais.

2.2 ESTUDOS VARIACIONISTAS RELEVANTES NO PORTUGUÊS BRASILEIRO

Os trabalhos a respeito da variação *haver/existir/ter* no Português Brasileiro são diversos e abordam desde aspectos não variacionistas até aspectos variacionistas, analisando ora *ter/haver*, ora *ter/haver/existir*, ora incluindo outras variantes, como *ter pessoal* (MARTINS; CALLOU, 2003; VITÓRIO, 2013). Há também trabalhos que investigam, sob a ótica histórica, outros contextos de *ter* e *haver*, além do existencial, como nos contextos de posse, participios e modais (ALMEIDA; CALLOU, 2003; MATOS E SILVA; MACHADO FILHO, 2009; OLIVEIRA, 2010).

Os trabalhos que consideramos não variacionistas são aqueles que, embora utilizem dados de estudos variacionistas, não têm objetivos de descrever o comportamento dos verbos existenciais numa comunidade de fala específica e nem de verificar quais os fatores linguísticos e/ou extralinguísticos influenciam no uso das variantes. Seus objetivos são diversos, entre eles: verificar o percurso histórico da variação (ALMEIDA; CALLOU, 2003; MATOS E SILVA; MACHADO FILHO, 2009; OLIVEIRA, 2010; CALLOU; AVELAR, 2012); analisar a estrutura das orações existenciais (FRANCHI; NEGRÃO; VIOTTI, 1998); contrastar a descrição dos estudos sociolinguísticos com a gramática normativa (VITÓRIO, 2010; COSTA *et al.*, 2011); discutir o processo de aquisição de linguagem, fazendo intertextualidade com o gerativismo (AVELAR, 2006; VITÓRIO, 2010).

É importante ressaltar que mesmo aqueles trabalhos que analisam os dados a partir do ponto de vista variacionista utilizam procedimentos metodológicos diferentes; alguns analisam somente a frequência de uso de cada verbo enquanto outros se interessam em descobrir os fatores linguísticos e/ou sociais responsáveis pela variação desses verbos, utilizando tanto a frequência de uso quanto outros valores estatísticos (pesos relativos, *input*, significância, entre outras).

No que diz respeito às variedades de fala e de escrita, as pesquisas variacionistas que encontramos também são variadas: **a)** língua escrita (VITÓRIO, 2007, 2008); **b)** língua falada e escrita (VITÓRIO, 2013); **c)** língua falada (SILVA, 2001; MARTINS; CALLOU, 2003; VITÓRIO, 2011, 2012; BATISTA, 2012; RAMOS; SANTOS, 2012; RIBEIRO; SOARES; LACERDA, 2013). Como nosso estudo analisa a língua falada, deter-nos-emos apenas na descrição dos estudos **b)** (utilizaremos apenas os dados de fala dessa pesquisa) e **c)**. Desses oito estudos variacionistas que descreveremos a seguir, somente Vitório (2013) e Ribeiro, Soares e Lacerda (2013) tratam da variável ternária *haver/existir/ter*, assim como

nós. Além desses, os dados gerais de Franchi, Negrão e Viotti (1998), apesar de não serem de cunho variacionista, serão mencionados em nosso trabalho.

A seguir, faremos uma pequena descrição das oito pesquisas variacionistas que tratam da língua falada e que contribuem para a descrição do fenômeno variável *ter/haver* (ou *ter/haver/existir*) no Português Brasileiro. Todos esses trabalhos são do início do século até o presente momento e estudam o fenômeno utilizando os bancos de dados dos projetos NURC (MARTINS; CALLOU, 2003; BATISTA, 2012), VALPB (SILVA, 2001), LUAL (VITÓRIO, 2011), ALiMA (RAMOS; SANTOS, 2012) e do Projeto Mineirês (RIBEIRO; SOARES; LACERDA, 2013). O *corpus* utilizado por Vitório (2012, 2013) foi constituído por ela própria com o objetivo de realizar seu doutoramento e uma amostra desse *corpus* foi utilizada em seu pós-doutorado. Todas as pesquisas expostas a seguir abrangem seis estados do Brasil, Rio de Janeiro e Minas Gerais (SUDESTE), Rio Grande do Sul (SUL), Bahia, Paraíba, Alagoas e Maranhão (NORDESTE). Dessa forma, temos 3 regiões do país representadas por todos esses estudos. Alguns trabalhos utilizaram o pacote de programas computacionais VARBRUL (MARTINS; CALLOU, 2003; SILVA, 2001; BATISTA, 2012; RAMOS; SANTOS, 2012; RIBEIRO; SOARES; LACERDA, 2013) e outros utilizaram a versão mais atualizada desse pacote de programas, o GOLDVARB X (VITÓRIO; 2011, 2012, 2013).

Silva (2001) analisou 1057 ocorrências de *ter/haver existenciais* de 60 informantes do projeto VALPB com o objetivo de descrever o comportamento desses verbos e verificar se, na comunidade estudada, esse fenômeno encontra-se em variação estável ou em mudança em progresso. No cômputo geral, 90% desses dados se referem à variante *ter* existencial, enquanto que 10% dizem respeito a *haver* existencial. Para a análise mais específica, foram testadas oito variáveis (Posição do SN Objeto em Relação ao Verbo, Realização do SN Objeto, Peso do SN Objeto, Concordância entre V e SN Objeto, Animacidade do SN Objeto, Anos de Escolarização, Faixa Etária, Sexo), mas somente quatro foram selecionadas pelo programa como significativas na variação *ter/haver existenciais*, apresentados aqui por ordem de relevância: anos de escolarização, faixa etária, animacidade do SN objeto e sexo.

Em relação à variável anos de escolarização, são os falantes não escolarizados (0,82) que mais aplicam o verbo *ter existencial*. Em relação à faixa etária, os falantes que mais favorecem o uso de *ter* são os falantes da faixa intermediária, de 26 a 49 anos (0,74), ao passo que os falantes mais jovens, de 15 a 25 anos (0,39) e os mais velhos, acima de 50 anos (0,36) inibem a sua aplicação. A variável sexo, quarto e último grupo selecionado como

relevante, indica que as mulheres (0,57) favorecem mais a variante *ter existencial* do que os homens (0,47). No que concerne à única variável linguística selecionada como influente no fenômeno em estudo, a animacidade do SN objeto ([+animado], [-animado]), os resultados apontam para o traço [+animado] (0,72) como o mais influente para a aplicação de *ter*. Depois de analisar as variantes nas faixas etárias e de fazer o cruzamento das variáveis sociais, a autora conclui que a variação *ter/haver* na fala de João Pessoa encontra-se em estado de estabilidade, não indicando, assim, indícios de mudança em progresso.

Martins e Callou (2003), com o objetivo de observar possíveis mudanças no tempo aparente e no tempo real, analisaram 2036 construções existenciais do projeto NURC das décadas de 70 e 90. As autoras incluíram no estudo as construções em que os verbos *ter* e *haver* apresentam o sentido existencial, o sentido de ocorrer ou acontecer e também incluíram as construções em que o verbo *ter* tem um sujeito expletivo, mas ainda assim com valor existencial, o que Vitório (2013) chama de *ter pessoal*. Os fatores que se mostraram relevantes no estudo nas duas décadas foram: o tempo verbal, a especificidade semântica do argumento interno, a faixa etária e o sexo. As autoras não afirmam se essas variáveis estão listadas em ordem de relevância pelo programa. Além disso, não há um detalhamento no estudo de cada variável. Acreditamos que isso se deve ao fato de este trabalho ser um artigo publicado em anais de congresso e por isso apresenta somente informações gerais da pesquisa.

Os resultados sugeriram uma mudança em progresso em favor da implementação de *ter existencial* como verbo preferido pelos falantes das duas capitais, pois o percentual de uso desse verbo aumenta da década de 70 para a década de 90: passa de 63% para 76% no Rio de Janeiro e de 74% para 86% em Salvador. Nesta capital, o uso de *ter* pelas mulheres de 25 a 35 anos (faixa etária 1) e de 36 a 55 anos (faixa etária 2) é categórico na década de 90 (100% das ocorrências) e as mulheres acima de 55 anos (faixa etária 3) apresentam percentual de 97%. Já os homens soteropolitanos da faixa etária 3 são mais resistentes a essa mudança, apresentando apenas 45% do uso de *ter existencial*. As faixas 1 e 2 do homens, apesar de não apresentarem uso categórico, se assemelham ao comportamento das mulheres dessas mesmas faixas, com percentuais de mais de 90% cada.

As autoras afirmam que o comportamento linguístico de homens e mulheres do Rio de Janeiro é semelhante, mas não especificam tal semelhança. Em relação à variável especificidade semântica do argumento interno (material, humano, evento e abstrato), o verbo *ter* é favorecido pelos traços material e humano e o verbo *haver*, embora com menos

ocorrências, é favorecido pelos traços evento e abstrato nas duas capitais. No que diz respeito à variável tempo verbal, no Rio de Janeiro, as construções no passado favorecem o verbo *haver*, enquanto que as construções do presente favorecem *ter* nas duas décadas. Em Salvador, as ocorrências no presente e no pretérito perfeito se comportam de igual modo às da fala carioca, isto é, as primeiras favorecem o verbo *ter* e as segundas favorecem *haver*. No entanto, o imperfeito apresenta comportamento diferente, pois favorece o verbo *ter* na década de 70 e *haver* na década de 90.

Vitório (2011) estuda o fenômeno variável *ter/haver* em Maceió, utilizando dados de 20 informantes do sexo feminino, estratificados em dois níveis de escolaridade (de 0 até a 4ª série e nível superior completo). A autora analisa 209 ocorrências das quais 93% são do verbo *ter* e somente 7% são do verbo *haver*, mostrando que a variante *ter* é a mais utilizada pelos falantes. As variáveis controladas foram tempo verbal, animacidade do SN objeto, natureza do SN objeto e escolaridade, mas somente duas destas foram selecionadas, por ordem de relevância: a escolaridade e o tempo verbal (presente e pretérito). No grupo de fatores escolaridade, o que mais beneficia o uso de *ter existencial* (com peso relativo de 0,73), como já esperado pela autora e já comprovado em pesquisas de outros autores (SILVA, 2001), é a escolaridade mais baixa (os falantes analfabetos ou os que cursaram até a 4ª série), enquanto que os falantes de nível superior favorecem o emprego de *haver existencial* (0,69). No que se refere ao tempo verbal, o presente favorece *ter* (0,68) e o passado privilegia *haver* (0,70), confirmando mais uma vez outra hipótese da autora.

Vitório (2012) analisou a variação *ter/haver* existenciais na fala alagoana com o objetivo de verificar se essa variação encontra-se em fase estável ou se se trata de uma mudança em progresso. A autora usou um *corpus* com 72 informantes, estratificados da seguinte forma: sexo (masculino e feminino), faixa etária (15-29 anos, 30-44 anos, mais de 44 anos) e escolaridade (ensino fundamental, ensino médio, ensino superior). Foram testadas as variáveis: animacidade do argumento interno, natureza do argumento interno, tempo verbal, sexo, faixa etária e escolaridade. Dessas, o VARBRUL selecionou, por ordem de relevância: escolaridade, tempo verbal, faixa etária e natureza do argumento interno.

O resultado geral indica a supremacia da variante *ter existencial* sobre *haver existencial* (95% e 5%, respectivamente). Quanto à escolaridade, o uso da variante não-padrão (*ter*) é quase categórico na educação básica (98%); na educação superior, o percentual também é alto, mas, ao analisar o peso relativo do verbo *haver existencial* nessa variável, Vitório (2012) constatou que, embora com menos ocorrência, essa variante apresentou um

peso relativo de 0,73. Isso mostra que a alta escolaridade é um fator favorecedor da variante padrão, corroborando outras pesquisas (DUTRA, 2000; SILVA, 2001; VITÓRIO, 2008), que atestam que quanto menor o nível de escolaridade do falante maior será o uso de *ter existencial*.

Ao cruzar as variáveis escolaridade e sexo, Vitório (2012) verificou que são os homens quem mais usam *haver existencial*. Com relação ao tempo verbal, o passado favorece mais o uso de *haver* (0,77), já o presente beneficia o uso de *ter* (0,57). Em relação à faixa etária, são os falantes de f1 (15 a 29 anos) (0,80) que mais favorecem a variante inovadora *ter*, ao passo que as outras duas faixas etárias (f2: 0,35 e f3: 0,31) não favorecem o uso dessa variante. No tocante à natureza do argumento interno (concreto ou abstrato), a autora constatou que os argumentos concretos (0,59) favorecem o uso de *ter existencial* enquanto que os argumentos abstratos (0,60) favorecem *haver existencial*. A autora conclui que, na comunidade estudada, a variação observada aponta para uma mudança em progresso em favor de *ter*, mas ressalta que ainda há variáveis que favorecem o uso de *haver*.

Batista (2012) descreve o uso de *ter/haver* no falar culto de três capitais das cinco que compõem o projeto NURC, a saber, Rio de Janeiro, Salvador e Porto Alegre. A autora analisou 1283 dados de construções existenciais prototípicas das décadas de 70 e 90, num estudo de tendência com o objetivo de fazer uma análise em tempo real de curta duração (LABOV, 2008 [1972]). O número de informantes, bem como a sua distribuição, não é mencionado no corpo do trabalho, mas eles foram estratificados por idade e gênero/sexo. As variáveis testadas foram: faixa etária, década de gravação, sexo/gênero, tempo verbal, natureza semântica do argumento e cidade do falante, não necessariamente na ordem que o programa as selecionou, pois essa informação não é dada no texto.

A autora apresenta os resultados em porcentagens, sem menção aos pesos relativos, e infere que o verbo *ter*, no *corpus* estudado, está em processo de mudança nas três cidades, apesar de haver algumas diferenças específicas em cada uma delas. No resultado geral (englobando as três cidades e as duas décadas), o verbo *ter* tem 75% das ocorrências contra 25% do verbo *haver*. Ao analisar a frequência de uso do verbo *ter*, levando em consideração sua aplicação em cada cidade e em cada década, a autora encontrou resultados coerentes com os percentuais gerais apresentados acima, pois essa variante sempre apresenta um percentual superior a *haver*. A menor frequência de uso de *ter* é 63% no Rio de Janeiro na década de 70 e a maior é 85% em Salvador na década de 90. Vale ressaltar que há sempre um aumento percentual do verbo *ter* de uma década para a outra nas três cidades.

Em todas as variáveis analisadas, o uso de *ter* é superior ao uso de *haver*, mas esse último ainda resiste em alguns contextos. De forma geral, juntando os dados das três cidades, os resultados revelam que a faixa 1 é a que mais aplica o uso de *ter* nas duas décadas (82% e 99%, respectivamente em 70 e 90), enquanto que a faixa 3 é a que menos aplica seu uso (59% e 68%, respectivamente em 70 e 90). Em relação ao tempo verbal, as maiores frequências de uso de *ter* nas três cidades são do presente e do pretérito imperfeito, sendo o maior percentual em Salvador (85% e 88%, respectivamente) e o menor no Rio de Janeiro (70% e 65%, respectivamente). Em relação à natureza semântica do argumento interno, os traços que mais promovem o uso de *ter* são [+humano] e [+material], ao passo que os traços [+abstrato] e [+evento] favorecem o uso de *haver*. Finalmente, a autora conclui que os resultados apontam para uma mudança em curso, nas três cidades, em favorecimento do verbo *ter*.

Ramos e Santos (2012) estudam o fenômeno variável *ter/haver* na fala maranhense, utilizando o banco de dados do ALiMA (Atlas Linguístico do Maranhão), que conta com 40 informantes, estratificados por localidade (8 informantes de São Luís e 4 informante das outras 8 localidades), sexo (feminino e masculino) e faixa etária (F1: 18-30 anos; F2: 50-65 anos). Os informantes do ALiMA possuem no máximo até a 4ª série do ensino fundamental (hoje, 5º ano), com exceção de 4 universitários de São Luís. Foram analisadas as variáveis localidade, sexo e faixa etária, animacidade do argumento interno e tempo verbal, com o objetivo de verificar a frequência de uso de *ter/haver* no *corpus*. Os resultados gerais, como nos trabalhos anteriores, indicam a supremacia de *ter* (97,5%) sobre *haver* (2,4%). Nos resultados da variável localidade, os autores constataram que o número maior de ocorrências (três) ocorreu em São Luís, exatamente na capital onde foram selecionados falantes universitários. Em relação à faixa etária e ao sexo, os idosos (6 ocorrências) e as mulheres (5 ocorrências) usam mais *haver* do que os jovens (2 ocorrências) e os homens (3 ocorrências), respectivamente. No que diz respeito às variáveis linguísticas, há também poucas ocorrências de *haver*, sendo o total de 8 ocorrências para cada variável. Na variável animacidade do argumento interno, o número de ocorrências é maior com o traço [-animado], com 5 ocorrências em relação ao traço [+animado], com 3 ocorrências. E em relação ao tempo verbal, *haver* ocorre mais no tempo presente (5) do que no tempo pretérito (3). Os autores ressaltam que as ocorrências de *haver existencial* nas duas variáveis linguísticas são favorecidas pelos falantes universitários do *corpus* e pela motivação do inquiridor quando este inclui essa variante na pergunta.

Ribeiro, Soares e Lacerda (2013) realizam sua pesquisa sobre os verbos *ter*, *haver* e *existir* no falar de Minas Gerais, utilizando o *corpus* que abrange 6 cidades: Arceburgo, Belo Horizonte, Mariana, Ouro Preto, Piranga e São João da Ponte. As autoras analisaram 1427 ocorrências das três variantes, das quais 93,3% são do verbo *ter*, 4,1% de *existir* e 2,6% de *haver*. Este trabalho não apresenta uma seleção das variáveis relevantes para o fenômeno em estudo, detendo-se apenas em expor as porcentagens de cada variável testada no estudo. Portanto, o que nos interessa aqui são somente esses dados gerais. Eventualmente, podemos usar algumas dessas frequências de uso para contrapor as dos nossos resultados.

Vitório (2013) analisa a variação *ter/ter pessoal/haver/existir* em dados de fala culta e escrita acadêmica com o objetivo de identificar evidências da implementação de mudanças da fala na escrita. As variáveis analisadas são: faixa etária, sexo, tempo verbal, especificidade semântica do argumento interno, ausência *versus* presença de elementos à esquerda do verbo, tipos de preenchedores à esquerda do verbo. Aqui só nos interessará os dados de língua falada. Vale ressaltar que Martins e Callou (2003) também analisaram construções de *ter pessoal*, no entanto, não as incluíram como uma variante. Dessa forma, esse tipo de ocorrência foi estudado como ocorrências de *ter*. Ao separar as ocorrências de *ter* das de *ter pessoal*, Vitório (2013) inclui mais uma variante em sua análise. Em nossos dados, só houve uma ocorrência de *ter pessoal* e por isso essa variante não foi analisada.

Os resultados gerais da análise de 381 dados indicam os seguintes resultados: *ter* (59%), *ter pessoal* (26%), *haver* (8%) e *existir* (7%). Em relação à variável tempo verbal, os dados mostram que o tempo presente apresenta o maior percentual de todos os verbos em relação aos demais tempos. Esse resultado não condiz com os de Vitório (2012), que apresenta o tempo passado como maior condicionador do verbo *haver*. No que concerne à especificidade semântica do argumento interno, o traço [+ abstrato] apresentou mais dados, dentre todos os verbos tanto na fala quanto na escrita. Esse resultado também não vai de encontro com outras pesquisas (AVELAR, 2006; VITÓRIO, 2008; VITÓRIO, 2012) que apontam o traço [-abstrato] como maior condicionador do verbo *ter*. Ainda em relação a essa variável, os verbos *haver* e *existir* apresentam dados muito semelhantes na fala e na escrita. O verbo *haver*, corroborando outras pesquisas, aumenta sua frequência de uso quando o argumento interno é [+abstrato] em ambas as modalidades. Em relação à variável ausência *versus* presença de elementos à esquerda do verbo, foram encontradas muito mais ocorrências do fator presença (85% fala; 84% na escrita) nos dados analisados. Os elementos de negação e

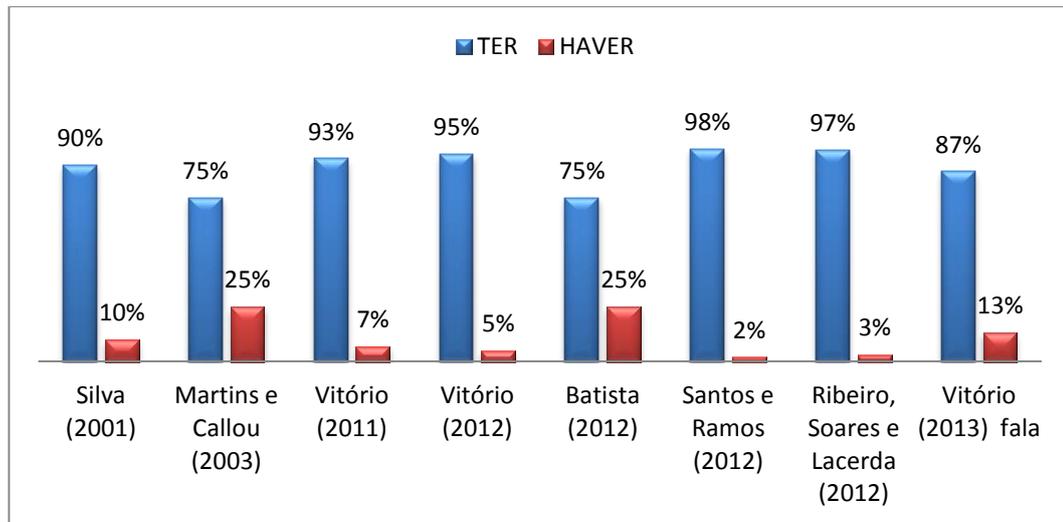
dos DPs¹ pronominais são os que apresentam maior percentual (23% e 20%, respectivamente). De acordo com a autora, esse resultado corrobora a afirmação de Duarte (1995) de que a tendência no Português Brasileiro é a perda do Princípio Evite Pronome e que isso possibilita inclusive o preenchimento pronominal de sentenças não-argumentais, como é o caso das existenciais. No que concerne às variáveis sociais, este estudo avaliou a faixa etária e o sexo, mas somente a primeira mostrou-se significativa em ambos os *corpora*, que foi dividida da seguinte forma: a faixa etária 1 agrega falantes de 15 a 29 anos; a faixa etária 2, de 30 a 44 anos e a faixa etária 3, acima de 44 anos. Os resultados apontam para a queda nas realizações do verbo *ter* à medida que a faixa etária aumenta (77%, 59% e 43%, respectivamente) e, em comportamento oposto, apontam para o aumento das ocorrências dos outros três verbos (*ter* pessoal – 19%, 24% e 32%; *haver* – 2%, 10% e 12%; *existir* – 2%, 7% e 10%) com o aumento da faixa etária. A autora conclui que as faixas 1 e 2 são as responsáveis pelas poucas realizações.

As sínteses das 8 pesquisas escolhidas como base para nossas análises nos revelam um problema para o estudo dos verbos existenciais: apesar de todas se fundamentarem nos procedimentos teórico-metodológicos da sociolinguística variacionista, cada pesquisador ou grupo de pesquisadores adota(m) um sistema analítico diferente (variável dependente e grupos de fatores condicionadores), que depende da intuição e/ou conhecimento acumulado acerca do tema que cada um possui. Esse fato não invalida os resultados de tais estudos, que são específicos de cada comunidade de fala, mas impossibilita comparações diretas entre os falares de todas as comunidades linguísticas representadas de modo que pudéssemos fazer generalizações mais precisas acerca do fenômeno. No entanto, é possível fazer triangulações acerca dos resultados, porque, mesmo utilizando procedimentos metodológicos distintos, os estudos apontam para resultados semelhantes.

Um exemplo de uma possível generalização que podemos fazer, a despeito das diferenças analíticas, é que os resultados gerais de todos os estudos para *haver* e *ter* indicam a grande preferência dos falantes pela variante *ter existencial* em detrimento de *haver existencial*, como podemos perceber pelo gráfico 1:

¹ Sigla do inglês (*determiner phrase*) usada pela autora, que significa “sintagma determinante”.

Gráfico 1 – Resultados gerais da frequência de uso para *ter* e *haver* em estudos variacionistas



Vemos que os únicos estudos que apresentam uma frequência de uso de *haver* superior a 20% são exatamente os de Martins e Callou (2003) e de Batista (2012), que trabalham com dados de fala de falantes que possuem um alto grau de escolarização (nível superior completo). E mesmo assim, o verbo *ter* é o verbo mais utilizado. A descrição desses oito estudos pautou a escolha de 10 dos 12 grupos de fatores analisados neste estudo.

3 A TEORIA DA VARIAÇÃO E DA MUDANÇA LINGUÍSTICA

Em 1966, William Bright publicou, sob o título de *Sociolinguistics*, os trabalhos que haviam sido apresentados dois anos antes num congresso da UCLA (Universidade da Califórnia em Los Angeles). Estava assim estabelecido um novo campo de estudos linguísticos denominado Sociolinguística, que já nasce interdisciplinar, como Calvet (2002) nos apresenta:

Os temas abordados [pela sociolinguística] eram variados: a etnologia da variação linguística (Gumperz), a planificação linguística (Haugen), a hipercorreção como fator de variação (Labov), as línguas veiculares (Smarin, Kelly), o desenvolvimento de sistemas de escrita (Sjoberg), a equação de situações sociolinguísticas dos Estados (Ferguson)... e os referenciais teóricos não eram menos variados. (CALVET, 2002, p. 29).

A respeito dessa heterogeneidade dos estudos sociolinguísticos na década de 1960, Alkmim (2008, p. 43) destaca algumas áreas também ligadas à Antropologia e à Sociologia, com seus respectivos representantes. Dentre estas, estão a Sociologia da Linguagem (J. Fishman), Sociolinguística Interacional (J. Gumperz), a Dialectologia Social (R. Shuy e P. Trudgil), a Etnografia da Comunicação (D. Hymes). A autora também inclui os “trabalhos de R. Bauman e J. Sherzer, voltados, particularmente, para a questão da arte verbal e da poética dos gêneros de fala”.

Atualmente, no Brasil, além da vertente Sociolinguística Variacionista oriunda dos estudos de Weinreich, Labov e Herzog (1968) e de Labov (1963, 2006 [1966], 1969, 1978, 2008 [1972], 1994, 2006 [2001], 2010), é importante citar, entre outras correntes, a Sociolinguística Educacional. Um de seus representantes é a linguista Maris Stella Bortoni-Ricardo, professora titular da Universidade de Brasília (UnB). Bortoni-Ricardo (2005, p. 147) define esse ramo de estudos linguísticos da seguinte forma:

[...] ramo das ciências sociais que faz interface com a linguística, a pragmática, a antropologia (na subárea de etnografia) e a sociologia (nas subáreas de etnometodologia e de análise da conversação), entre outras. Trata-se de um paradigma de base fenomenológica, interpretativa, que apresenta um arcabouço teórico interdisciplinar e uma metodologia bastante refinada para a descrição dos fenômenos da interação humana. (BORTONI-RICARDO, 2005, p. 147)

Voltando ao início desta seção, Bright (1966, p.17), ao fazer uma conjectura dos estudos sob a ótica da interação língua e sociedade da época, apresenta a tarefa da sociolinguística, que “é, portanto, demonstrar a covariação sistemática das variações

linguística e social, e, talvez, até mesmo, demonstrar uma relação causal em uma ou outra direção”. Estava pronto o terreno propício para a reflexão sobre os fundamentos que norteavam os estudos sociolinguísticos, que tinham como missão descreverem as línguas naturais como um fenômeno indiscutivelmente social e heterogêneo.

Dessa forma, as línguas naturais não deveriam ser estudadas fora de sua comunidade de fala e a heterogeneidade, antes um problema para a análise linguística, passou a ser vista como constitutiva das línguas. Dentre as vertentes que associavam a língua à sociedade estavam a Sociologia da Linguagem, a Etnografia da Comunicação e a Sociolinguística Variacionista. Nesta pesquisa, nos interessará somente a Sociolinguística Variacionista ou Teoria da Variação e Mudança Linguística.

A Teoria da Variação e Mudança Linguística surge a partir de um ensaio de Uriel Weinreich, William Labov e Marvin I. Herzog intitulado *Empirical foundations for a theory of language change*. Esse texto foi escrito para uma conferência na Universidade do Texas, Estados Unidos, e foi publicado em 1968, juntamente com outros textos dessa conferência que recebeu o nome de *Directions for Historical Linguistics: a symposium*, editados por Winfred P. Lehmann and Yakov Malkiel.

As ideias apresentadas por esses autores englobam trabalhos do próprio Labov (1963, 1966), dos estudos de dialetologia de Herzog (1965 *apud* LABOV, 2008[1972], p. 15) e da visão ampla sobre história da linguística de Weinreich, autor do atlas lingüístico e cultural dos judeus asquenazes (LABOV, 2008[1972], p. 15). Weinreich, Labov e Herzog (1968) lançam os princípios empíricos que governam a teoria que se propõe analisar a variação e a mudança linguística, levando em consideração que esses dois fenômenos eram inerentes a qualquer sistema linguístico e a mudança poderia ser observada mesmo antes de se concretizar. A partir dessa nova concepção do fazer da linguística, o estudo da linguagem, até então baseado em uma concepção de língua abstrata, homogênea e ideal (SAUSSURE, 2006 [1916]; CHOMSKY, 1957) voltou-se para a fala, um objeto concreto, que fora excluído dos estudos linguísticos estruturalistas e gerativistas:

Existe uma crescente percepção de que a base do conhecimento intersubjetivo na lingüística tem de ser encontrada na **fala** – a língua tal como usada na vida diária por membros da ordem social, este veículo de comunicação com que as pessoas discutem com seus cônjuges, brincam com seus amigos e ludibriam seus inimigos. (LABOV, 2008 [1972], p. 13, grifo nosso)

Weinreich, Labov e Herzog (1968) dão enfoque a duas questões centrais em relação à mudança linguística, a saber: a) mudanças completas de um século ou dois, no

máximo; b) processos de mudança em curso que podem ser observados em uma ou duas gerações. Para analisar essas questões no que diz respeito à variação e à mudança linguística os autores discutem 5 problemas, que devem ser resolvidos para a construção de fundamentos empíricos para uma teoria que pretende tratar da mudança. São eles: a restrição, a transição, o encaixamento, a avaliação e a implementação.

O problema da restrição se refere aos fatores sociais e/ou linguísticos que restringem ou favorecem o conjunto de mudanças possíveis no sistema e a direção que essa mudança pode seguir.

O problema da transição diz respeito à trajetória da mudança, isto é, como as formas linguísticas mudam de um estágio para outro. Essa transição não ocorre de maneira brusca. De acordo com os autores, ela passa por três fases: o falante aprende uma forma nova; ela passa a coexistir com outra forma já inserida no sistema; a forma antiga se torna obsoleta e cai em desuso.

O problema do encaixamento procura analisar como a mudança estudada se encaixa na estrutura linguística e na estrutura social da comunidade de fala, uma vez que essas duas realidades são interdependentes para os estudos variacionistas.

O problema da avaliação é relativo à consciência linguística dos falantes, ou seja, como eles avaliam os usos linguísticos da comunidade à qual pertencem. Essa avaliação pode interferir no processo de mudança, uma vez que os falantes tendem a evitar as formas linguísticas estigmatizadas e supervalorizar as formas que gozam de prestígio no seio social.

Por último, o problema da implementação tem como objetivo descrever os fatores sociolinguísticos responsáveis pela mudança e por que essa mudança ocorre em determinada época e não em outra.

Esses cinco problemas brevemente expostos acima representam os obstáculos que o pesquisador sociolinguista deve cuidar de superar para que os fenômenos da variação e mudança linguísticas sejam empiricamente analisados.

Nossa pesquisa aborda apenas um desses problemas, que é o da restrição, ou seja, investigamos quais os fatores favorecem as variantes menos usadas (*haver* e *existir*) em nossa amostra e, conseqüentemente, restringem o uso da variante mais usada (*ter*).

Um dos princípios pelos quais a Teoria da Variação e Mudança Linguística se norteia é o fato de que toda mudança pressupõe variação, mas nem toda variação acarreta mudança (TARALLO, 2003, p. 63). Dito de outra forma, se um uso linguístico ganhou a batalha de outro uso referencialmente equivalente, isso implicará dizer que antes da

concretização da mudança, as duas formas coexistiam, ou seja, encontravam-se em variação. No entanto, é possível observar que nem todas as formas linguísticas que variam no sistema culminarão em mudança definitiva.

Há duas maneiras de observar se o fenômeno escolhido para o estudo representa uma variação estável ou uma mudança em curso: o estudo em tempo real e o estudo em tempo aparente.

O estudo em tempo real pode ser de longa duração ou de curta duração. Adotar o primeiro caso significa analisar a mudança de uma comunidade linguística no intervalo de um ou mais séculos. A maior dificuldade desse tipo de estudo é a viabilidade de coletar dados em períodos tão longos. Tarallo (2003) oferece algumas saídas para o pesquisador, dentre as quais estão: a consulta de atlas linguísticos, textos escritos em prosa que possam ser um retrato do vernáculo da comunidade estudada e as gramáticas da época, uma vez que elas contêm sempre comentários sobre a língua falada, fornecendo pistas ao pesquisar. Vale ressaltar aqui que os atuais projetos descritos na justificativa deste trabalho (NURC, NORPOFOR, etc.) serão fundamentais para a superação dessas dificuldades, uma vez que os pesquisadores do futuro se servirão dos dados desses projetos para verificar processos de mudança completos.

A segunda forma de estudar a mudança em tempo real diz respeito ao intervalo de curta duração (duas ou três décadas). Esse tipo de estudo pode ser de tendência ou de painel. O primeiro é quando, na volta ao campo para a coleta de dados, o pesquisador entrevista pessoas diferentes das que participaram do primeiro *corpus* e o segundo é quando as mesmas pessoas são entrevistadas novamente. Neste estudo, a maior dificuldade pode ser em encontrar os mesmos falantes, uma vez que muitos da faixa etária mais velha podem não estar mais vivos. Apesar desses obstáculos, os dados podem fornecer ao pesquisador resultados muito relevantes para a explicação da mudança.

O estudo em tempo aparente é aquele que leva em consideração o comportamento do fenômeno em estudo nas diferentes faixas etárias dos informantes. Os resultados apresentados pelas diferentes faixas etárias podem mostrar indícios de momentos passados da língua. No entanto, através desse estudo, não é possível afirmar categoricamente a existência da mudança observada, ficando o pesquisar apenas com a possibilidade de levantar hipóteses acerca do fenômeno. É o que faremos neste trabalho, uma vez que o *corpus* utilizado, o NORPOFOR, não apresenta dados de duas épocas diferentes, porque seu período de coleta não completou nem uma década ainda (agosto de 2003 a julho de 2006).

Em relação aos processos de mudança em curso ou variação estável, Martelotta (2011, p. 48-49) nos fornece exemplos bem esclarecedores desses dois fenômenos:

Os usos de *bicicleta/bicicreta*, *eles fizeram/eles fez*, *nós falamos/nós fala*, por exemplo, podem ser considerados casos de variação estável, pois as formas sem prestígio de cada par acima não costumam aparecer na fala de grupos com mais escolaridade. Usos como *a gente estuda/nós estudamos* e *falarei/vou falar* poderiam ser considerados (com base em pesquisas como as de Lopes, 2004 e Tesch, 2011), casos de mudança em curso. Os novos usos (*a gente e ir + infinitivo*) foram se espalhando e hoje são muito frequentes na fala de pessoas de todos os níveis de escolaridade. (MARTELOTTA, 2001, p. 48-49, grifos do autor)

Além desses dois termos específicos da Teoria da Variação e Mudança Linguística (tempo real e tempo aparente), existem outros que facilitam a propagação dos resultados dos trabalhos da área e, portanto, precisam ser explicados. O principal deles é o conceito de regra variável. Labov (2008[1972]) discute dois tipos de regras que podem ser observadas na heterogeneidade linguística. Portanto, há regras que são categóricas porque não sofrem mudança e apresentam a mesma regularidade na língua e há regras que são variáveis e são responsáveis pelos processos de mudança. É esse segundo tipo que interessa à Teoria da Variação e Mudança Linguística.

Outros dois conceitos importantes são: variável e variante. A respeito desses termos, Tarallo (2003, p. 8) nos fornece um conceito bem didático:

Em toda comunidade de fala são frequentes as formas linguísticas em variação. Como referimos, anteriormente, a essas formas em variação dá-se o nome de “variantes”. “Variantes linguísticas” são, portanto, diversas maneiras de se dizer a mesma coisa em um mesmo contexto, e com o mesmo valor de verdade. A um conjunto de variantes dá-se o nome de “variável linguística”. TARALLO (2003, p. 8, aspas do autor)

Para exemplificar essa definição, podemos usar o objeto de estudo desta pesquisa: a variação *haver*, *existir* e *ter*. Em contextos existenciais, esses três verbos passam a mesma informação referencial, ou seja, indicam a presença de algum novo elemento do discurso, estabelecendo sua existência. Portanto, *haver*, *existir* e *ter* são três variantes da mesma variável, que é o uso desses três verbos em contextos existenciais.

No que concerne à variável dependente, isto é, o fenômeno linguístico a ser estudado, é importante destacar que a definição de variável exposta por Tarallo (2003) é plenamente aceita quando se trata de variáveis fonético-fonológicas. Em relação às variáveis morfossintáticas, que é o nosso objeto de estudo, a teoria laboviana sofre uma grande crítica

de Lavandera (1978), que sugere a extensão do termo “significado representacional”, pois ele não levaria em conta a funcionalidade das variantes. Partindo desse ponto de vista, cada construção morfosintática apresenta um significado próprio, não sendo possível, portanto, falar-se em variação linguística nesse nível de análise.

Em resposta, Labov (1978) utiliza a expressão “estado de coisas” para reforçar a noção de significado referencial e sustenta que duas ou mais estruturas sintáticas que expressam o mesmo estado de coisas possuem o mesmo valor de verdade e, portanto, estão em variação dentro do sistema linguístico. O autor apresenta dois questionamentos que dizem respeito ao uso da língua no dia a dia e que reforçam sua definição:

Como identificamos que alguém fala como uma pessoa do campo a menos que saibamos que existem formas rurais e formas urbanas com o mesmo significado? Como identificamos que alguém falou conosco educadamente a menos que saibamos que ela escolheu uma das diversas formas de se dizer a mesma coisa, neste caso a variante mais atenuante?² (LABOV, 1978, p. 2, tradução nossa)

O autor admite que, obviamente, duas variantes linguísticas podem não ter o mesmo significado referencial em todos os contextos linguísticos, mas cabe ao pesquisador usar critérios rigorosos para sua análise. Em nossa pesquisa, por exemplo, não consideramos estruturas com os verbos *haver/ter* com sentido de *acontecer* porque, neste contexto, o verbo *existir* não expressa o mesmo estado de coisas, portanto, não se refere à mesma informação (cf. metodologia).

As variáveis podem ser ainda dependentes ou independentes. Essa terminologia utilizada nos estudos variacionistas, como nos esclarece Guy e Zilles (2007, p. 74), é oriunda da estatística e serve para estabelecer uma relação de dependência das formas linguísticas com os fatores linguísticos e sociais.

A variável dependente é o fenômeno em variação que o pesquisador deseja estudar, ou seja, duas ou mais formas linguísticas (variantes) que, no mesmo contexto, apresentam o mesmo valor de verdade (TARALLO, 2003, p. 8). No caso da presente pesquisa, a variável dependente é a variação *haver/existir/ter* existenciais, que podem se alternar, mantendo o mesmo significado referencial, ou seja, expressando a mesma informação.

²Texto original: How do we know that someone talks like a countryman unless we know that there are rural forms and urban forms with the same meaning? How do we know that someone has spoken politely to us, unless we know that he chose one of several ways of saying the same thing, in this case the more mitigating variant?

As variáveis independentes são os fatores linguísticos e sociais que condicionam a variável dependente. A correlação entre esses fatores, interpretada adequadamente pelo pesquisador, é que vai traçar todo o perfil da mudança linguística numa determinada comunidade linguística. Como veremos na metodologia, nossa pesquisa estuda 3 variáveis independentes sociais e 7 variáveis independentes linguísticas.

Definidos os principais termos com os quais opera a sociolinguística variacionista, é de suma relevância entendermos que não há teoria de sucesso que não sofra críticas por parte de outros estudiosos. Na verdade, dado o caráter não definitivo e incompleto do fazer científico, podemos afirmar que o fato de uma teoria ser criticada é, em toda a ciência, lugar comum, o que torna a busca pela compreensão do mundo que nos cerca algo muito cativante, emocionante e dinâmico, fazendo com que a busca do saber seja incessante.

No que concerne à teoria em questão, Bagno (2011) nos oferece uma breve crítica à Sociolinguística Variacionista. Segundo o autor, ela, embora muito importante na linguística contemporânea e merecedora de respeito por parte dos linguistas, carece de uma teoria social consistente. O autor afirma que o verdadeiro objetivo dessa corrente de estudos linguísticos é a descrição das línguas e, para a realização dessa descrição, utiliza a análise do contexto social somente como pretexto para explicar o fenômeno linguístico. Ele prossegue:

A própria divisão, tão característica dos estudos sociolinguísticos, entre *fatores internos (linguísticos, sistêmicos)* e *fatores externos (sociais)* como vetores da variação e da mudança, entre *variáveis dependentes* e *variáveis independentes*, revela uma dissociação entre entidades que deveriam ser vistas como uma coisa só – ou seja, na sociolinguística variacionista predomina uma concepção *dualista*, que separa o *social* do *linguístico*. No entanto, **o social é constitutivo do linguístico**, assim como **o linguístico é constitutivo do social** – não existem “influências” da sociedade sobre a língua, nem tampouco a língua “reflete” a sociedade: as relações entre esses dois termos, *língua* e *sociedade*, são muitíssimo mais amplas, mais íntimas e mais complexas do que uma mera influência ou um mero reflexo. (BAGNO, 2011, p. 62-63, grifos e aspas do autor)

De fato, essa observação de Bagno (2011) é muito pertinente e mostra que a sociolinguística variacionista, assim como toda teoria, possui limites de análise, que devem ser alcançados por outros ramos dos estudos da linguagem. No contexto dos estudos que levam em consideração o caráter social das línguas naturais, o autor cita a sociologia da linguagem que, embora usada como sinônimo da sociolinguística, tem outros objetivos em relação à linguagem. Como o próprio Bagno (2011) complementa, na sociologia da linguagem,

o que está em jogo é a língua como *instituição*, como instrumento de *controle social* de uma parcela da sociedade sobre as demais, como *arma* empregada para o exercício da *violência simbólica* de uns cidadãos contra os outros, como palco de *disputas de poder* e de *conflitos entre grupos sociais*, como elemento constitutivo da *identidade de indivíduos e de coletividades*, como *moeda* e *mercadoria* na economia das trocas linguísticas conforme delineada por Bourdieu... etc.” (BAGNO, 2011. p. 63 – grifos do autor)

Apesar de suas limitações, como qualquer outra ciência, é inegável a enorme contribuição que a Sociolinguística Variacionista tem dado aos estudos linguísticos, no que diz respeito à descrição das línguas naturais e, conseqüentemente, à construção de outras ideologias acerca dos usos linguísticos que servem de aporte para a formação de políticas linguísticas e a criação de materiais didáticos para a escola.

4 METODOLOGIA

4.1 TIPO DE PESQUISA

Esta é uma pesquisa de caráter quantitativo, pois analisa uma quantidade grande de dados de língua falada com o auxílio do programa de análise estatística **GoldVarb X**, que fornece análises multivariadas dos grupos de fatores em estudo. Para a realização desse estudo, utilizamos os pressupostos teórico-metodológicos da Teoria da Variação e Mudança Linguística de Weinreich, Labov e Herzog (1968) e de Labov (1963, 2006 [1966], 1969, 1978, 2008 [1972], 1994, 2006 [2001], 2010). Essa área de estudos, também conhecida por Sociolinguística Variacionista ou Sociolinguística Quantitativa, permite-nos analisar, através do estudo da língua em uso na comunidade linguística (ou comunidade de fala), as relações existentes entre as categorias sociais e as categorias linguísticas de um determinado fenômeno variável.

Alkmim (2008, p. 31, grifo da autora), afirma que o objeto de investigação da Sociolinguística Variacionista é a análise da língua falada, estudada em situações reais de uso e acrescenta que o ponto de partida desses estudos é a *comunidade linguística*, que ela também denomina de comunidade de fala.

Labov (2008 [1972], p. 150), a respeito da comunidade de fala, faz a seguinte observação:

A comunidade de fala não é definida por nenhuma concordância marcada no uso de elementos linguísticos, mas sim pela participação num conjunto de normas compartilhadas; estas normas podem ser observadas em tipos de comportamento avaliativo explícito e pela uniformidade de padrões abstratos de variação que são invariantes no tocante a níveis particulares de uso. (LABOV, 2008 [1972], p. 150)

Essa “uniformidade de padrões abstratos de variação” já foi constatada na comunidade de fala fortalezense (RODRIGUES, 2013; GUIMARÃES, 2014), uma vez que seus falantes compartilham um conjunto de normas bem definidas e que seguem padrões linguísticos de variação e/ou mudança linguística característicos da comunidade.

4.2 COMUNIDADE DE FALA: A CIDADE DE FORTALEZA

Nesta seção, esboçaremos um retrato geográfico da cidade de Fortaleza na década de 2000, época em que foram gravadas as entrevistas que compõem nossa amostra.

Quando as primeiras entrevistas do NORPOFOR começaram a ser realizadas em agosto de 2003, Fortaleza já havia completado 277 anos e 4 meses de existência, desde que foi elevada à condição de vila, em 13 de abril de 1729. Devido ao seu crescimento econômico e político, ela foi oficializada como cidade em 1823, ou seja, 94 anos depois (BRUNO; FARIAS, 2011, p. 1). Portanto, na década de 2000, Fortaleza tinha uma história de formação linguística de quase 3 séculos, incluindo influências indígenas, portuguesas e holandesas.

Em 1997, foram criadas 7 Secretarias Executivas Regionais com o objetivo de descentralizar a administração da cidade. Na década de 2000, essas secretarias abrangiam 114 bairros e Fortaleza tinha uma população de 2.141.402 habitantes de acordo com o senso do IBGE de 2000. Atualmente, a cidade possui 119 bairros divididos da seguinte forma, de acordo com o site da prefeitura³:

Quadro 1 - Distribuição dos bairros de Fortaleza por Secretarias Executivas Regionais (SERs)

SERs	BAIRROS
SER I 15 bairros	A Vila Velha, Jardim Guanabara, Jardim Iracema, Barra do Ceará, Floresta, Álvaro Weyne, Cristo Redentor, Ellery, São Gerardo, Monte Castelo, Carlito Pamplona, Pirambu, Farias Brito, Jacarecanga e Moura Brasil
SER II 20 bairros	Aldeota, Cais do Porto, Cidade 2000, Cocó, De Lourdes, Dionísio Torres, Engenheiro Luciano Calvalcante, Guararapes, Joaquim Távora, Manuel Dias Branco, Meireles, Mucuripe, Papicu, Praia de Iracema, Praia do Futuro I e II, Salinas, São João do Tauape, Varjota, Vicente Pinzon
SER III 17 bairros	Amadeu Furtado, Antônio Bezerra, Autran Nunes, Bonsucesso, Bela Vista, Dom Lustosa, Henrique Jorge, João XXIII, Jóquei Clube, Olavo Oliveira, Padre Andrade, Parque Araxá, Pici, Parquelândia, Presidente Kennedy, Rodolfo Teófilo e Quintino Cunha
SER IV 19 bairros	São José Bonifácio, Benfica, Fátima, Jardim América, Damas, Parreão, Bom Futuro, Vila União, Montese, Couto Fernandes, Pan Americano, Demócrito Rocha, Itaoca, Parangaba, Serrinha, Aeroporto, Itaperi, Dendê e Vila Pery
SER V 18 bairros	Conjunto Ceará, Siqueira, Mondubim, Conjunto José Walter, Granja Lisboa, Granja Portugal, Bom Jardim, Genibaú, Canindezinho, Vila Manoel Sátiro, Parque São José, Parque Santa Rosa, Maraponga, Jardim Cearense,

³<http://www.fortaleza.ce.gov.br/regionais>

	Conjunto Esperança, Presidente Vargas, Planalto Ayrton Senna e Novo Mondubim
SER VI 29 bairros	Aerolândia, Ancuri, Alto da Balança, Barroso, Boa Vista (unificação do Castelão com Mata Galinha), Cambeba, Cajazeiras, Cidade dos Funcionários, Coaçu, Conjunto Palmeiras (parte do Jangurussu), Curió, Dias Macedo, Edson Queiroz, Guajerú, Jangurussu, Jardim das Oliveiras, José de Alencar (antigo Alagadiço Novo), Messejana, Parque Dois Irmãos, Passaré, Paupina, Parque Manibura, Parque Iracema, Parque Santa Maria (parte do Ancuri), Pedras, Lagoa Redonda, Sabiaguaba, São Bento (parte do Paupina) e Sapiranga
Sercefor 1 bairro	Centro

O senso de 2010 mostrava Fortaleza com uma população de 2.452.185 e uma estimativa de 2.571.896 pessoas para 2014. Portanto, da época da construção do NORPOFOR até os dias de hoje, o crescimento populacional foi de 20%. É relevante levarmos isso em consideração para entendermos que os resultados desta pesquisa serão válidos para a Fortaleza da década de 2000 e não para a Fortaleza de hoje.

4.3 O *CORPUS* E A AMOSTRA

Os trabalhos sociolinguísticos da vertente variacionista têm como principal objetivo analisar o vernáculo, definido por Labov (2006, p. 86 – tradução nossa) como “a língua primeira adquirida pelo falante, controlada perfeitamente, e usada primordialmente entre amigos íntimos e membros da família⁴”. Essa língua pode ser observada na fala mais espontânea, ou seja, quando o falante não está preocupado em monitorar seu uso linguístico. No entanto, segundo Labov (2008[1972], p. 244), o único meio pelo qual se pode conseguir bons dados do vernáculo pretendido pelo pesquisador é por meio da entrevista individual, gravada. Dito isso, uma pergunta surge imediatamente: como podemos colher a fala espontânea de um informante em uma situação não espontânea de comunicação, que é a gravação de uma entrevista diante de uma pessoa desconhecida? Esse impasse foi denominado pelo próprio Labov como o paradoxo do observador e logo ele sugeriu algumas soluções:

⁴ Texto original: “the language first acquired by the language learner, controlled perfectly, and used primarily among intimate friends and family members.”

Uma maneira de superar o paradoxo é romper os constrangimentos da situação de entrevista com vários procedimentos que desviem a atenção do falante e permitam que o vernáculo emerja. [...]. Também podemos envolver a pessoa com perguntas e assuntos que recriem emoções fortes que ela experimentou no passado, ou envolvê-la em outros contextos. Uma das perguntas desse tipo que tem dado mais resultado é a que lida com o “risco de vida”: “Você já viveu uma situação em que correu sério risco de morrer?”. As narrativas produzidas em resposta a essa pergunta quase sempre exibem uma mudança de estilo que se distancia da fala monitorada e se aproxima do vernáculo. (LABOV (2008[1972], p. 244-245, aspas do autor)

Neste trabalho, utilizamos o *corpus* do projeto NORPOFOR (Norma Oral do Português Popular de Fortaleza)⁵ por ter sido constituído com muita seriedade e competência, seguindo os critérios rigorosos da Sociolinguística Variacionista no que diz respeito à escolha dos informantes e às entrevistas. Além disso, esse projeto se encaixa perfeitamente em nosso objetivo, que é estudar o falar popular. O NORPOFOR foi concretizado sob a coordenação da profa. Dra. Aluiza Alves de Araújo da Universidade Estadual do Ceará e contou com a colaboração de 4 bolsistas do programa de iniciação científica da instituição (Girleene Moreira da Silva – voluntária; Neyla Denize de Souza – voluntária; Francisco Cláudio Rodrigues – FUNCAP e Maria de Fátima do Nascimento Silva – IC-UECE) além de, na época, o professor substituto da UECE Kilpatrick Müller Campelo. Os alunos do Curso de Letras da UECE também tiveram participação voluntária e fundamental na realização das entrevistas e transcrição dos inquéritos, conforme Araújo (2011).

Com o objetivo de neutralizar a influência de falares de outras regiões do país, os informantes foram escolhidos, em obediência aos seguintes critérios: 1) são fortalezenses natos ou vieram morar nesta cidade com, no máximo, cinco anos de idade; 2) possuem pais cearenses; 3) nunca se ausentaram de Fortaleza por um período superior a dois anos consecutivos; e 4) mantêm residência fixa na capital cearense. (ARAÚJO, 2011. p. 838).

O NORPOFOR conta com 198 informantes, que foram estratificados por gênero, faixa etária (I: 15 a 25 anos; II: 26 a 49 anos; III: a partir de 50 anos), escolaridade (A: 0 a 4 anos; B: 5 a 8 anos; C: 9 a 11 anos) e tipo de registro (DID: Diálogo entre Informante e Documentador; D2: Diálogo entre Dois Informantes; EF: Elocução Formal).

⁵Araújo (2007) utiliza o termo norma popular na acepção entendida por Bagno (2003, p. 59): “variedades linguísticas relacionadas a falantes sem escolaridade superior completa, com pouca ou nenhuma escolarização”

Quadro 2 - Distribuição dos informantes do NORPOFOR por gênero, idade, tipo de registro e escolaridade

NORPOFOR	Gênero																	
	Masculino									Feminino								
Registro	DID			D2			EF			DID			D2			EF		
Esc. → F. E. ↓	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
I	5	4	5	4	4	5	2	2	4	4	5	4	2	5	5	0	0	2
II	4	5	5	4	4	4	4	3	4	5	5	5	4	5	5	0	5	4
III	5	6	5	3	3	4	3	2	1	4	5	4	4	5	4	1	1	1

Fonte: Araújo (2011, p. 839)

Para a realização desta pesquisa, excluímos os tipos de registro EF e DID. O primeiro tipo foi excluído por se tratar do uso da língua num contexto formal e, apesar dos falantes não possuírem alto nível de escolarização, acreditamos que, nesse tipo de contexto, eles monitoram seus usos linguísticos, claro, dentro dos limites de sua língua vernacular. O segundo tipo de registro serviria aos nossos propósitos (estudar o falar popular de Fortalezaense), pois os cuidados tomados na entrevista (e.g. contato prévio com o entrevistado, entrevista realizada de maneira informal, omissão do real objetivo da entrevista, temas muito pessoais, etc.) minimizaram esse efeito de monitoramento. Porém, esse tipo de inquérito foi excluído do trabalho por questões de tempo e, dessa forma, optamos pela análise dos inquéritos D2, visto que são os mais espontâneos possíveis, portanto, os que mais se aproximam do vernáculo. Nesse tipo de inquérito, os informantes, que eram amigos ou familiares, escolhiam o tema da conversa, o dia, o local e o horário da gravação que, na maioria das vezes, não tinha participação do documentador, reduzindo muito os efeitos inibidores do gravador.

Para definir nossa amostra, escolhemos 3 informantes de cada célula, com exceção da célula F-D2-A-I do *corpus* (informantes femininos, em inquéritos de dois informantes, de faixa etária entre 15 a 25 anos e de grau de escolaridade I), porque no *corpus* completo, essa célula só apresenta 2 informantes. Isso não representa um problema uma vez que o GoldVarb X, como destaca Guy e Zilles (2007, p. 106), tolera esse tipo de desvio nas células.:

Como vemos, nossa amostra é bem equilibrada no que diz respeito à distribuição dos informantes por célula. Os 53 informantes nos forneceram uma quantidade significativa

de dados (931) em comparação com o estudo mais representativo que encontramos sobre o tema, que é a tese de Vitório (2012). Essa autora trabalha com 72 informantes, mas só encontrou 772 ocorrências existenciais. Se retirarmos os 28 dados do verbo *existir* da nossa amostra (verbo que não foi estudado pela autora) ainda temos uma quantidade superior a dela, com 903 dados.

Como mostra o quadro 3, nossa amostra resultou num total de 29 inquéritos do tipo D2 e 53 informantes, distribuídos por sexo, faixa etária e escolaridade. Não há simetria entre os números de inquéritos D2 e o número de informantes porque, em alguns inquéritos, selecionamos apenas 1 informante devido a pouca quantidade de dados. O quadro 3 mostra como ficou dividida nossa amostra:

Quadro 3 - Distribuição dos informantes por sexo, escolaridade e faixa etária em nossa amostra

29 inquéritos	SEXO					
53 informantes	MASCULINO			FEMININO		
Escolaridade → Faixa etária ↓	0 a 4 anos	5 a 8 anos	9 a 11 anos	0 a 4 anos	5 a 8 anos	9 a 11 anos
15 a 25 anos	3	3	3	2	3	3
26 a 49 anos	3	3	3	3	3	3
A partir de 50 anos	3	3	3	3	3	3

Essa distribuição de falantes também nos forneceu uma distribuição equilibrada de ocorrências, como vemos no quadro 4:

Quadro 4 - Distribuição das ocorrências por sexo, escolaridade e faixa etária em nossa amostra

29 inquéritos	SEXO					
53 informantes	MASCULINO (474)			FEMININO (457)		
Escolaridade → Faixa etária ↓	0 a 4 anos (180)	5 a 8 anos (104)	9 a 11 anos (190)	0 a 4 anos (178)	5 a 8 anos (125)	9 a 11 anos (154)

15 a 25 anos (209)	45	24	36	8	60	36
26 a 49 anos (359)	103	41	81	42	35	57
A partir de 50 anos (363)	32	39	73	128	30	61

4.4 AS ENTREVISTAS D2 DO NORPOFOR

As entrevistas do NORPOFOR foram realizadas no período de agosto de 2003 a julho de 2006

com o objetivo de constituir um banco de dados sobre o falar popular dos fortalezenses. Face à inexistência de um corpus que fosse suficientemente representativo dessa variedade do ponto de vista quantitativo e que controlasse as variáveis gênero, faixa etária, escolaridade e tipo de registro, surgiu a motivação para elaboração do NORPOFOR. (ARAÚJO, 2011, p. 836)

Ainda de acordo com Araújo (2011, p. 836), as gravações dos inquéritos tipo D2 foram feitas com base no relato de experiência pessoal, abordando temas como família, lazer, infância, entre outros. Esses temas eram escolhidos pelos próprios informantes no momento da gravação e após uma conversa prévia entre informantes e documentador (quando este estava presente), para a criação de um laço menos estranho. Ao todo, foram 92 documentadores.

Deste total, 03 eram professores (Kilpatrick Müller Campelo, Aluiza Alves de Araújo e Socorro Abreu), 19 eram universitários e 70 eram universitárias, todos vinculados ao Curso de Graduação em Letras da UECE. Os documentadores recebiam as mesmas orientações sobre como realizar a gravação dos inquéritos. Cada um tinha a responsabilidade de estimular o(s) informante(s) a falar e controlar os fatores que pudessem prejudicar a qualidade da gravação (intervenção de terceiros, ruídos externos, manuseio do gravador e fitas cassete, entre outros). (ARAÚJO, 2011, p. 841)

Vale ressaltar que vários inquéritos D2 não possuíam a presença do documentador. Um gravador era entregue aos dois informantes, que recebiam instruções básicas do processo de gravação, e eles decidiam o dia, a hora, o local e o tema sobre o qual eles iriam falar. Esse detalhe torna esse tipo de inquérito ainda mais espontâneo e vernacular do que os outros dois (DID e EF).

4.5 OS INFORMANTES DE NOSSA AMOSTRA

Nossa amostra é composta de 29 inquéritos do tipo D2, que representam 53 informantes, cujos dados podem ser visualizados no quadro 5, que elaboramos:

Quadro 5 - Distribuição dos informantes da amostra por nº do inquérito/informante, sexo, idade, escolaridade, bairro, Secretaria Executiva Regional (SER) e atividade exercida

Nº	Inq.	Inf.	Sexo	Idade	Escol.	Bairro	SER	Atividade exercida
1	04	Inf. 1	Masculino	44	9-11	Messejana	VI	Bancário
2		Inf. 2	Feminino	42	9-11	Messejana	VI	Prendas do lar
3	14	Inf. 1	Masculino	38	9-11	Mondubim	V	Militar
4		Inf. 2	Masculino	35	9-11	Maraponga	V	Oficial de Justiça
5	15	Inf. 1	Masculino	25	9-11	Vila União	IV	Comerciante
6		Inf. 2	Masculino	24	9-11	Fátima	IV	Analista de suporte
7	35	Inf. 1	Masculino	21	5-8	Messejana	IV	Mecânico
8	49	Inf. 1	Feminino	41	9-11	Henrique Jorge	III	Costureira
9		Inf. 2	Feminino	40	9-11	Henrique Jorge	III	Desempregada
10	50	Inf. 1	Feminino	17	9-11	Parangaba	IV	Estudante
11		Inf. 2	Feminino	17	9-11	Parangaba	IV	Estudante
12	51	Inf. 1	Feminino	15	9-11	Parangaba	IV	Estudante
13		Inf. 2	Masculino	15	9-11	Parangaba	IV	Estudante
14	52	Inf. 1	Masculino	34	5-8	Parque São José	V	Vigilante
15		Inf. 2	Masculino	35	0-4	Planalto do Pici	III	Vigilante
16	60	Inf. 1	Feminino	51	5-8	Demócrito Rocha	IV	Prendas do lar
17		Inf. 2	Feminino	56	5-8	Demócrito Rocha	IV	Vendedora ambulante
18	71	Inf. 1	Masculino	51	9-11	Antônio Bezerra	III	Funcionário público
19	93	Inf. 1	Feminino	59	0-4	Álvaro Weyne	I	Aposentada
20		Inf. 2	Feminino	63	0-4	Álvaro Weyne	I	Lavadeira
21	94	Inf. 1	Masculino	34	0-4	Barra do Ceará	I	Pedreiro
22		Inf. 2	Masculino	31	0-4	Barra do Ceará	I	Auxiliar de pedreiro
23	99	Inf. 1	Feminino	42	5-8	Tancredo Neves	VI	Prendas do lar
24	101	Inf. 2	Feminino	26	5-8	Bonsucesso	III	Prendas do lar

25	106	Inf. 1	Feminino	76	0-4	Messejana	VI	Aposentada
26		Inf. 2	Masculino	76	0-4	Messejana	VI	Aposentado
27	108	Inf. 1	Masculino	46	5-8	Quintino Cunha	III	Serviços gerais
28		Inf. 2	Masculino	21	0-4	Barra do Ceará	I	Estudante
29	111	Inf. 1	Masculino	58	9-11	Carlito Pamplona	I	Comerciante
30		Inf. 2	Masculino	70	9-11	Álvaro Weyne	I	Aposentado
31	114	Inf. 1	Feminino	20	5-8	Serrinha	IV	Estudante
32		Inf. 2	Feminino	38	0-4	Serrinha	IV	Prendas do lar
33	118	Inf. 1	Feminino	17	0-4	Monte Castelo	I	Estudante
34		Inf. 2	Feminino	26	5-8	Bela Vista	III	Estudante
35	129	Inf. 1	Feminino	60	9-11	Parquelândia	III	Aposentada
36		Inf. 2	Feminino	49	0-4	Parquelândia	III	Merendeira escolar
37	132	Inf. 1	Masculino	52	0-4	Farias Brito	I	Operário
38		Inf. 2	Masculino	60	5-8	Farias Brito	I	Pintor aposentado
39	139	Inf. 1	Feminino	28	0-4	Alto Alegre	IV	Doméstica
40		Inf. 2	Feminino	19	0-4	Alto Alegre	IV	Doméstica
41	141	Inf. 1	Masculino	15	5-8	Parque Santa Rosa	V	Estudante
42		Inf. 2	Masculino	16	5-8	Conjunto Esperança	V	Estudante
43	152	Inf. 1	Feminino	55	5-8	Antônio Bezerra	III	Prendas do lar
44		Inf. 2	Masculino	51	0-4	Antônio Bezerra	III	Comerciante
45	153	Inf. 1	Masculino	24	0-4	Parque São José	V	Balconista
46		Inf. 2	Masculino	18	0-4	Bom Jardim	V	Balconista
47	154	Inf. 1	Feminino	18	5-8	José Walter	V	Estudante
48		Inf. 2	Masculino	53	5-8	José Walter	V	Aposentado
49	155	Inf. 1	Masculino	55	5-8	Álvaro Weyne	I	Motorista
50		Inf. 2	Feminino	15	5-8	Pan-americano	IV	Estudante
51	157	Inf. 1	Feminino	51	9-11	Álvaro Weyne	I	Cozinheira
52		Inf. 2	Feminino	52	9-11	Praia de Iracema	II	Gerente administrativa
53	160	Inf. 1	Masculino	46	5-8	Quintino Cunha	III	Auxiliar de serviços gerais

Como vemos neste quadro, nossos informantes representam 29 bairros de Fortaleza e abrangem 6 das 7 Secretarias Executivas Regionais. Os bairros mais representados são Messejana (SER VI) e Álvaro Weyne (SER I), com 5 informantes cada. Os menos

representados têm apenas 1 informante. São 15 bairros: Carlito Pamplona e Monte Castelo (SER I); Praia de Iracema (SER II); Planalto do Pici, Bela Vista e Bonsucesso (SER III); Fátima, Vila União e Pan-americano (SER IV); Maraponga, Bom jardim, Mondubim, Conjunto Esperança e Parque Santa Rosa (SER V); Tancredo Neves (SER VI). Os outros 12 bairros são representados por 2, 3 ou 4 informantes. (ver APÊNDICE B).

Em relação à atividade exercida, nossa amostra apresenta 27. As três mais representativas são: estudante, com 12 informantes; aposentado(a) e prendas do lar, com 6 informantes cada. As demais são representadas por 1, 2 ou 3 informantes.

4.6 LEVANTAMENTO DOS DADOS

Após definido o tipo de inquérito a ser utilizado nesta pesquisa, as gravações do tipo D2, a seleção da amostra foi feita com base nas células existentes no NORPOFOR e a partir de uma busca prévia das variantes em estudo para identificar quais seriam os melhores inquéritos a serem escolhidos. Depois dessa etapa, passamos ao levantamento de todos os dados, que foi feito utilizando o recurso “Localizar” do *Word*, buscando todas as raízes possíveis de cada verbo. Ao passo que íamos identificando as ocorrências, também íamos ouvindo cada inquérito para checar e/ou tirar eventuais dúvidas que apareceram durante o processo. Vale lembrar que foram utilizados 29 inquéritos D2 e, desses, apenas 6 possuem entre 28 e 29 minutos de gravação. Todos os demais têm entre 30 minutos e 1 hora de material gravado, perfazendo um total de 20 horas e 30 minutos de dados de fala analisados.

Ao identificar cada ocorrência, salvávamo-la em arquivo *txt* para que o **GoldVarb X** pudesse fazer a leitura dos dados. Uma vez identificadas todas as ocorrências, foi possível decidir o sistema analítico que iríamos utilizar, isto é, todas as variáveis independentes que possivelmente eram relevantes para o fenômeno em estudo.

4.7 VARIÁVEIS

4.7.1 Variável dependente

Nesta pesquisa, estudamos o fenômeno variável *haver, existir e ter* na acepção de existência, levando em consideração a definição de significado sintático de Labov (1978, p. 7 – tradução nossa): “duas sentenças que se referem ao mesmo estado de coisas têm o mesmo

valor de verdade⁶. Acreditamos que se substituirmos cada verbo utilizado nas ocorrências (1), (2) e (3) pelos outros dois do nosso sistema analítico, teríamos a mesma informação (“estado de coisas”) sendo veiculada.

- (1) é porque naquele tempo HAVIA respeito pelos pais (NORPOFOR, D2, 60)
- (2) EXISTE alguns ambiente que são mais fácil de se encontrar certo?
(NORPOFOR, D2, 04)
- (3) TEM gente que vota né só mesmo por votar... (NORPOFOR, D2, 118)

No entanto, é importante selecionar somente aquelas ocorrências que apresentam uma variante cuja substituição por outra (ou em nosso caso, pelas outras duas, já que trabalhamos com uma variável ternária), mantenha o mesmo sentido sintático exposto por Labov (1978).

Por esta razão, excluimos de nossa análise as sentenças de *haver* e *ter* no sentido de *ocorrer*, *acontecer*, pois a substituição por *existir* em tais contextos não é possível, ou porque isso acarretará em outro significado ou porque a sentença não parecerá natural.

- (4) ei hoje TEM culto não né?... é a mudança né? (NORPOFOR, D2, 50)
- (5) o que que HOUVE com ele? vocês brigaram... alguma coisa assim?
(NORPOFOR, D2, 157)

A inserção de *existir* em (4) dá outro sentido ao enunciado, isto é, denota *existência* mesmo e não *ocorrer*. Já em (5), essa inserção não é possível porque torna o enunciado não natural, não vernacular. É relevante dizer que não obtivemos nenhuma ocorrência de *existir* no sentido de *ocorrer*, o que força nossa escolha metodológica.

Nós excluimos também de nossa análise a única ocorrência de *ter pessoal*, pois em (6) a substituição de *ter* por *haver* ou *existir* não é possível.

- (6) hoje o pobre tá abandonado em todos os sentidos... na saúde mesmo você não TEM mais nada... você não TEM um hospital público (NORPOFOR, D2, 111)

⁶ Texto original: [...] two utterances that refer to the same state of affairs have the same truth-value [...]

4.7.2 Variáveis independentes

São 10 variáveis independentes analisadas neste estudo, das quais 3 são sociais e 7 são linguísticas. Elas foram escolhidas levando em consideração três critérios. O primeiro deles diz respeito às 3 variáveis sociais, que foram selecionadas com base na divisão do NORPOFOR (**sexo**, **faixa etária** e **escolaridade**); a variável tipo de registro não foi investigada porque só trabalhamos com os inquéritos do tipo D2, pelas razões apresentadas na seção 3.3. O segundo e o terceiro critérios foram utilizados para a escolha de todas as variáveis linguísticas e são, respectivamente, a literatura existente acerca do tema e a análise mais detalhada das ocorrências. Com base na literatura (SILVA, 2001; MARTINS; CALLOU, 2003; VITÓRIO, 2012, 2013; BATISTA, 2012; SANTOS; RAMOS, 2012), decidimos analisar as 6 variáveis linguísticas seguintes: **animacidade do Sintagma Nominal** (doravante, SN), **forma verbal**, **peso do SN**, **posição do SN em relação ao verbo**, **concordância entre o verbo e o SN**, **presença x ausência de elementos à esquerda do verbo**. Com base nos dados de nossa amostra, decidimos incluir mais uma variável linguística: **repetição do verbo no mesmo enunciado**. Para cada uma desses grupos de fatores, existe uma hipótese, criada a partir dos resultados encontrados em outros trabalhos.

4.7.2.1 Variáveis sociais

As variáveis sociais desempenham um papel importantíssimo nos estudos sociolinguísticos variacionistas, pois revelam como a língua varia e/muda de acordo com as características sociais dos falantes. A seguir, abordaremos cada uma das 3 variáveis sociais analisadas nesta pesquisa:

- SEXO
- FAIXA ETÁRIA
- ESCOLARIDADE

4.7.2.1.1 *Sexo*

Essa variável é muito testada nos estudos variacionistas devido ao princípio geral de que homens falam diferentemente do que as mulheres. No entanto, existe uma grande discussão sobre o melhor termo com o qual devemos trabalhar para analisar essa categoria, se

sexo ou gênero. Eckert e McConnell-Ginet (2003, p. 10 – tradução nossa) afirmam que “o sexo é uma categorização baseada primordialmente no potencial reprodutivo, enquanto que o gênero é a elaboração social do sexo biológico⁷”. Em outras palavras, nascemos homens ou mulheres, mas com o passar do tempo formamos nosso papel social com base em nosso sexo biológico, que pode ser completamente diferente dessa divisão tradicional. A sociedade hoje apresenta homens e mulheres, em vários setores da escala social, fazendo tarefas que antes eram primordialmente executadas por um sexo apenas.

Acreditamos que essa discussão é pertinente e muito frutífera uma vez que pode mudar o foco da constituição dos principais *corpora* do Brasil (e.g. NURC, VARSUL, PORCUFORT, NORPOFOR). No entanto, atualmente, todos esses corpora, apesar de incluírem, nas fichas dos informantes, informações acerca dos papéis sociais (e.g. atividade exercida) desempenhados por eles, são estratificados pela separação biológica entre homens e mulheres. Por esta razão, o presente trabalho aborda a variável **sexo** como uma categoria essencialmente biológica e investiga como os homens e as mulheres se comportam diante da variação *haver/existir/ter* de nossa amostra. Portanto, a variável **sexo** está dividida, de acordo com o *corpus* estudado, em:

- Masculino
- Feminino

4.7.2.1.2 Faixa etária

A variável **faixa etária** também é bastante controlada nos estudos sociolinguísticos. Por meio da análise dessa variável, além de entender como as variantes se comportam nas diferentes divisões etárias, é possível também inferir, através dos indícios revelados pelos dados, o *status* do fenômeno variável em estudo: variação estável ou mudança em progresso. Essa análise dos índices apresentados nas diferentes faixas etárias é conhecida nos estudos de Sociolinguística Variacionista como **estudo do tempo aparente** (ver seção 5.6 Análise do tempo aparente, p. 88).

As faixas etárias controladas aqui são:

- 15 a 25 anos (jovens)
- 26 a 49 anos (adultos)

⁷Texto original: Sex is a biological categorization based primarily on reproductive potential, whereas gender is the social elaboration of biological sex.

- A partir de 50 anos (idosos)

4.7.2.1.3 *Escolaridade*

A observação do dia a dia confirma que a escola gera mudanças na fala e na escrita das pessoas que as frequentam e das comunidades discursivas. Constata-se, por outro lado, que ela atua como preservadora de formas de prestígio, face a tendência de mudança em curso nessas comunidades. Veículo de familiarização com a literatura nacional, a escola incute gostos, normas, padrões estéticos e morais em face da conformidade de dizer e de escrever. (VOTRE, 2012, p. 51)

Diante das palavras acima, entendemos que o papel da escola no funcionamento de uma língua é extremamente decisivo. Por isso, a variável **escolaridade** é uma das mais importantes variáveis sociais analisadas nos estudos variacionistas.

No Brasil, a variante inovadora *ter* existencial é muito estigmatizada na escrita. No entanto, esse estigma não se aplica à fala, visto que ela já está presente na variedade falada das pessoas mais letradas (MARTINS; CALLOU, 2003; BATISTA, 2012). Esses estudos utilizam dados do projeto NURC, que representa a língua falada por pessoas que possuem nível superior completo e, mesmo assim, o verbo *ter* tem a maior frequência de uso em todas as variáveis testadas. As escolaridades controladas no NORPOFOR e, conseqüentemente, analisadas nesta pesquisa são:

- 0 a 4 anos
- 5 a 8 anos
- 9 a 11 anos

4.7.2.2 Variáveis linguísticas

Vale ressaltar que é exatamente o estudo das variáveis sociais e sua correlação com as variáveis linguísticas que caracteriza a Sociolinguística Variacionista, pois Labov (2008 [1972]) acredita que a linguística deve ser necessariamente uma ciência social. Torna-se evidente, portanto, que “As variáveis, tanto linguísticas quanto não linguísticas, não agem isoladamente, mas operam num conjunto complexo de correlações que inibem ou favorecem o emprego de formas variantes semanticamente equivalentes.” (MOLLICA, 2012, p. 27)

As 7 variáveis linguísticas deste trabalho foram escolhidas levando em consideração todas as pesquisas expostas até aqui, independente de sua relevância, bem como a observação minuciosa dos dados. São elas:

- FORMA VERBAL
- ANIMACIDADE DO SN
- PESO DO SN
- POSIÇÃO DO SN EM RELAÇÃO AO VERBO
- CONCORDÂNCIA ENTRE O VERBO E O SN
- PRESENÇA X AUSÊNCIA DE ELEMENTOS À ESQUERDA DO VERBO
- REPETIÇÃO DO VERBO NO MESMO ENUNCIADO

4.7.2.2.1 *Forma verbal*

Na literatura sobre o tema, a denominação dessa variável é sempre **tempo verbal**, no entanto, esse termo é analisado com significados diferentes, sem muito detalhamento nos procedimentos adotados para tal designação. Dos estudos encontrados sobre o tema, Vitória (2012, p. 77, grifos da autora) é a única que justifica sua escolha, optando por analisar “o tempo expresso pela forma do verbo” e, assim, a autora distribui as formas verbais nos tempos “passado e presente”. Batista (2012), por exemplo, define essa variável de **tempo verbal**, mas estuda as formas verbais, uma vez que inclui o infinitivo em sua análise.

Optamos por denominar essa variável **forma verbal** porque encontramos ocorrências das formas nominais **infinitivo** e **gerúndio** de forma que nem sempre ficava claro o tempo verbal que essas formas representam. Encontramos ocorrências de 9 formas verbais, que estão listadas abaixo:

- Presente do indicativo
- Presente do subjuntivo
- Pretérito imperfeito do indicativo
- Pretérito imperfeito do subjuntivo
- Pretérito perfeito do indicativo
- Futuro do pretérito do indicativo
- Futuro do subjuntivo
- Infinitivo
- Gerúndio

4.7.2.2.2 *Animacidade do SN*

A variável **animacidade do SN** foi subdividida nos traços [+humano] e [-humano]. Não há uma homogeneidade na literatura acerca da investigação das propriedades semânticas do SN que complementa o verbo existencial. Martins e Callou (2003) estudam a variável **especificidade semântica do argumento interno** e a divide em quatro: material (concreto), humano, abstrato e evento. Silva (2001) analisa a variável **animacidade do SN** e a subdivide nos traços [+animado] e [-animado]. Já Vitório (2012) investiga duas variáveis em relação à semântica do complemento dos verbos existenciais, a saber, **a natureza do argumento interno** (concreto/abstrato) e a **animacidade do argumento interno** ([+animado] e [-animado]).

Optamos por não estudar as denominações “animado, inanimado, concreto e abstrato” por considerarmos conceitos muito filosóficos, o que torna a classificação do SN muito difícil. A denominação “evento” também não foi analisada por nós porque as ocorrências de *haver* e *ter* no sentido de *acontecer* foram excluídas dos nossos dados.

Portanto, optamos por analisar sob o termo **Animacidade do SN** os seguintes fatores, que são mais palpáveis e, portanto, de fácil classificação:

- [+humano]

(7) porque EXISTE pessoas que não se conforma em você não querer ela
(NORPOFOR, D2, 04)

- [-humano]

(8) porque TEM as escolas diferenciadas... zero beneficio pra eles...
(NORPOFOR, D2, 71)

Esse termo SN (sintagma nominal) ao invés de argumento interno foi escolhido meramente por uma opção terminológica (PERINI;1996, 2008, 2010; BAGNO, 2011; CASTILHO, 2012), no entanto, trata-se do mesmo elemento da sentença estudado na literatura: o complemento dos verbos *haver*, *existir* e *ter* na acepção de existência.

Como vemos até o momento, cada estudioso(a) divide as variáveis a serem estudadas, principalmente as de caráter semântico-discursivo, da forma que mais lhe agrada

de modo que a comparação entre os resultados oriundos de tais estudos devem levar isso em consideração.

4.7.2.2.3 *Peso do SN*

Essa variável foi controlada por Silva (2001) e decidimos estudá-la também, seguindo a mesma divisão da autora, que utilizou como base a abordagem de Perini (1996) sobre o SN:

- SN simples

(9) mas... HOUVE vezes em que eu... eu observei que... os cinco conversando:...
(NORPOFOR, D2, 157)

- SN complexo

(10) TEM muitas mulheres que vão pela novela... (NORPOFOR, D2, 60)

- SN nulo

(11) não... TEM vários... diversos... [referente: *cantor*] (NORPOFOR, D2, 50)

- Núcleo do SN nulo

(12) antes num TINHA... [referente: *professores*] (NORPOFOR, D2, 04)

4.7.2.2.4 *Posição do SN em relação ao verbo*

Essa variável também foi testada em Silva (2001) e dividida nos seguintes fatores: **SN à direita do verbo**, que é a estrutura prototípica, e **SN à esquerda do verbo**, que é a inversão de posições. Nessa segunda estrutura, pode ocorrer a topicalização do SN (PERINI, 2010, p. 331) ou em estruturas relativas na forma **SN + que + verbo existencial**. Essa segunda estrutura de SN à esquerda foi a mais encontrada em nossa amostra. Segue um exemplo de cada fator:

- SN à direita do verbo

(13) TEM muitA gatinha lá (NORPOFOR, D2, 141)

- SN à esquerda do verbo

(14) as irmã fica lá perto do North North Shopping por lá numa capelinha que TEM (NORPOFOR, D2, 129)

4.7.2.2.5 *Concordância entre o verbo e o SN*

Esta foi mais uma variável analisada por Silva (2001). A autora controlou as seguintes estruturas **V (singular) → AI (singular)**, **V (singular) → AI (plural)** e **V (plural) → AI (plural)**. Esta última não ocorreu nenhuma vez em nossa amostra, portanto, essa variável ficou com apenas dois fatores:

- V (singular) → AI (singular)

(15) quer dizer num HÁ pressa pra gente nunca chegar lá né?
(NORPOFOR, D2, 93)

- V (singular) → AI (plural)

(16) TINHA umas casinha lá em ci::ma... (NORPOFOR, D2, 93)

Em (16), estamos considerando o SN plural, uma vez que a marca morfológica do determinante indica claramente que o falante está se referindo a mais de uma casinha. Levando em consideração que estamos trabalhando com o falar popular, esse tipo de ocorrência representa a maioria dos dados.

4.7.2.2.6 *Presença x ausência de elementos à esquerda do verbo*

Essa variável foi controlada por Vitório (2013) e diz respeito ao preenchimento da posição anterior ao verbo. Esse espaço pode ser preenchido por advérbios, palavras de negação, pronomes relativos, entre outros elementos. Vitório (2013) encontrou 85% das ocorrências na fala com presença de elementos à esquerda do verbo. No entanto, a autora não

diz se essa variável foi selecionada pelo **GoldVarb X** como relevante em seu estudo, apenas aborda os resultados percentuais.

Em nossa amostra, também encontramos um alto percentual de ocorrências com a presença de elementos à esquerda do verbo (70%). No entanto, nosso objetivo primeiro com a escolha dessa variável é descobrir se a presença desses elementos realmente favorece uma das variantes em estudo. Feito essa ressalva, essa variável ficou assim:

- Presença de preenchedor
(17) não TEM condições... tu é assim é? (NORPOFOR, D2, 101)
- Ausência de preenchedor
(18) TEM todo um valor né? (NORPOFOR, D2, 15)

4.7.2.2.7 *Repetição do verbo no mesmo enunciado*

Essa última variável linguística foi definida após a coleta de todos os dados, pois isso permitiu percebermos um fenômeno que poderia ser relevante para a variação *haver/existir/ter*: em várias ocorrências, havia a repetição do verbo no mesmo enunciado. Por isso, analisamos os seguintes fatores:

- Repetição do verbo
(19) o certo é esse L. mas não tem condições... não TEM condições...
(NORPOFOR, D2, 101)
- Não repetição do verbo
(20) agora TEM muita gente que faz isso né... tipo essas meninazinha eu pensando que era tudo crentinha da igreja...

4.8 CODIFICAÇÃO DOS DADOS

Depois de selecionadas as variáveis que serviram de base para nossa análise, o passo seguinte foi criar uma chave de codificação, isto é, todos os fatores de todas as variáveis devem ser codificados para que o programa computacional possa fazer as análises estatísticas.

Os códigos a serem utilizados nos fatores ficam a critério do pesquisador. É importante criá-los de modo que ele possa lembrar o fator facilmente durante toda a análise, pois todos os resultados gerados pelo programa são expostos por meio desses códigos. Um exemplo dessa codificação pode ser observado em (21) e para a visualização de toda a chave de codificação, ver APÊNDICE C:

(21) (tm29-o*dziu< tem que TER umas palmada

A codificação da ocorrência acima (iniciada por um parêntese por exigência do programa) deve ser lida da seguinte forma:

t → trata-se da variante *ter*,

m → enunciada por um falante do sexo masculino,

2 → da faixa etária de 26 a 49 anos,

9 → com escolaridade de 9 a 11 anos,

- → seguido de SN com traço [-humano],

o → a variante *ter* não ocorre na fala do interlocutor,

***** → há elementos à esquerda do verbo,

d → o SN está posicionado à direita do verbo,

z → trata-se de um SN simples,

i → a variante *ter* está na forma de infinitivo,

u → a variante *ter* aparece anteriormente no mesmo enunciado e

< → a concordância se dá por meio da estrutura V (singular) + SN (plural).

São esses dados que servirão de *input* para o **GoldVarb X** fazer todas as análises. Portanto, as sentenças expostas após a codificação não são lidas pelo programa, servindo somente para a leitura do pesquisador.

4.9 GOLDFARB X

Este trabalho utiliza o programa computacional **GoldVarb X**, que é uma versão atualizada do pacote de programas VARBRUL. De acordo com Guy e Zilles (2007),

O varbrul é um conjunto de programas computacionais de análise multivariada, especificamente estruturado para acomodar dados de variação sociolinguística. A

análise se chama ‘multivariada’ porque permite investigar situações em que a variável linguística em estudo é influenciada por vários elementos do contexto, ou seja, múltiplas variáveis independentes. (GUY; ZILLES, 2007. p.105, aspas dos autores).

É importante destacar que a análise é “multivariada”, ou seja, o programa leva em consideração não cada variável isoladamente, mas avalia a relação estatística que as variáveis têm entre si em diferentes contextos. Essa avaliação é representada por meio de outras medidas além das porcentagens, como os pesos relativos, que fornecem dados mais consistentes do que as porcentagens. Os autores afirmam que fazer uma análise somente levando em consideração a frequência de uso e as porcentagens é “fazer uma análise univariada (cf. a seção 5 acima) e, portanto, suscetível aos possíveis resultados incorretos de qualquer método que não controle variáveis intervenientes.” (GUY; ZILLES, 2007. p. 106–parênteses dos autores). É possível, numa rodada do programa, encontrarmos um determinado fator que apresente alta porcentagem, mas que os dados estatísticos nos mostram que ele não é favorável ao fenômeno variável ao estudo.

O peso relativo é estabelecido de 0 a 1,00 e é calculado de acordo com o tipo de variável dependente (binária, ternária ou eneária) de modo que a soma do peso de cada variável seja 1,00. Por exemplo, numa variável binária, a média do peso relativo é 0,50; numa ternária, é 0,33 e assim por diante. Dizemos que o fator favorece a variante quando o peso relativo é maior que a média; quando ele for menor, desfavorece a variante em análise, mas favorece a outra variante; e quando ele for igual à média, não favorece nenhuma das variantes. Por exemplo, na análise de dois fatores de uma variável, onde o primeiro apresenta peso relativo de 0,78 e o segundo tem peso relativo de 0,22, o primeiro fator favorece a variante em análise, enquanto que o segundo não favorece.

Guy e Zilles (2007, p. 106-107) admitem que outros métodos multivariados podem ser usados para satisfazer a pesquisa sociolinguística, mas apresentam três vantagens que fazem do VARBRUL/GOLDVARB X o método mais utilizado pelos sociolinguistas variacionistas. A primeira vantagem é que ele foi projetado exatamente para a estruturação da linguagem natural, por isso, pode-se trabalhar com células que apresentem quantidades diferentes de ocorrências. A segunda vantagem é que o programa permite o manuseio de células que não apresentam nenhum dado. E, finalmente, a terceira vantagem diz respeito à possibilidade de recodificação dos dados no próprio sistema. É por todos esses motivos que o VARBRUL/GOLDVARB X é o pacote de programas mais utilizado nas pesquisas variacionistas. Ele realiza diferentes tipos de análises, mas as que usamos para a realização

desta pesquisa foram as seguintes: a rodada geral e a análise binomial (rodadas *step up* e *step down*). O primeiro tipo de análise nos fornece a frequência de uso de cada variante em todos os fatores de todas as variáveis independentes analisadas. Portanto, esse passo nos dá uma visão geral dos dados e mostra os fatores onde houve nocautes, que devem ser resolvidos para que o programa faça a análise binomial. Os nocautes ocorrem quando, por exemplo, em uma determinada variável, um fator não apresenta dados de uma variante. Dizemos então que, no referido contexto, a regra é categórica porque só ocorre em um dos fatores (ver Tabela 2).

Há três modos básicos de resolvermos os nocautes, que dependem da escolha do pesquisador: ou desprezamos os dados, ou eliminamo-los ou amalgamamos o fator com nocaute com outro que não os tenha apresentado. Amalgamar fatores significa juntá-los em um só (ver Tabela 2), porém, para que o pesquisador possa fazer esse procedimento, é necessário que os fatores a serem amalgamados apresentem uma semelhança linguística ou social além de uma semelhança quantitativa, ou seja, na frequência de uso ou no peso relativo (GUY; ZILLES, 2007, p. 188-189). Outra forma de amálgama que o **GoldVarb X** nos permite fazer são os cruzamentos. Quando queremos investigar a influência de dois grupos de fatores considerados juntos, formamos um novo grupo de fatores, cruzando cada fator do primeiro grupo com cada fator do segundo grupo.

Por exemplo, um dos cruzamentos que realizamos neste trabalho foi o dos grupos **sexo** e **faixa etária** na análise *haver/ter*. Como o grupo **sexo** tem 2 fatores (masculino e feminino) e o grupo **faixa etária** tem 3 (15 a 25 anos; 26 a 29 anos e a partir de 50 anos), o novo grupo criado (vale ressaltar que esse grupo deve ser desconsiderado isoladamente) terá 6 fatores (2 x 3), que serão: homens de 15 a 25 anos; homens de 26 a 29 anos; homens a partir de 50 anos; mulheres de 15 a 25 anos; mulheres de 26 a 29 anos; mulheres a partir de 50 anos.

O segundo tipo de análise que utilizamos aqui foi a binomial. Nesta análise, o programa realiza rodadas *step up* e *step down*, além de fornecer, para cada rodada, dados estatísticos relevantes para a interpretação do pesquisador, como a convergência da interação (*convergence at interaction*), o *input*, a significância estatística (*significance*) e o logaritmo de verossimilhança (*log likelihood*). Todos esses dados são apresentados em cada análise binomial que realizamos. Antes de explicarmos o que são rodadas *step up* e *step down*, vamos entender um pouco esses dados estatísticos.

A convergência da interação (*convergence at interaction*) diz respeito à quantidade de interações que o programa teve que fazer para chegar à melhor análise da

rodada. O *GoldVarb X* realiza até 20 interações e caso chegue a esse número sem ter encontrado convergência, ele para e nos informa que, naquela rodada, não houve convergência. “Normalmente esse número não significa muito para o pesquisador, desde que consiga convergência.” (GUY; ZILLES, 2007, p. 198). Mesmo assim, decidimos mostrar essa informação, para que o leitor saiba que só analisamos rodadas onde houve convergência dos dados.

O *input* de cada rodada nos revela a probabilidade da variante analisada ocorrer no(s) grupo (s) de fator(es) analisado(s) na referida rodada. Esse valor é fornecido em números de 0,000 a 1,000. Portanto, um *input* de 0,024 nos diz que a variante analisada tem 2,4% (24/1000) de probabilidade de ocorrer no contexto da rodada.

A significância estatística (*significance*) diz respeito à probabilidade de uma determinada análise condizer com os dados do mundo real. Esse cálculo, assim, como o *input*, também é fornecido em números de 0,000 a 1,000. No entanto, existe um limite para considerarmos a rodada estatisticamente significativa, o que podemos, *grosso modo*, dizer que é a margem de erro da análise: 0,050, ou seja, 5%. Rodadas que apresentam uma significância maior que 0,050 não têm significância estatística porque apresentam uma margem de erro maior que 5% e, portanto, não é confiável.

O último valor estatístico que apresentamos em nossas análises é logaritmo de verossimilhança (*log likelihood*). Guy e Zilles (2007, p. 238) definem-no como sendo

um número calculado pela rotina do Varbrul que mede a qualidade da aproximação entre o modelo (os fatores que caracterizam os contextos, os pesos associados com os fatores, o *input* e o modelo matemático logístico) e os dados observados.” (GUY; ZILLES, 2007, p. 238, grifo dos autores)

Esse número é sempre negativo e devemos interpretá-lo da seguinte forma: quanto maior o *log likelihood* for, i.e., mais próximo de 0, melhor é a análise da rodada. Dessa forma, se tivermos que escolhermos entre duas análises distintas, por exemplo, uma de *log likelihood* -111, 786 e outra de *log likelihood* -109, 456, ficamos com a segunda opção. No entanto, na análise binomial, o programa já seleciona a melhor rodada a partir da avaliação dos valores estatísticos explicitados acima (*convergence*, *input*, *log likelihood* e *significance*). Isso facilita muito o trabalho do pesquisador.

Na análise binomial, o programa revela os dados em níveis. Nas rodadas *step up*, ele começa no nível 0, onde fornece o *input* (que, neste nível, é exatamente igual à porcentagem da regra de aplicação) e um *log likelihood* iniciais, sem a análise de nenhum

grupo de fatores. No nível 1, o programa analisa cada grupo de fatores isoladamente, fornecendo, além do *input* e do *log likelihood*, a quantidade de interações realizadas e significância estatística de cada grupo. Nesse nível ocorre a seleção do grupo que apresenta a melhor significância estatística (o menor número entre 0,000 e 0,050). No nível 2, o programa analisa a interação desse grupo selecionado com cada um dos demais. Se houver a seleção de outro grupo de fatores, o programa segue na análise, fazendo rodadas dos dois grupos selecionados com cada um dos restantes no nível 3 e assim sucessivamente até não haver mais a seleção de grupos por falta de significância estatística. Após esse processo, o programa revela a melhor rodada de todas, que é aquela que contém os grupos de fatores mais favorecedores do fenômeno investigado.

As rodadas *step down* fazem um processo inverso do procedimento *step down* e, começando pelo nível maior de análise (que depende do número de grupos de fatores analisados), vai excluindo os grupos não significativos em cada nível. No caso de nossa primeira análise *haver/ter*, trabalhamos com 10 grupos e, portanto, a análise *step down* começa no nível 10, analisando todos os grupos. Depois, as rodadas do nível 9, analisam todas as combinações de 9 grupos e exclui um grupo de fatores, as do nível 8 analisam todas as combinações de 8 grupos e excluem mais um grupo, e assim sucessivamente. Ao final, o programa fornece também a melhor rodada *step down*, ou seja, aquela que apresenta a pior significância estatística. Os pesos relativos da melhor rodada *step up* (melhor significância estatística) devem ser iguais aos pesos relativos da melhor rodada *step down* (pior significância estatística). Em outras palavras, as análises *step down* servem para confirmar as análises *step up*, dando mais credibilidade às análises do **GoldVarb X**. Foi por tudo isso que podemos fazer com esse programa que o escolhemos para o estudo estatístico de nossos dados.

5 DESCRIÇÃO DOS DADOS E ANÁLISE DOS RESULTADOS

Nesta seção, mostraremos os resultados da análise dos dados da interação entre *haver*, *existir* e *ter*. No entanto, antes de iniciarmos a análise propriamente dita, fizemos uma rodada experimental para podermos obter uma visão geral das ocorrências dentro de cada grupo de fatores e, assim, termos condições de fazer alguns ajustes em nosso sistema analítico.

É comum nos trabalhos variacionistas a comparação com outros estudos sobre o mesmo tema de diferentes regiões, com o objetivo de entender o fenômeno de forma macro e, conseqüentemente, colaborar para uma descrição mais ampla do Português Brasileiro, trabalho que linguistas como Perini (1996, 2008, 2010), Azeredo (2010), Bagno (2011) e Castilho (2012) já vem fazendo com muita dedicação, como já anteriormente mencionado. No entanto, devido à falta de homogeneidade das metodologias (e.g. diferentes divisões de escolaridade, não apresentação de pesos relativos em alguns estudos, entre outras diferenças) dos 8 trabalhos que serviram de base para esta pesquisa (1 dissertação, 2 teses e 5 artigos), decidimos apenas descrever nossos dados e discutir nossos resultados sem fazer comparações detalhadas entre os falares de Fortaleza e de outras cidades. Deter-nos-emos a comparar apenas resultados gerais, como as porcentagens, porque nesse aspecto a metodologia parece não interferir nos resultados, uma vez que todos os trabalhos convergem para uma mesma tendência: a supremacia do verbo *ter* em relação a *haver* e *existir*, quando analisado. No ponto onde as metodologias convergem (caso das faixas etárias), mesmo que aproximadamente, teceremos alguns comentários sobre os dados, mas sem fechar conclusões sobre os falares das cidades em questão.

5.1 RODADA EXPERIMENTAL

Nessa rodada teste, o **GoldVarb X** contabilizou 943 ocorrências, que estão divididas da seguinte forma: *haver* (27), *existir* (28) e *ter* (888). Com esse teste foi possível observar algumas peculiaridades nos grupos de fatores **peso do SN** e **forma verbal**. No primeiro grupo, 2 dos 4 fatores comparados apresentaram comportamento semelhante: o fator **SN nulo** e o fator **núcleo do SN nulo**, como ilustram as ocorrências retiradas de nossa amostra:

- (1) TINHA... tava todo mundo lá na igreja (referente: *gente*): **SN nulo**. (NORPOFOR, D2 50)
- (2) TEM uns que o atendimento é ho-rrível (referente: *restaurantes*): **núcleo do SN nulo** (NORPOFOR, D2 15)

Nosso objetivo, ao controlar esse dois fatores, era perceber se havia alguma diferença significativa entre os dois tipos de nulidade, no entanto, como mostra a tabela 1, essa divisão não se mostrou produtiva na amostra, uma vez que verificamos uma regra categórica no fator **núcleo do SN nulo**, onde só houve ocorrências do verbo *ter* (39).

Tabela 1 - Distribuição das ocorrências no grupo **peso do SN** nos três verbos

FATORES	haver	Existir	Ter	total	%
SN Simples	19	19	564	602	63,8
SN Complexo	5	7	247	259	27,5
SN nulo	3	2	38	43	4,6
Núcleo do SN nulo	0	0	39	39	4,1
TOTAL	27	28	888	943	100%

Como esses dois fatores apresentam frequências de uso aproximadas (4,6% e 4,1%, respectivamente) e têm a nulidade como característica linguística comum, optamos por recodificar esse grupo de modo que somente a nulidade (do SN ou do núcleo) passou a ser analisada. Dessa forma, esse grupo passou de 4 para 3 fatores (SN simples, SN complexo e SN nulo).

No segundo grupo que apresentou peculiaridades, **forma verbal**, existem 9 fatores. Esse grupo destoa dos demais em quantidade de fatores porque queríamos averiguar se outras formas verbais, exceto as mais usadas na literatura acerca do tema (**presente**, **pretérito perfeito** e **pretérito imperfeito**), apresentariam dados significativos em nosso estudo. As ocorrências nesse grupo são expostas na tabela 2.

Tabela 2 - Distribuição das ocorrências no grupo **forma verbal** nos três verbos.

FATORES	haver	existir	Ter	total	%
Presente do indicativo	10	20	578	608	64,5
Pretérito imperfeito do indicativo	2	6	229	237	25,1
Infinitivo	9	1	39	49	5,2
Pretérito perfeito do indicativo	6	1	30	37	3,9
Gerúndio	0	0	5	5	0,5
Futuro do subjuntivo	0	0	4	4	0,4
Presente do subjuntivo	0	0	1	1	0,1
Futuro do pretérito do indicativo	0	0	1	1	0,1
Pretérito imperfeito do subjuntivo	0	0	1	1	0,1
TOTAL	27	28	888	943	100%

Como fica claro ao observarmos a tabela 3, as formas mais encontradas em nossa amostra foram o **presente do indicativo** e o **pretérito imperfeito do indicativo**, totalizando juntas 89,8% de todos os dados. As formas **infinitivo** e **pretérito perfeito do indicativo** representam somente 9,0% dos dados, mas resolvemos mantê-las no grupo por apresentarem ocorrências em todos os verbos e ainda representarem um percentual estatisticamente relevante. No entanto, as demais formas (**gerúndio**, **futuro do subjuntivo**, **presente do subjuntivo**, **futuro do pretérito do indicativo** e **pretérito imperfeito do subjuntivo**) ocorreram muito pouco na amostra, representando apenas um total de 1,2%. Portanto, decidimos excluir esses 5 fatores desse grupo e, conseqüentemente, seus 12 dados da amostra, uma vez que esse valor é insignificante estatisticamente. Guy e Zilles (2007, p. 62, maiúsculas dos autores) afirmam que o objetivo do pesquisador em Sociolinguística Quantitativa deve ser o de obter uma análise “EFICIENTE, GERAL E PODEROSA” e que para isso “temos que identificar os fatores não significativos, que em nada contribuem para a análise, e eliminá-los”. Foi exatamente isso o que fizemos no grupo **forma verbal**, que ficou com 4 fatores (**infinitivo**, **presente do indicativo** [doravante, **presente**], **pretérito perfeito do indicativo** [doravante, **perfeito**] e **pretérito imperfeito do indicativo** [doravante, **imperfeito**]). Em nossa amostra, restaram 931 dados para análise. Seguem as ocorrências que foram excluídas do nosso estudo:

a. Gerúndio

- (3) que tá TENDO um empecilho porque... eu tou desempregado há:: um... tempo né?... (NORPOFOR, D2 155)
- (4) respeito de aluno com os professores nã/ quase não tá TENDO... (NORPOFOR, D2 106)
- (5) tá TENDO vaga? (NORPOFOR, D2 114)
- (6) as coisas são muito mais difícil por causa da água porque TENDO água acaba com a dificuldade (NORPOFOR, D2 04)
- (7) com certeza TENDO um grupo bom (NORPOFOR, D2 14)

b. Futuro do subjuntivo

- (8) se não TIVER o castigo... (NORPOFOR, D2 04)
- (9) eu acho que tem que ter conversado com ele porque se não TIVER isso aí... (NORPOFOR, D2 04)
- (10) agora se não TIVER ordem? (NORPOFOR, D2 04)
- (11) cê está habilitado quando TIVER vaga você VAI agora todo mundo sabe (NORPOFOR, D2 14)

c. Presente do subjuntivo

- (12) é mais difícil ter droga dentro da igreja né? pode até ser que TENHA ninguém sabe (NORPOFOR, D2 04)

d. pretérito imperfeito do subjuntivo

- (13) mas se TIVESSE empregozinho de qualquer coisa eu enfrentava... (NORPOFOR, D2 94)

e. futuro do pretérito do indicativo

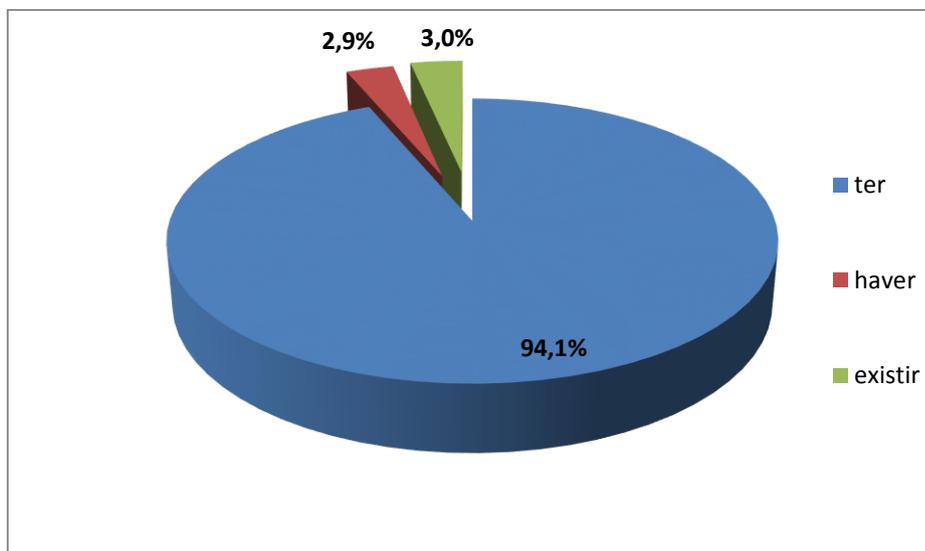
- (14) ela disse que se ela levasse a turma pra lá::... aí TERIA uma comissão:: um abatimento::... (NORPOFOR, D2 157)

5.2 RODADA TERNÁRIA: VISÃO GERAL DAS VARIANTES

Após recodificarmos o grupo **peso do SN** e excluirmos as 12 ocorrências do grupo **forma verbal**, passamos a análise propriamente dita dos 931 dados restantes.

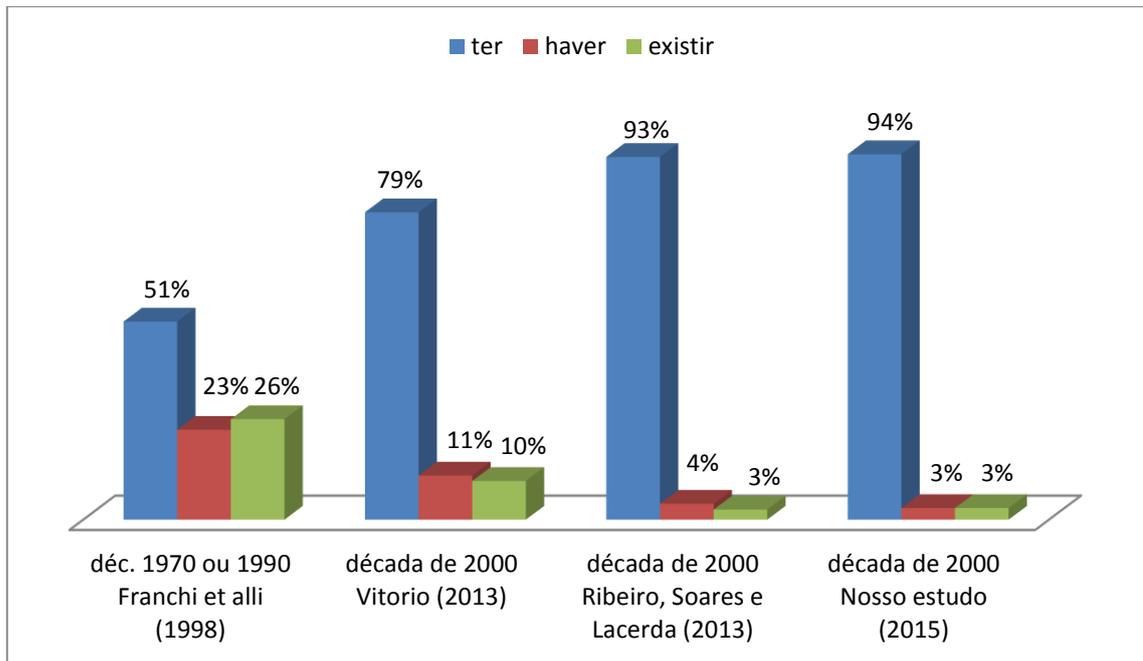
O **GoldVarb X** somente produz análises com seleção de fatores relevantes, pesos relativos, significância estatística, entre outros, quando trabalhamos com variáveis binárias e, como nossa variável é ternária (*haver*, *existir* e *ter*), tivemos que fazer as análises investigando dois verbos por vez. No entanto, antes de procedermos com as rodadas binárias, resolvemos fazer uma rodada inicial ternária somente para sabermos a frequência de uso de cada verbo na amostra, o que nos deu uma visão geral do fenômeno. Essa rodada não apresentou nenhum nocaute.

Gráfico 2 – Frequência de uso das variantes *haver*, *existir* e *ter*



Como vemos no gráfico 2, o verbo *ter* é o verbo mais utilizado pelos falantes da nossa amostra, com 94,1% (876 ocorrências), enquanto que os verbos *haver* e *existir* disputam a segunda preferência, com, respectivamente, 2,9% (27 ocorrências) e 3,0% (28 dados). Esse resultado corrobora os resultados de Franchi, Negrão e Viotti (1998), Ribeiro, Soares e Lacerda (2013) e Vitória (2013), que também apresentam o verbo *ter* como o que possui o maior percentual em contraste com *haver* e *existir*, como ilustra o gráfico 3.

Gráfico 3 - Comparação das frequências de uso de *haver*, *existir* e *ter* em quatro estudos



Fonte: elaborado pelo autor

Nos principais estudos que servem de base a esta pesquisa, apresentados na seção 2., o verbo *ter* foi a variante escolhida como valor de aplicação da regra, isto é, todos os resultados gerados pelo **GoldVarb X** são dados de *ter* e o foco da investigação desses trabalhos é descobrir o papel dele em relação à variante *haver*. Em nosso estudo, optamos pelos verbos *haver* e *existir* porque eles apresentaram uma frequência de uso muito baixa e, portanto, o que nos interessa saber é qual o espaço que esses dois verbos possuem dentro das orações existenciais em contraste com *ter*. Também nos é pertinente investigar a disputa acirrada que parece haver entre os verbos *haver* e *existir*, que, ao fazermos o arredondamento, apresentam a mesma frequência de uso. Dessa forma, o verbo *haver* é a regra de aplicação nas análises *haver/ter* (903 dados) e *haver/existir* (55 dados), e o verbo *existir* na análise *existir/ter* (904 dados).

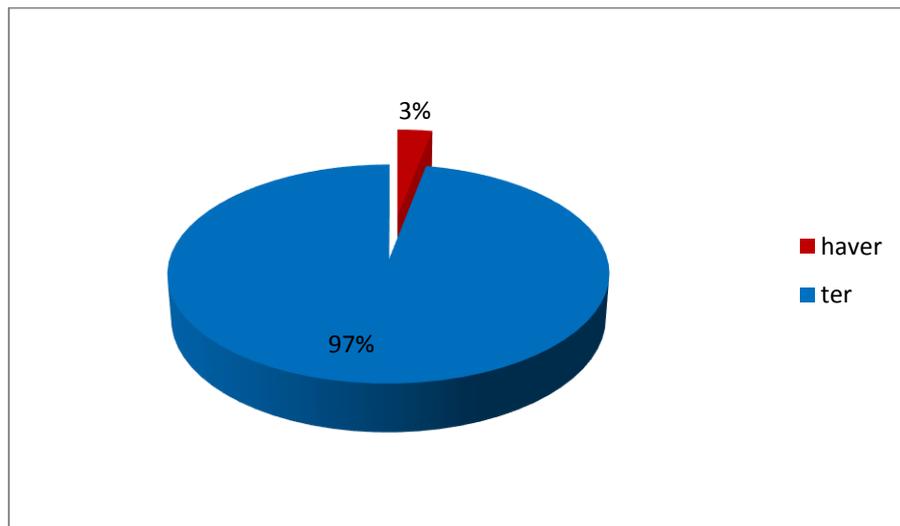
5.3 ANÁLISES DE *HAVER/TER*

Depois de termos a visão geral do fenômeno, passamos a analisar o comportamento dos verbos *haver*, *existir* e *ter* em análises binárias. Primeiramente, faremos o estudo de *haver* e *ter*, depois de *existir* e *ter* e, por último, apesar dos poucos dados de que dispomos, analisaremos os contextos de uso de *haver* e *existir*.

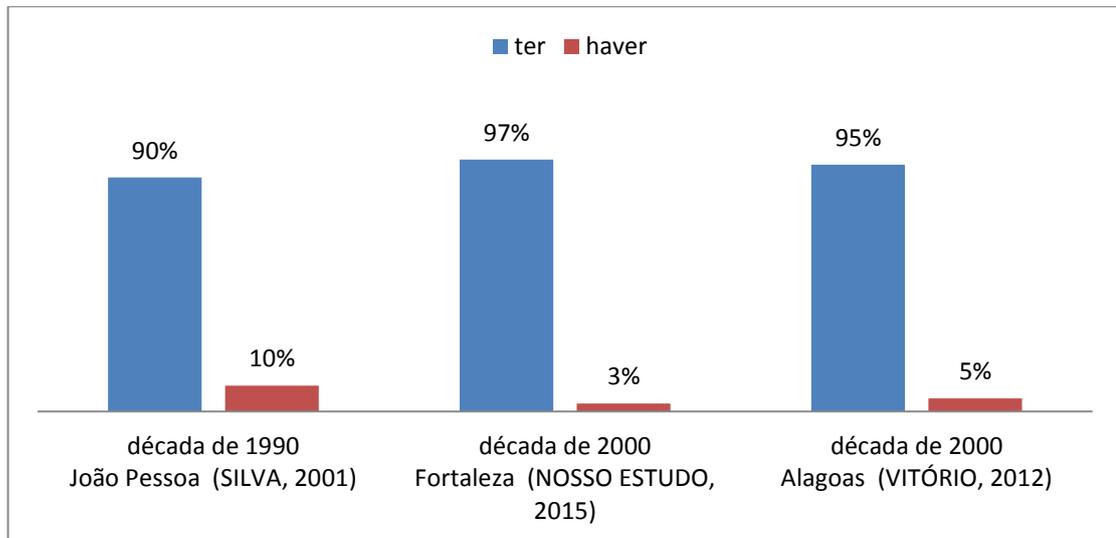
5.3.1 Primeira análise: visão geral da regra

Ao realizarmos a rodada inicial para analisar a variação *haver/ter*, não houve nenhum nocaute e a distribuição dos dados nas variantes ficou da seguinte forma: 28 são do verbo *haver* e 876 são do verbo *ter*. O gráfico 4 mostra a frequência de aplicação de cada verbo e revela uma regra quase categórica em favor de *ter* na amostra em estudo.

Gráfico 4 - Frequência de uso das variantes *haver* e *ter*

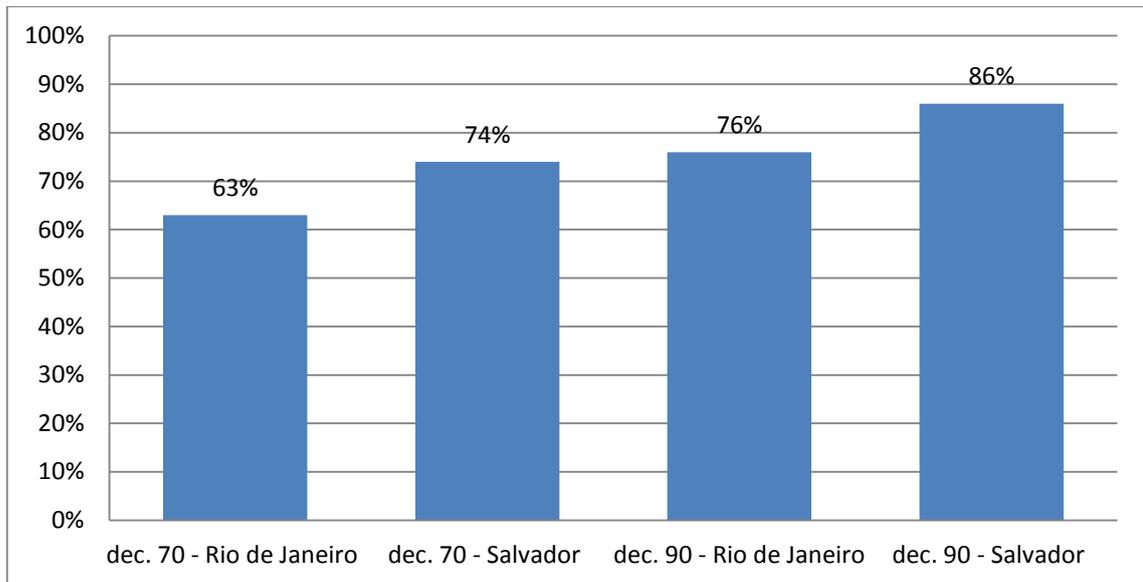


Bagno (2011, p. 624-625) sustenta que o verbo *ter* é o verbo existencial mais utilizado pelos falantes do Português Brasileiro. Vale ressaltar que essa afirmação de Bagno (2011) diz respeito a uma lista de 7 verbos existenciais (*ser, ter, haver, existir, fazer, estar, tratar-se de*). O resultado inicial de nossa amostra, apresentado no gráfico 2, corrobora essa afirmação, considerando tanto o verbo *haver* quanto o verbo *existir* (resultados serão discutidos adiante). Os resultados dos estudos de Silva (2001), que representa o falar de João Pessoa, e de Vitório (2012), que representa o falar de Maceió também vão de encontro ao exposto por Bagno (2011) no contraste entre *ter* e *haver*.

Gráfico 5 - Comparação dos resultados gerais de *haver/ter* de Silva (2001), nosso estudo (2015) e Vitória (2012)

Fonte: elaborado pelo autor

O gráfico 5 ilustra bem a preferência dos falantes desses três estudos pelo verbo *ter*. Podemos ver ainda que nossos dados superam os dos outros dois, mostrando quase um uso categórico dessa variante. Uma possível explicação para isso pode ser o fato de essas duas autoras terem incluído, em suas amostras, falantes de nível superior, respectivamente, 20% e 23% do número total de falantes, enquanto que em nossa amostra não há nenhum falante de nível superior. Os estudos variacionistas têm mostrado que há uma relação entre o nível de escolarização e o uso de formas consideradas padrão (no caso desta análise, o verbo *haver*) no sentido de que quanto maior a escolaridade do falante, maior a frequência de uso das formas padronizadas. Isso explica, de certa forma, o porquê de o verbo *haver* ter tido uma frequência tão baixa em nosso estudo. Vale ressaltar aqui que o verbo *ter*, apesar de não ser considerado uma forma padrão nas gramáticas normativas, não sofre estigma social por parte dos falantes e, talvez por isso, também apresente índices altos em estudos que analisam dados representativos da fala de pessoas com alto grau de escolarização, como os falantes do projeto NURC. Martins e Callou (2003), ao analisarem os dados desse projeto das décadas de 70 e 90 em Salvador e no Rio de Janeiro, encontraram os seguintes resultados:

Gráfico 6 - Frequências de uso de *ter* nos dados do projeto NURC

Fonte: Elaborado pelo autor a partir dos dados de Martins e Callou (2003).

O gráfico 6 confirma a alta frequência de uso de *ter* tanto na década de 70 quanto na década de 90 nos dados das duas cidades. Isso reforça que essa variante não sofre estigma por parte dos falantes escolarizados, uma vez que todos os falantes do projeto NURC possuem alto grau de escolarização (nível superior completo).

5.3.2 Segunda análise: visão mais detalhada da regra

Apesar dos resultados iniciais mostrarem uma regra quase categórica da aplicação de *ter* em relação a *haver*, é importante descobrirmos quais os fatores responsáveis pela permanência do verbo *haver* no falar popular. Para isso, prosseguimos na análise mais detalhada dos dados. Nessa etapa, o programa selecionou o *step up* #13 como a melhor análise (*convergence at interaction 8; input 0,014; log likelihood = -94,645; significance = 0.001*). Os grupos de fatores selecionados como condicionadores da aplicação de *haver* na amostra analisada foram, por ordem de relevância, a **forma verbal** e a **faixa etária**. O que nos chama a atenção aqui é o *input* de 0,014, que, como explicado na metodologia, representa a probabilidade de aplicação da regra (no caso, *haver*) na interação das variáveis favorecedoras. Neste caso, o *input* de 0,014 significa que a probabilidade de *haver* ocorrer quando os grupos de fatores **forma verbal** e **faixa etária** interagem é de 1,4%, ou seja, muito baixa. Isso reforça os resultados percentuais quase categóricos de *ter* e justifica nosso interesse em investigar o porquê dessa variação ainda existir na amostra em estudo. É isso que passamos a analisar a

partir de agora. Discutiremos, a seguir, os resultados obtidos para cada uma das variáveis selecionadas:

5.3.2.1 Forma verbal

Tabela 3 - Atuação da variável **forma verbal** sobre o verbo *haver*

FATORES	Aplic./Total	%	P.R.
Infinitivo	9/48	18,8	0,956
Perfeito	6/36	16,7	0,929
Presente	10/588	1,7	0,508
Imperfeito	2/231	0,9	0.245

Como ilustra a tabela 4, o **infinitivo** e o **perfeito** são as formas verbais que mais favorecem o verbo *haver*, enquanto que o **imperfeito** favorece *ter*. O **presente** se manteve no nível da neutralidade, nem favorecendo um nem outro, apesar de minimamente tender a favorecer *haver*. Os altos pesos relativos do **infinitivo** e do **perfeito** não deixam dúvidas de que, na amostra em estudo, esses fatores são dois dos responsáveis pela existência de *haver* nos dados analisados. Vale ressaltar que a forma verbal TINHA já se estabilizou como favorita dos falantes em estruturas em que funciona como auxiliar, formando o pretérito mais que perfeito (TINHA + particípio passado). Portanto, é lógico ter se expandido também em outras estruturas dentro do sistema, e.g., nas estruturas existenciais. Como mostram os resultados da tabela 4, essa forma verbal ganha de HAVIA também nas estruturas existenciais.

5.3.2.2 Faixa etária

A **faixa etária** é uma variável muito relevante nos estudos variacionistas e quando ela é selecionada pelo programa estatístico como um dos fatores condicionadores do fenômeno, possibilita-nos fazer inferências sobre o processo de uma possível mudança em curso observada no **tempo aparente**. Essa discussão será feita na seção 5.6, na página 78. Alguns estudos que serviram de literatura para nosso trabalho, abordaram a análise da **faixa etária** para inferir se a variável estudada representa mudança em curso ou variação estável.

Silva (2001) observou, ao estudar a comunidade de João Pessoa, uma possível variação estável enquanto que Martins e Callou (2003), analisando a fala do Rio de Janeiro e de Salvador, e Vitório (2012), analisando a comunidade de Maceió, detectaram uma provável mudança em curso.

Na análise *haver/ter* desta pesquisa, a **faixa etária** foi a segunda variável selecionada como relevante e seus resultados estão expostos na tabela 5.

Tabela 4 - Atuação da variável **faixa etária** sobre o verbo *haver*

FATORES	Aplic./Total	%	P.R.
A partir de 50 anos	16/350	4,6	0,733
15 a 25 anos	5/204	2,5	0,389
26 a 49 anos	6/349	1,7	0,321

Por meio desses resultados, podemos perceber que a faixa etária **a partir de 50 anos** é a grande responsável pela manutenção de *haver* na amostra estudada ao passo que os falantes de **15 a 25 anos** e os falantes de **26 a 49 anos** são responsáveis pela restrição de seu uso, conseqüentemente, são favorecedores de *ter*. Batista (2012, p. 22) afirma que essa variante predomina na fala dos jovens. Como podemos ver, em nosso estudo, tanto os jovens quanto os adultos favorecem *ter*, com pesos relativos 0,611 e 0,679, respectivamente.

5.3.2.3 Faixa etária x sexo

Para termos uma percepção melhor desse fenômeno, fizemos o cruzamento dessa variável com as outras duas variáveis sociais (sexo e escolaridade). No entanto, o cruzamento **faixa etária x escolaridade** apresentou regras categóricas para os falantes de **0 a 4 anos de escolaridade das faixas etárias 1 e 2** e para os falantes de **5 a 8 anos de escolaridade da faixa etária 1**. Nesses contextos, só houve ocorrências do verbo *ter*. Portanto, somente o cruzamento **faixa etária x sexo** nos ofereceu pesos relativos, porque não apresentou regras categóricas e ainda foi selecionado pelo programa estatístico, no *step up* #18 (*convergence at interaction* 9; *input* 0,013; *log likelihood* = -93,669; *significance* = 0.009), como o segundo grupo de fator relevante para o fenômeno variável *haver/ter*.

Tabela 5 – Atuação do cruzamento **sexo x faixa etária** sobre o verbo *haver*

FATORES	15 a 25 anos			26 a 49 anos			A partir de 50 anos		
	Apl./total	%	P.R.	Apl./total	%	P.R.	Apl./total	%	P.R.
Masculino	2/3	66,7	0,386	5/15	33,3	0,375	6/14	42,9	0,667
Feminino	3/6	50,0	0,404	1/2	50,0	0,202	10/18	55,6	0,793

A tabela 5 mostra que tanto os **homens** quanto as **mulheres a partir de 50 anos** favorecem *haver*, enquanto que ambos os sexos dos **jovens** e **adultos** restringem essa variante. Os dados desse cruzamento reforçam a relevância que a faixa etária dos **idosos** tem para a existência de *haver* na amostra analisada.

Após realizarmos os cruzamentos da **faixa etária** com o **sexo** e com a **escolaridade**, também cruzamos essas duas últimas variáveis sociais. No entanto, não obtivemos os pesos relativos porque houve nocautes em três fatores: **jovens com 5 a 8 anos de escolaridade**, **jovens e adultos com 0 a 4 anos de escolaridade**.

5.3.3 Considerações sobre as análises de *haver/ter*

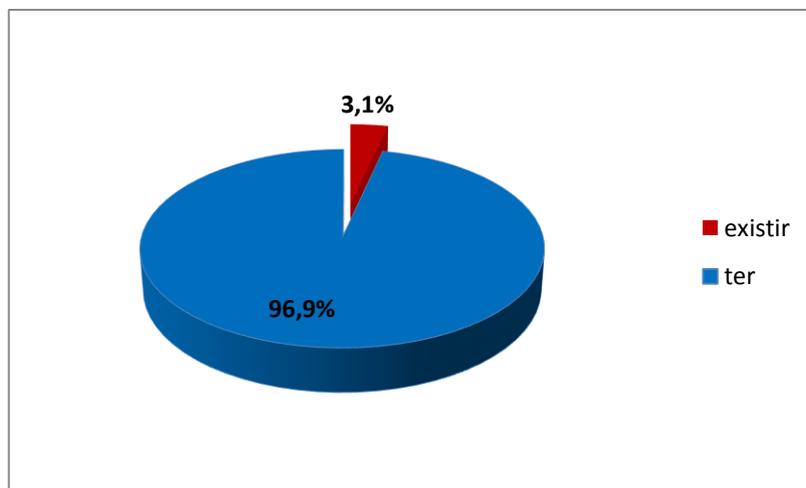
A análise dos dados de *haver* e *ter* revelou que, apesar da pouca representatividade em nossa amostra, o verbo *haver* ainda encontra contextos favoráveis a sua aplicação. As formas verbais **infinitivo** e **perfeito** foram os fatores mais favorecedores desse verbo de acordo com a seleção do **GoldVarb X**. Além disso, esses dois fatores apresentaram pesos relativos muito altos, respectivamente (0,956) e (0,929), o que demonstra que a atuação dessas duas formas verbais sobre a regra é quase categórica. O terceiro fator condicionador de *haver* é a faixa etária **a partir de 50 anos** (p.r. 0,733). Esse fator mostrou-se forte no cruzamento **faixa etária x sexo**, pois tanto os homens quanto as mulheres dessa faixa etária se mostraram favoráveis ao fenômeno, estas favorecendo mais do que aqueles.

5.4 ANÁLISES DE *EXISTIR/TER*

5.4.1 Primeira análise: visão geral da regra

Desconsiderando os dados da variante *haver*, ficaram 904 dados para serem rodados no **GoldVarb X** na análise *existir/ter*. Na primeira rodada, não houve nocautes e a frequência de uso de *existir*, assim como *haver*, é muito baixa, tratando-se, nessa análise, também de uma regra quase categórica em favor da variante *ter*. O gráfico 7 ilustra esses resultados iniciais.

Gráfico 7 - Frequência de uso das variantes *existir* e *ter*



5.4.2 Segunda análise: visão mais detalhada da regra

Nosso objetivo aqui é, assim como na análise *haver/ter*, identificar quais fatores são responsáveis pela manutenção da variante *existir* na amostra em estudo em contraste com a variante *ter*. Na análise multivariada dessas duas variantes, o programa selecionou, no *step up* #7 (*convergence at interaction 5; input 0,027; log likelihood = -118,170; significance = 0.000*), uma única variável como significativa na variação *existir/ter*. Essa variável selecionada foi a **posição do SN**. Nenhuma variável social foi selecionada, indicando que os princípios que regem a variação *existir/ter* são mais influenciados pela estrutura interna da língua do que pela estrutura social. O *input* dessa rodada reforça o resultado percentual, uma vez que indica uma probabilidade de 2,1% de *existir* ocorrer quando esse grupo está em foco. A seguir, explicaremos a relevância desse grupo de fator.

5.4.2.1 Posição do SN

Tabela 6 - Atuação da variável **posição do SN** sobre o verbo *existir*

FATORES	Aplic./Total	%	P.R.
SN à esquerda	8/60	13,3	0,848
SN à direita	20/844	2,4	0,469

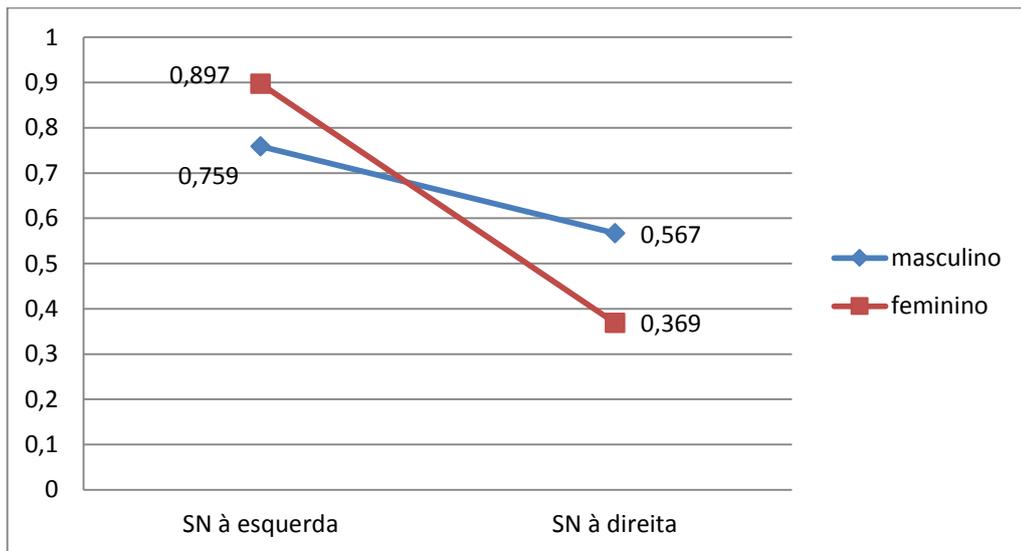
Controlamos a variável **posição do SN** com o intuito de verificar duas estruturas existenciais. A primeira é considerada, na literatura variacionista sobre o tema, como prototípica (V + SN) e representa a maior parte das ocorrências de nossa amostra, com 93,4% contra 6,6% da estrutura não prototípica (SN + V). Essa variável não se mostrou relevante na variação *haver/ter*, mas foi a primeira selecionada pelo **GoldVarb X** como favorecedora da variação *existir/ter*. A estrutura existencial não prototípica, representada pelo fator **SN à esquerda**, aparece em 28,57% das ocorrências do verbo *existir* e apenas em 5,93% das ocorrências do verbo *ter*. Isso justifica seu favorecimento àquela variante em detrimento dessa, como ilustra a tabela 6. Podemos concluir, portanto, que a inversão dos elementos V e SN nas estruturas existenciais contribui para a manutenção de *existir* em nossa amostra, apesar de a estrutura prototípica apresentar maior aplicação de ambos os verbos.

Aqui também realizamos cruzamentos com o objetivo de aprofundar mais a análise. Fizemos dois grupos de cruzamentos, a saber: variável linguística **posição do SN** com todas as variáveis sociais, que não foram selecionadas como relevantes pelo programa e as variáveis sociais entre si. Do primeiro grupo, os cruzamentos **posição do SN x sexo** e **posição do SN x escolaridade** foram os únicos nos quais não ocorreram nocautes e, portanto, puderam ser analisados. O cruzamento **posição do SN x faixa etária** apresentou nocautes e não prosseguimos na análise. Do segundo grupo, aquele que apresentou nocautes foi o cruzamento **faixa etária x escolaridade**. Já os cruzamentos **sexo x faixa etária** e **sexo x escolaridade** puderam ser analisados já que não tiveram nocautes. Portanto, são quatro cruzamentos da análise *existir/ter* que pudemos analisar. Desses quatro, apenas **sexo x faixa etária** não foi selecionado pelo programa como influenciador da variação em estudo. Os outros três foram selecionados pelo **GoldVarb X** como relevantes para a variação *existir/ter*: Os cruzamentos **posição do SN x sexo** e **posição do SN x escolaridade** foram selecionados como o primeiro grupo de fatores relevante e o cruzamento **sexo x escolaridade** foi o segundo em ordem de relevância. A seguir apresentamos os resultados de todos esses três cruzamentos.

5.4.2.2 Posição do SN x sexo

Este cruzamento foi selecionado como o primeiro grupo de fatores relevante na variação em estudo no *step up* #10 (*convergence at interaction 5; input 0,025; log likelihood = -115,926; significance = 0.000*).

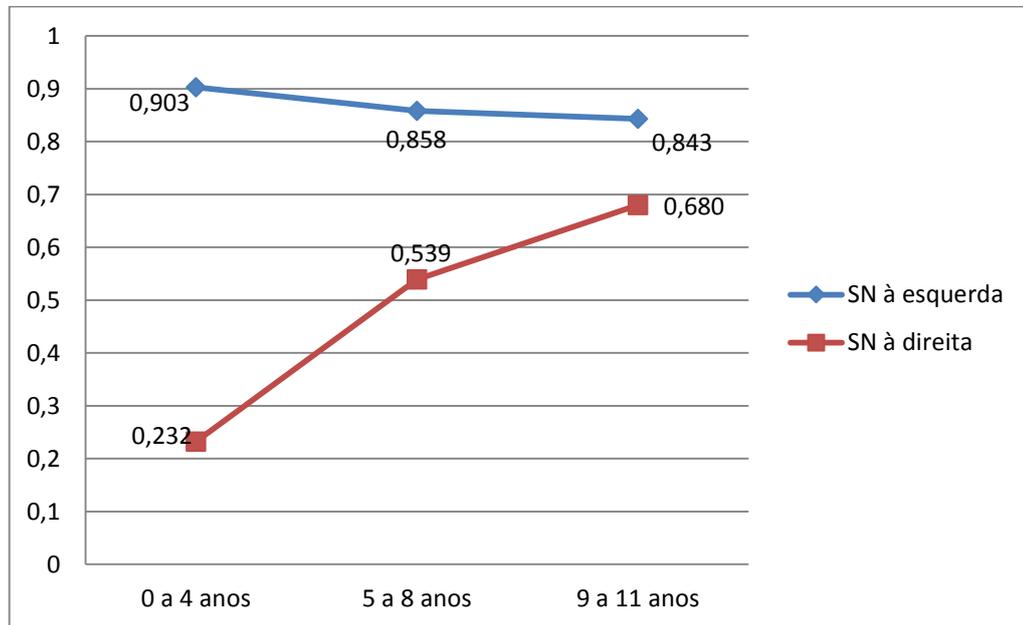
Gráfico 8 - Atuação do cruzamento **posição do SN x sexo** sobre o verbo *existir*



Como vemos no gráfico 8, tanto os **homens** quanto as **mulheres** são favorecedores de *existir* quando usam estruturas com **SN à esquerda do verbo**, com as mulheres avançando um pouco mais nesse favorecimento. Em relação às estruturas com **SN à direita do verbo**, os **homens** favorecem timidamente a aplicação de *existir*, enquanto que as **mulheres** retêm o uso dessa variável. Vemos pelos resultados acima que o padrão do gráfico de ambos os sexos são semelhantes, porém, mais uma vez aqui, as **mulheres** se mostram mais extremistas, visto que a diferença dos pesos relativos entre as duas estruturas são bem maiores nelas do que nos **homens** (0,528 contra 0,192).

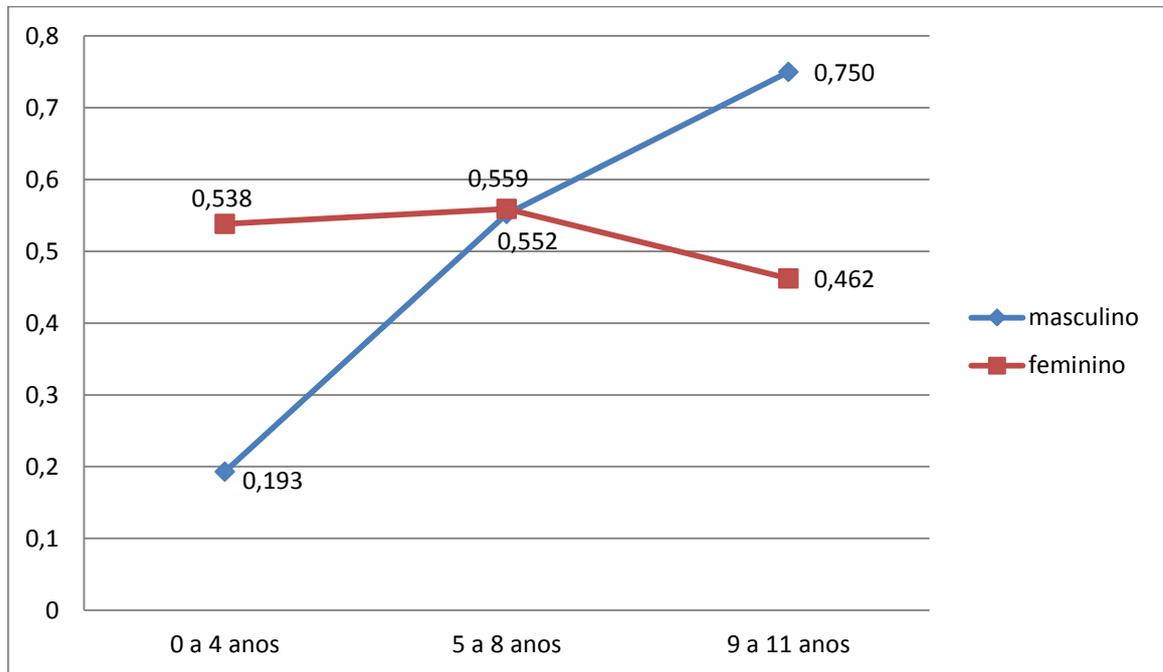
5.4.2.3 Posição do SN x escolaridade

Gráfico 9 - Atuação do cruzamento **posição do SN x escolaridade** sobre o verbo *existir*



O gráfico 9 nos revela um comportamento peculiar no que diz respeito à **posição do SN** e à **escolaridade**. Esses resultados foram atingidos no *step up* #10 (*convergence at interaction 5; input 0,020; log likelihood = -113,008; significance = 0.000*). Por meio deles, percebemos que a variável **escolaridade** apresenta uma correlação com a variável **posição do SN**, só que essa correlação opera de modo diferente nas duas estruturas. O fator **SN à esquerda do verbo**, favorecedor da aplicação de *existir*, continua favorecendo essa variante em todos os graus de escolarização, com índices muito próximos, mas com um padrão decrescente, ou seja, à medida que a escolaridade aumenta, o peso relativo desse fator decresce minimamente. O inverso ocorre com as estruturas com **SN à direita do verbo**, que apresenta crescimento de seus índices ao passo que a escolaridade aumenta. Em outras palavras, quanto maior a escolaridade do falante, mais ele usa *existir* na estrutura canônica **V + SN**, sendo que os falantes de maior escolaridade (**5 a 11 anos**) favorecem a variável em questão, enquanto que os falantes com menor escolaridade (**0 a 4 anos**) não favorecem essa variável.

5.4.2.4 Sexo x escolaridade

Gráfico 10 - Atuação do cruzamento **sexo x escolaridade** sobre o verbo *existir*

Este último cruzamento mostrou-se relevante para a variação *existir/ter* no *step up* #18 (*convergence at interaction* 6; *input* 0,021; *log likelihood* = -112,192; *significance* = 0.039). Como observamos neste gráfico, homens e mulheres apresentam comportamentos bem diferentes. Em relação aos homens, observamos que há uma correlação clara em relação ao grau de escolaridade no sentido de que quanto maior o grau de escolaridade maior o índice de uso de *existir*, sendo que os homens com baixa escolaridade (**0 a 4 anos**) não favorecem essa variante, ao passo que os homens com maior escolaridade (**5 a 11 anos**) favorecem-na. Já em relação às mulheres essa correlação não existe, pois da escolaridade **0 a 4 anos** para a de **5 a 8 anos** há um pequeno aumento no peso relativo e dessa última escolaridade para a de **9 a 11 anos** há uma queda um pouco maior. E enquanto que os homens com baixa escolaridade (**0 a 4 anos**) são os únicos que não favorecem *existir*, são as mulheres com maior escolaridade (**9 a 11 anos**) as únicas que não condicionam a aplicação dessa variante, apesar de apresentarem um peso relativo próximo à linha da neutralidade.

5.4.3 Considerações sobre as análises de *existir/ter*

Assim como os dados de *haver* na análise *haver/ter*, os de *existir* na análise *existir/ter* também apresentou pouca representatividade em nossa amostra. No entanto, essa variante também ainda encontra contextos favoráveis a sua aplicação, principalmente nas estruturas onde o SN aparece à esquerda do verbo, isto é, nas estruturas prototípicas. Esse foi o único fator relevante na aplicação de *existir*. O fator **SN à esquerda** se mostrou forte inclusive nos cruzamentos, onde as variáveis não selecionadas pelo programa como favorecedoras do fenômeno ganharam peso ao interagirem com esse fator. Foi o que aconteceu com ambos os sexos e os três níveis de escolaridade, como mostram os gráficos 8 e 9. Podemos inferir que as variáveis sociais parecem não atuarem diretamente na variação em estudo, pois só tornam-se influentes ao serem cruzadas com a variável **posição do SN**, variável linguística selecionada como única favorecedora do fenômeno, com exceção da variável **faixa etária**, que apesar de relevante na variação *haver/ter*, sequer foi forte nos cruzamentos.

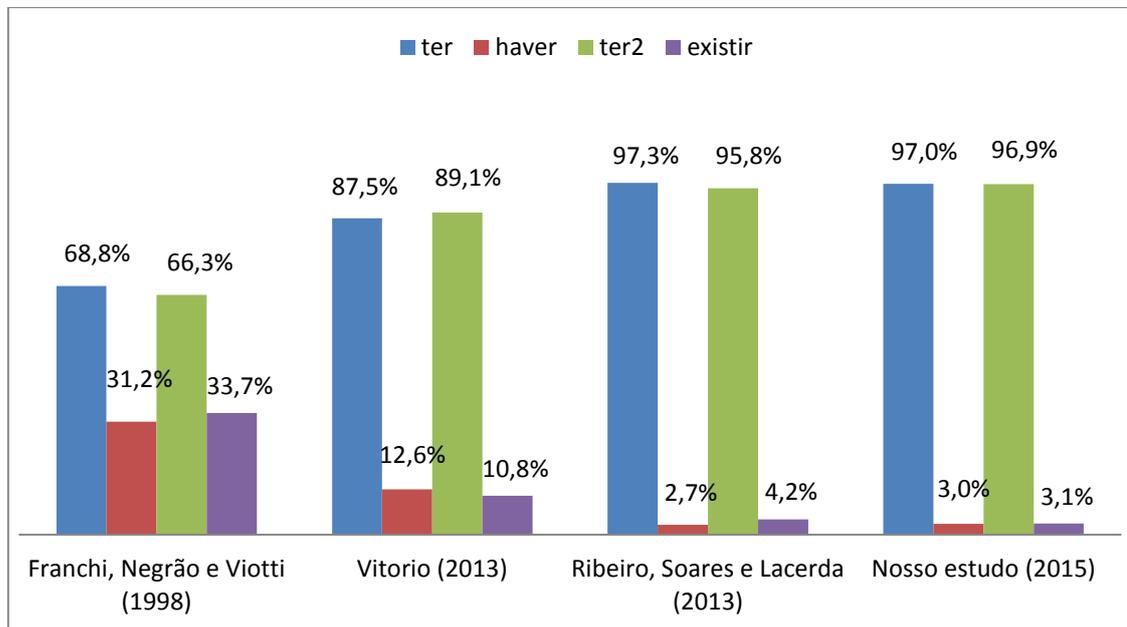
O **sexo** e a **escolaridade**, além de passarem a influenciar o fenômeno quando estão em interação com a **posição do SN**, também se mostraram fortes quando unidas, pois dos 6 fatores do cruzamento, 4 passaram a favorecer a variante *existir*, a saber: **homens e mulheres** de escolaridade intermediária (**5 a 8 anos**), homens com maior grau de escolarização (**9 a 11 anos**) e mulheres com menor grau de escolarização (**0 a 4 anos**).

5.5 ANÁLISES DE *HAVER/EXISTIR*

5.5.1 Primeira análise: visão geral da regra

Apesar dos poucos dados de que dispomos (apenas 55), decidimos analisar a variação *haver/existir*, porque, pelos resultados obtidos na rodada ternária, acreditamos que a grande disputa para ocupar o espaço nas construções existenciais no Português Brasileiro não está mais entre os verbos *haver* e *ter* nem entre *existir* e *ter*, mas sim entre os verbos *haver* e *existir*. É quase categórico o uso da variante *ter* em nossa amostra, quando contrastado com essas duas variantes. Isso é atestado tanto pela frequência de uso em nosso trabalho (97% contra *haver*; 96,9 contra *existir*) como também pela frequência de uso nos trabalhos de Franchi, Negrão e Viotti (1998), Ribeiro, Soares e Lacerda (2013) e Vitória (2013).

Gráfico 11 - Frequência de uso das análises *ter/haver* e *ter/existir* em três estudos e no nosso



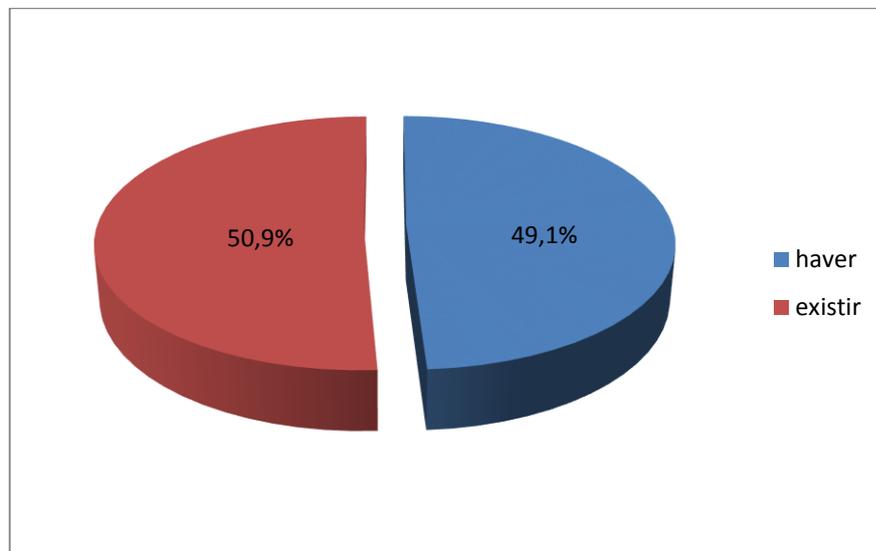
Fonte: dados de Franchi, Negrão e Viotti (1998), Ribeiro, Soares e Lacerda (2013), Vítório (2013) e nosso estudo (2015).

O gráfico 11, além de mostrar a supremacia de *ter* em relação a *haver* e *existir* em todos os estudos, também mostra uma proporção inversa em relação a esses três verbos. À medida que a frequência de uso do verbo *ter* aumenta consideravelmente de um estudo para outro, a frequência dos verbos *haver* e *existir* diminui. Também é possível observar que, mesmo apresentando índices bem menores, esses dois últimos verbos têm sempre frequências de uso muito próximas. Por isso, acreditamos que a grande questão no estudo da variação dos verbos existenciais não é mais o papel de *ter*, que, pelos dados que temos, já se consolidou como o verbo existencial mais utilizado no Português Brasileiro, mas sim os papéis desempenhados por *haver* e *existir* na disputa pela segunda preferência dos falantes, que apontam para dois caminhos: ou um deles vai ganhar a disputa para ocupar o segundo lugar no contexto existencial ou ambos vão perder essa função na língua. Claro que a resposta a essa questão só poderá ser dada após inúmeras outras pesquisas sobre o tema e comparações com dados de outras épocas futuras. Empreendimento muito complexo e prolongado, e por isso, não intencionado por esta dissertação.

O que podemos fornecer aqui, devido à já mencionada escassez de dados, são *insights* preliminares que podem ser confrontados com estudos futuros que analisem uma maior quantidade de dados.

Ao realizarmos a rodada inicial da variação *haver/existir* com os 55 dados (não houve nocautes) e percebemos que a distribuição deles é muito equilibrada, pois 27 são de *haver* e 28 são de *existir*. A representação percentual, exposta no gráfico 12, mostra uma diferença de apenas 0,9 pontos percentuais, reafirmando esse equilíbrio. Entretanto, faz-se necessário enfatizar que todos os resultados dessa análise, apresentados a seguir, servem apenas como indícios do comportamento desses dois verbos, haja vista o baixo número de dados analisados. É preciso, pois, ampliar a amostra, para que essa análise faça jus ao arcabouço teórico-metodológico da Sociolinguística Quantitativa, que como o próprio nome deixa claro, deve trabalhar com quantidades significativas de dados.

Gráfico 12 - Frequência de uso das variantes *haver* e *existir*



5.5.2 Segunda análise: visão mais detalhada da regra

Na análise multivariada, a melhor rodada *step up* foi a de número #12 (*convergence at interaction 12; input 0,513; log likelihood = -25,732; significance = 0.007*). Temos, a partir do *input*, a constatação do equilíbrio, já mencionado, que existe entre essas duas variantes. Nessa rodada, assim como houve com a análise *haver/ter*, a **forma verbal** foi a primeira variável considerada relevante para a variação *haver/existir*. A novidade nesta análise foi a presença da variável **sexo** (segunda selecionada em ordem de relevância), enquanto que na análise *haver/ter*, a variável social selecionada foi a **faixa etária**. Até então, o **sexo** só havia se mostrado relevante em cruzamentos.

5.5.2.1 Forma verbal

Tabela 7 - Atuação da variável **forma verbal** sobre o verbo *haver*

FATORES	Aplic./Total	%	P.R.
Infinitivo	9/10	90,0	0,947
Perfeito	6/7	85,7	0,934
Presente	10/30	33,3	0,279
Imperfeito	2/8	25,0	0,087

Como vemos na tabela 7, as formas verbais **infinitivo** (p.r. 0,947) e **perfeito** (p.r. 0,934), definitivamente, são os fatores mais condicionadores da variante *haver* em todo o nosso estudo, pois favorecem esse verbo tanto na análise com *ter* quanto na análise com *existir*, ambas com pesos relativos acima de 0,900. Dessa forma, já podemos afirmar que essas duas formas verbais são os principais fatores responsáveis por *haver* ainda se fazer presente na amostra analisada.

5.5.2.2 Sexo

Tabela 8 - Atuação da variável **sexo** sobre o verbo *haver*

FATORES	Aplic./Total	%	P.R.
Feminino	14/26	53,8	0,769
Masculino	13/29	44,8	0,254

É sabido que homens e mulheres são diferentes em inúmeros aspectos: aparência, papéis sociais, modo de pensar e agir em situações diversas, entre outras coisas. Portanto, é de se esperar que ambos apresentem comportamentos linguísticos distintos. Os estudos variacionistas em geral têm mostrado que esse princípio é seguido pelas variantes estudadas. Como vemos na tabela 8, os resultados dessa pequena amostra também refletem essa diferença de comportamento linguístico, uma vez que o sexo **feminino** (p.r. 0,769) condiciona a variante *haver* e o sexo **masculino** se revela uma restrição ao uso dessa variável, pois favorece a variante *existir* (p.r. 0,746). O fato dos pesos relativos dos sexos que favorecem as variantes opostas serem muito próximos indica uma divisão bem nítida na aplicação das variantes *haver/existir* por parte dos homens e das mulheres.

Não houve nenhum cruzamento para ser analisado na variação *haver/existir*, pois os únicos que não apresentaram nocautes, não foram selecionados pelo programa como relevantes nesta análise.

5.5.3 Considerações sobre as análises de *haver/existir*

Esta análise, sem dúvida, é a mais equilibrada de todas as realizadas nesta pesquisa, apesar de não nos dar informações confiáveis acerca do fenômeno pelos poucos dados encontrados na amostra. Mesmo assim, esses resultados são importantes por seu pioneirismo, pois não encontramos nenhum trabalho variacionista que faça essa análise que realizamos. Dessa forma, podemos presenciar uma mudança de foco nos estudos sobre os verbos existenciais *haver*, *existir* e *ter*, com aprofundamento maior na análise dos dois primeiros.

Feita essa ressalva, percebemos que dois dos três fatores que atuam no favorecimento de *haver* na análise *haver/ter* também favorecem essa variante na variação *haver/existir*. São eles: **infinitivo** e **perfeito**. Isso reforça a importância dessas duas formas verbais na manutenção de *haver* no falar popular. A diferença da análise *haver/ter* para a *haver/existir* é que, enquanto, na primeira, a **faixa etária** foi a variável social selecionada pelo programa, na segunda o **sexo** revelou-se relevante. As **mulheres** são as grandes responsáveis pelo uso de *haver* em contraste com *existir*, favorecida pelos **homens**.

Abaixo, segue um quadro com o resumo dos resultados das três análises que realizamos para uma visualização mais global da variação *haver/existir/ter*.

Quadro 6 - Síntese dos resultados das três análises

ANÁLISES	FATORES FAVORECEDORES DE <i>H AVER</i>
<i>haver/ter</i>	Análise binomial
	1. Infinitivo 2. Perfeito 3. Falantes a partir de 50 anos
	Cruzamento
	4. Homens e mulheres a partir de 50 anos
<i>existir/ter</i>	Análise binomial
	1. SN à esquerda do verbo
	Cruzamentos 2. SN à esquerda do verbo realizado por ambos os sexos 3. SN à esquerda do verbo realizado por falantes de todas as escolaridades

	<ol style="list-style-type: none"> 4. SN à direita do verbo realizado por homens 5. SN à direita do verbo realizado por falantes de 5 a 11 anos de escolaridade 6. Homens de 5 a 11 anos de escolaridade 7. Mulheres de 0 a 8 anos de escolaridade.
<i>haver/existir</i>	Análise binomial
	<ol style="list-style-type: none"> 1. Infinitivo 2. Perfeito 3. Mulheres

5.6 ANÁLISE DO TEMPO APARENTE

Os estudos em tempo aparente de variação e mudança envolvem agrupar falantes de diferentes idades e comparar a frequência de uma variante na fala das sucessivas gerações. O tempo aparente é um modo de simular e formular mudança em tempo real, usando dados sincrônicos, quando os *corpora* diacrônicos que discutimos na seção anterior não estão disponíveis para os pesquisadores ou quando estes não têm o tempo ou o dinheiro para construir seu próprio *corpus* de tempo real⁸. (MEYERHOFF, 2006, p. 133, tradução nossa)

Essa definição deixa claro que o estudo em tempo aparente é uma simulação do tempo real. Acredita-se que a estrutura básica da língua adquirida pelo falante quando ele é jovem permanece durante as outras fases da vida de modo que a fala de um idoso atual pode fornecer pistas de como a língua era há uma ou duas gerações. Dessa forma, quando a frequência de uso de uma determinada variante decresce à medida que a idade do falante aumenta, podemos estar diante de uma mudança em curso. Caso essa relação não seja refletida pelos dados, podemos estar diante de uma variação estável.

Labov (1994, p. 72, tradução nossa), após a demonstração de vários dados de seus estudos, entende que “as inferências sobre mudança em progresso a partir do tempo aparente não podem ser desprezadas e que este tipo de análise pode ser realizada com proveito quando não há dados de tempo real disponíveis⁹”. É importante destacar que, para se chegar a conclusões mais precisas, é necessário investigar a faixa etária em relação a outras variáveis

⁸ Texto original: Apparent time studies of variation and change involve sampling speakers of different ages and comparing the frequency of a variant in the speech of successive generations. Apparent time is a way of simulating and modelling real time change using synchronic data, when the diachronic corpora we discussed in the previous section are not available to researchers or when researchers do not have the time or money to construct their own real time corpus.

⁹ Texto original: It appears that the inferences to be drawn about change in progress from apparent time are not negligible, and that this type of analysis can be pursued profitably when no real-time data are available.

sociais e que os resultados encontrados em relação ao tempo aparente devem ser oportunamente contrastados com dados de tempo real.

Além de definir o estudo em tempo aparente, Meyerhoff (2006) destaca vários estudos que fazem esse tipo de análise. Alguns estudos apontados pela autora utilizam a frequência de uso das variantes enquanto outros utilizam os pesos relativos. Labov (1994), na análise de seus dados, utiliza tanto frequência de uso quanto índices estatísticos. Silva (2001) e Vitória (2012), para fazerem inferências sobre o status da variação nas comunidades estudadas, utilizam em algumas análises os dois tipos de informações estatísticas: a frequência e o peso relativo. Nesta seção, apresentaremos a análise das faixas etárias somente para a análise *haver/ter* porque foi a única cuja **faixa etária** foi selecionada pelo programa como relevante.

Para o estudo do tempo aparente, utilizaremos os dados da variante *ter* porque essa é a variante considerada inovadora na literatura, uma vez que, no decorrer da história, foi ocupando espaço nas orações existenciais e aumentando sua frequência de uso em relação aos verbos *haver* e *existir*. Portanto, essas últimas são consideradas variantes mais conservadoras. Tarallo (2003, p. 65) nos informa que se “o uso da variante mais inovadora for mais frequente entre os jovens, decrescendo em relação à idade dos outros informantes, você terá presenciado uma situação de mudança em progresso”.

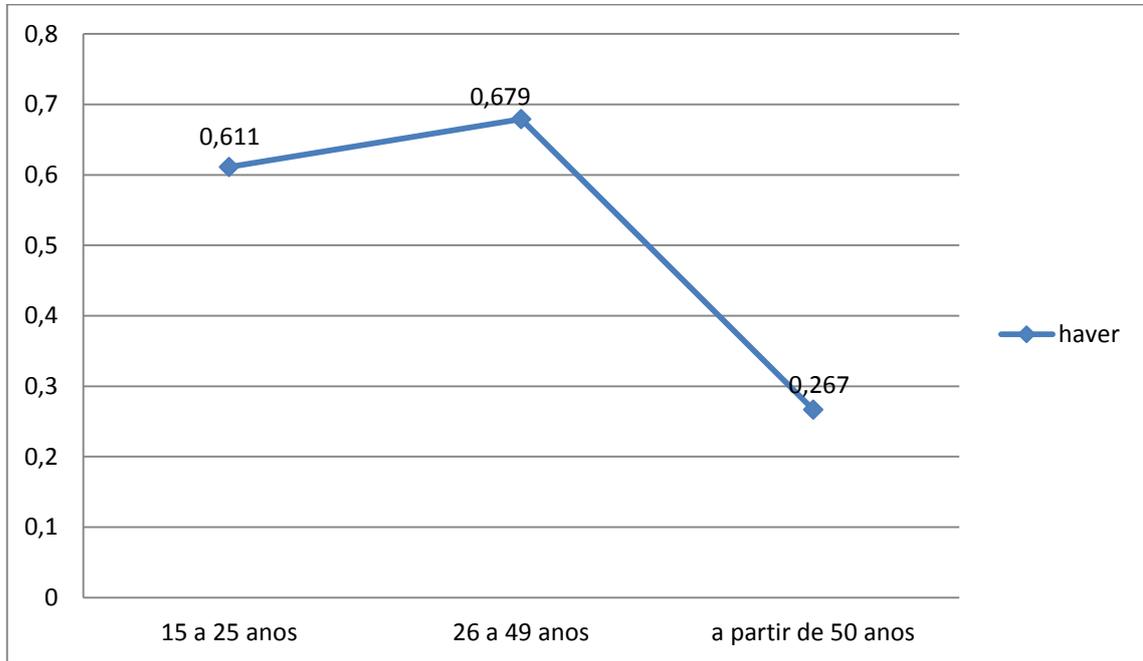
5.6.1 O tempo aparente na variação *haver/ter*

O gráfico 13 a seguir mostra que há um aumento discreto de peso relativo da faixa jovem para a faixa adulta no que diz respeito ao favorecimento da variante inovadora *ter*. No entanto, esse valor diminui drasticamente da faixa adulta para a idosa, que não favorece essa variante. Apesar de não termos um padrão perfeito de mudança em curso nos termos de Tarallo (2003), isto é, o peso relativo de *ter* decrescendo da faixa jovem para a faixa adulta e desta para a idosa, podemos inferir que é possível estarmos diante de um fenômeno em mudança.

Se observarmos os dados, percebemos que as faixas etárias dos jovens e dos adultos apresentam pesos relativos muito próximos e que ambas favorecem a variante inovadora *ter*. Dessa forma, o grande contraste está entre essas duas faixas etárias e a faixa dos idosos, pois o peso relativo desta é bem menor, favorecendo a variante conservadora *haver*. Podemos inferir ainda a partir desta análise que a faixa dos adultos é a grande restrição,

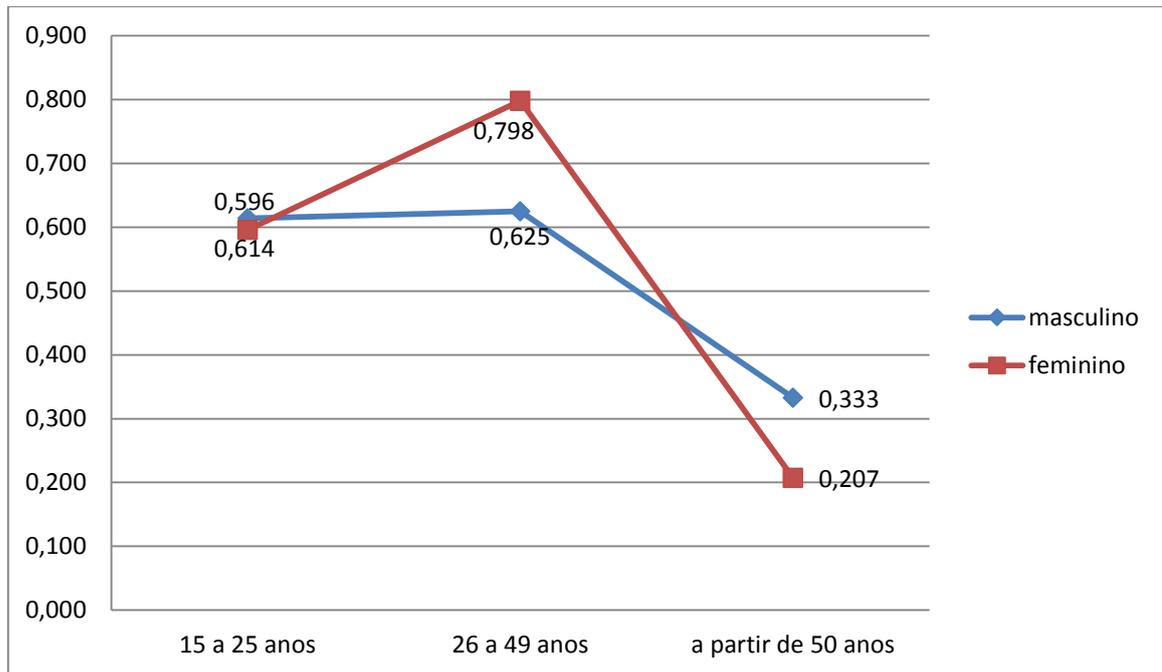
para que essa mudança não esteja mais avançada, porém, não podemos dizer por que isso acontece.

Gráfico 13 - Atuação da variável **faixa etária** sobre *ter*



Ao observarmos o padrão do gráfico 14, resultado do cruzamento das variáveis **sexo x faixa etária**, podemos afirmar que ambos os sexos se comportam de modo semelhante e, em coerência com os dados do gráfico 13: tanto as mulheres quanto os homens jovens e adultos favorecem *ter* enquanto que ambos os sexos dos idosos favorece *haver*. Além disso, há também um aumento de peso relativo entre as duas primeiras faixas etárias e uma queda significativa entre as duas últimas faixas etárias. Partindo desses padrões gerais, podemos afirmar que tanto os homens quanto as mulheres apresentam comportamento semelhante. No entanto, percebemos que existe uma diferença entre os dois sexos em relação tanto ao aumento do peso relativo da faixa jovem para a faixa adulta quanto à queda desse valor da faixa adulta para a idosa: esses valores são mais acentuados nas mulheres, mostrando, assim, que as mulheres são as responsáveis pela grande restrição da faixa etária intermediária para o avanço da mudança a que nos referimos no gráfico 13.

Gráfico 14 - Atuação do cruzamento **faixa etária x sexo** sobre a variante *ter*



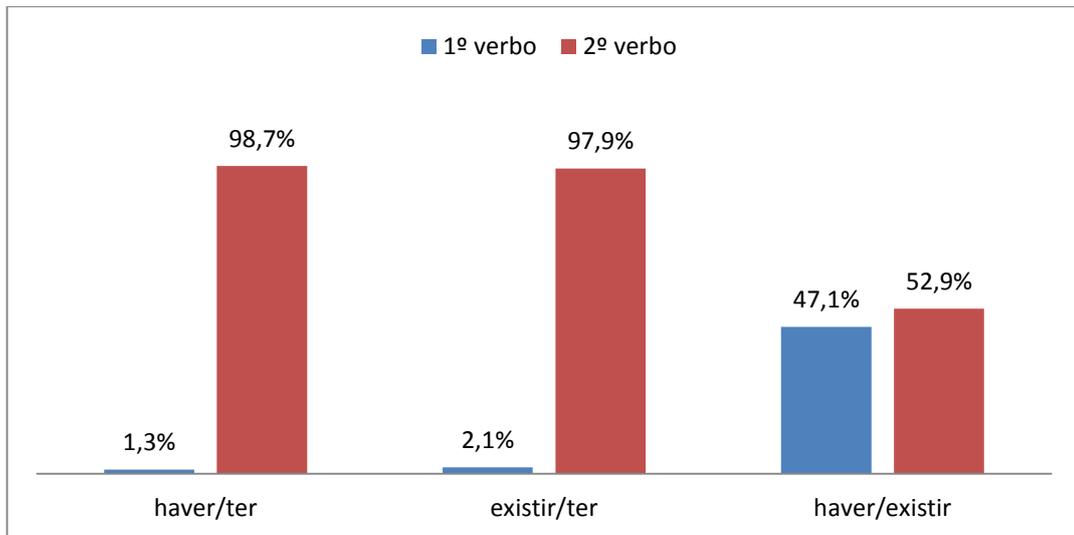
Existe uma grande discussão sobre o papel das mulheres em relação à mudança linguística. Alguns estudos têm mostrado que as mulheres são as principais responsáveis por alguns processos de mudança em relação às formas inovadoras. Outros têm mostrado que elas utilizam mais as formas prestigiadas do que os homens. No entanto, Labov (2008 [1972], p. 348), afirma que a “generalização correta [...] não é a de que as mulheres lideram a mudança linguística, mas sim que a diferenciação sexual da fala frequentemente desempenha um papel importante no mecanismo da evolução linguística”. Podemos afirmar então que, na análise *haver/ter*, a diferença de sexo não atua de forma decisiva no fenômeno, uma vez que essa variável não foi selecionada como relevante. Entretanto, ao ser cruzada com a faixa etária, essa variável revela um aspecto significativo: os homens estão mais avançados nessa possível mudança em curso do que as mulheres, pois apresentam um padrão gráfico mais próximo daquilo que nos informa Tarallo (2003, p. 65).

6 CONCLUSÃO

O objetivo geral desta pesquisa foi investigar qual o lugar dos verbos *haver* e *existir* nas construções existenciais em contraste com a variante inovadora *ter*. Nossa hipótese é de que a supremacia do verbo *ter* em frequência de uso em nossa amostra apresentará índices muito altos uma vez que trabalhamos com o falar popular. Os trabalhos que abrangem o falar culto (MARTINS; CALLOU, 2003; BATISTA, 2012) já mostram altas frequências de uso dessa variante. Para atingir nosso objetivo geral, traçamos os seguintes objetivos específicos: a) Identificar os fatores sociais e/ou linguísticos responsáveis pela manutenção de *haver* e *existir* no falar popular de Fortaleza; b) Investigar, a partir de análises no tempo aparente (faixa etária), se há indícios de esses verbos estarem passando por um processo de mudança em curso no sentido de serem substituídos pelo verbo *ter*.

Inicialmente, nossos dados mostraram que o lugar dos verbos *haver* e *existir* nas construções existenciais do falar popular de Fortaleza é muito restrito devido à frequências de uso desses verbos em relação a *ter*: *haver* (2,9%) e *existir* (3,0%). Depois, essa inferência foi reforçada pelas análises binárias dessas variantes, pois essa porcentagem não mudou muito. Na análise *haver/ter*, *haver* ficou com 3,0% das ocorrências, enquanto na análise *existir/ter*, *existir* ficou com 3,1%. Em outras palavras, não houve mudança percentual considerável entre a rodada ternária e as rodadas binárias. Esses dois *insights* iniciais foram confirmados pelos *inputs* das rodadas que forneceram uma visão geral de cada análise. Como já é sabido, o *input* de uma rodada *step up* é a probabilidade da variante designada como aplicação ocorrer quando os fatores selecionados como relevantes interagem, ou seja, a probabilidade de ocorrência dessa variante no contexto que lhe é mais favorável. Com base nos *inputs* das rodadas gerais, dispusemos, no gráfico 15, a probabilidade de cada verbo ocorrer nas três análises.

Gráfico 15 - Probabilidade de ocorrência das variantes nas três análises realizadas



Observamos, a partir desses resultados, que a variante *ter* está quase totalmente implementada nas construções existenciais de nossa amostra, ocupando os espaços de *haver* e *existir*. Estatisticamente, poderíamos dizer que essa batalha, como metaforicamente Tarallo (2003) denomina a existência de duas ou mais variantes, já foi vencida por *ter*, uma vez que a probabilidade desses dois verbos ocorrerem, nos contextos que lhes são mais favoráveis, é muito baixa. No entanto, como revelaram as análises *haver/ter* e *existir/ter*, as variantes conservadoras ainda são produzidas pelos falantes da amostra em estudo e fatores específicos ainda as favorecem, funcionando como restrição para a total implementação de *ter*. A partir desses achados, podemos tratar agora das duas interpretações possíveis para a questão proposta pelo título de nosso trabalho. Conseqüentemente, ao respondê-las, teremos atingido o objetivo geral e o primeiro objetivo específico desta pesquisa.

A primeira interpretação pode ser a seguinte: Tem chance de [o verbo] *haver* ainda existir no falar popular? E a resposta é sim. O verbo *haver* ainda encontra espaço nas estruturas existenciais, condicionado pelas formas verbais **infinitivo** e **pretérito perfeito** e pela faixa etária dos idosos, **falantes a partir de 50 anos**. A segunda interpretação possível para o nosso título é a seguinte: Tem chance de haver ainda [o verbo] *existir* no falar popular? E a resposta mais uma vez é afirmativa. O verbo *existir* leva uma pequena vantagem probabilística sobre o verbo *haver*, como mostra o gráfico 18, e ainda encontra espaço nas estruturas existenciais com um número maior de fatores condicionadores, a saber: sentenças com **SN à esquerda**, sentenças com **SN [-humano]**, sentenças com a estrutura **verbo no singular + SN no plural** e entre falantes mais escolarizados (**5 a 11 anos**).

Satisfeitos nosso objetivo geral e nosso primeiro objetivo específico, passemos ao segundo objetivo específico deste trabalho. A análise em tempo aparente do fenômeno variável *haver/existir/ter* não pode ser plenamente abarcada pelos nossos dados, pois, das três análises realizadas aqui, somente uma apresentou a variável faixa etária como relevante, a saber, *haver/ter*. Nessa análise, os dados revelaram que existem indícios de uma possível mudança em curso no português popular de Fortaleza, que aponta para a substituição de *haver* por *ter* em sentenças existenciais. É necessário reiterar que essa mudança trata-se apenas de uma inferência, possível através da análise do comportamento das variantes na divisão etária do *corpus*, e que, para chegarmos a conclusões mais precisas, precisaríamos de dados de tempo real.

Após termos cumprido com nossos objetivos de pesquisa, resolvemos investigar a variação *haver/existir* e um fato interessante nos foi revelado pelos poucos dados desses dois verbos: a grande disputa que parece ocorrer na sentenças existenciais não está entre o verbo *ter* e os verbos *haver* e *existir*, mas sim entre estes dois últimos. Dizemos “parece” porque, como informado antes, os dados dessas duas variantes não são estatisticamente significativos e podem ter, de alguma forma, desviado os resultados de nossa análise. Portanto, para uma confirmação dessa hipótese, é necessário coletarmos uma quantidade maior de dados desses dois verbos para checarmos se obteremos os mesmos resultados.

Eis aqui uma ideia de pesquisa futura que certamente faremos e esperamos que outros pesquisadores, utilizando *corpora* de outros falares, também empreendam essa investigação. Outra questão, diretamente ligada à variação *haver/existir*, que pode ser respondida em pesquisas futuras é a seguinte: as variantes *haver* e *existir* tendem a desaparecer totalmente ou a hipotética luta entre elas fará com que uma só desapareça e ocupe o segundo lugar na preferência dos falantes, ascendendo para uma nova batalha com *ter*?

Além desses questionamentos, outros podem surgir, se ultrapassadas as limitações deste trabalho, que, como qualquer outro sobre língua natural, está longe de ser completo. Uma dessas limitações diz respeito à análise de somente um tipo de inquérito, o D2: os inquéritos do tipo DID e EF poderiam nos ter fornecido uma quantidade maior de dados dos verbos *haver* e *existir* de modo que a variação entre eles pudesse ser melhor fundamentada. Além disso, a inclusão dessas outras entrevistas nos possibilitaria o controle de mais uma variável social, que seria o estilo, pois acreditamos que cada um desses três tipos de entrevistas apresentam graus de monitoramento de linguagem distintos.

Para concluir, podemos afirmar que, apesar das restrições desta pesquisa, nossos resultados fornecem informações importantes acerca do falar popular de Fortaleza. Apesar de, na maior parte do trabalho, nos referirmos sempre a nossa amostra, acreditamos que ela seja representativa do universo maior que é o falar popular dos fortalezenses. No entanto, é importante destacar que nosso tema não representa uma marca específica da fala de Fortaleza, como o enfraquecimento da fricativa /v/, estudado por Rodrigues (2013, p. 150), que, segundo a autora, é “uma das peculiaridades do falar fortalezense, que enriquece os elementos sócio-culturais que caracterizam esta comunidade”. Guimarães (2014, p. 218) também destacou as formas nominais de tratamento “*macho, rapaz e mulher*” como “variantes muito frequentes no falar de Fortaleza”. Nosso tema parece ser mais geral do que específico, indicando, talvez, uma tendência da expansão de *ter* não só no português popular de Fortaleza, mas também no português tanto culto quanto popular falado em outras capitais do país, como mostraram os estudos descritos na seção 2.

REFERÊNCIAS

- ALKMIM, T. M. Sociolinguística: parte I. In: MUSSALIM, F.; BENTES, A. C. (orgs.) **Introdução à linguística: domínios e fronteiras**. v. 1. 8. ed. São Paulo: Cortez, 2008. cap. 1, p. 21-47.
- ALMEIDA, Erica Sousa de; CALLOU, Dinah. Estruturas com ter e haver em textos jornalísticos: do século XIX ao XX. In: ENCONTRO DO CÍRCULO DE ESTUDOS LINGUÍSTICOS DO SUL , 5., 2002, Curitiba,PR. **Anais Eletrônicos...** Curitiba: Mídia Curitibana, 2003. p. 510-515. Disponível em: <<http://www.celsul.org.br/Encontros/05/pdf/069.pdf>>. Acesso em: 15 maio 2014.
- ARAÚJO, Alves de Araújo. **As vogais médias pretônicas no falar popular de Fortaleza: uma abordagem variacionista**. 2007. 152 f. Tese (Doutorado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade Federal do Ceará, Fortaleza, 2007. Disponível em: < http://www.repositorio.ufc.br/bitstream/riufc/3597/1/2007_tese_ALAraujo.pdf >. Acesso em: 23 jul. 2015.
- _____. O projeto Norma Oral do Português Popular de Fortaleza – NORPOFOR. In: CONGRESSO NACIONAL DE LINGUÍSTICA E FILOLOGIA, 15., 2011, Rio de Janeiro. **Cadernos do CNLF**. Rio de Janeiro: CiFEFiL, 2011, v. XV, n. 5, t. 1. p. 835-845. Disponível em: < www.filologia.org.br/xv_cnlf/tomo_1/72.pdf >. Acesso em: 10 jul. 2014
- ARGOTE, C. de. **Regras da Língua Portuguesa: espelho da Latina**. Lisboa Occidental: Officina da Musica, 1725. Disponível em: <http://purl.pt/10/4/1-601-p_PDF/1-601-p_PDF_24-C-R0150/1-601-p_0000_1-398_t24-C-R0150.pdf>. Acesso em: 10 maio 2015
- AVELAR, Juanito. Gramática, competição e padrões de variação: casos com ter/haver e de/em no português brasileiro. **Revista de Estudos da Linguagem**, Belo Horizonte, v. 14, n. 2, p. 99-143, dez. 2006. Disponível em: <<http://periodicos.letras.ufmg.br/index.php/-relin/article/view/2425/2379>>. Acesso em: 28 out. 2014.
- AZEREDO, J. C. de. **Gramática Houaiss da Língua Portuguesa**. 3. ed. São Paulo: Publifolha, 2010.
- BAGNO, M. **Gramática Pedagógica do Português Brasileiro**. São Paulo: Parábola Editorial, 2011.
- BARBOSA, J. S. **Grammatica philosophica da lingua portuguesa**. Lisboa: Typographia Academia das Sciencias, 1882.
- BATISTA, Priscila Guimarães. **Ter e Haver existenciais na fala culta de Rio de Janeiro Salvador e Porto Alegre: do social ao linguístico**. 2012. 70 f. Dissertação (Mestrado em Letras Vernáculas) – Programa de Pós-Graduação em Letras Vernáculas, Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2012. Disponível em: < <http://www.letras.ufrj.br/posverna/mestrado/BatistaPG.pdf> >. Acesso em: 12 ago. 2014

BECHARA, E. **Moderna Gramática Portuguesa**. 37. ed. Rio de Janeiro: Nova Fronteira, 2009.

BORTONI-RICARDO, S. M. **Nós cheguem na escola, e agora?**: Sociolinguística na sala de aula. São Paulo: Parábola Editorial, 2005.

BRIGHT, W. As dimensões da sociolinguística. Tradução de Elizabeth Neffa Araújo Jorge. In: **Sociolinguistics**. In: PROCEEDING OF THE UCLA SOCIOLINGUISTICS CONFERENCE, 1964. 3. ed. Mouton: The Hague, 1966.

BRUNO, A.; FARIAS, A. de. **FORTALEZA**: 285 anos. p. 01-11, 2011. Disponível em: <www.arturbruno.com.br/images/conteudo/file/cartilhaHFortaleza.pdf>. Acesso em: 17 abr. 2015.

CALLOU, D.; AVELAR, J. Preservação e mudança na história do português: de possessivo e mudança. **Revista Matruga**, Rio de Janeiro, v. 19, n. 30, p. 224-235, jan./jun. 2012. Disponível em: <<http://www.-pgletras.uerj.br/matruga/matruga30/arqs/matruga30a11.pdf>>. Acesso em 09 ago. 2014.

CALVET, L. **Sociolingüística**: uma introdução crítica. Tradução de Marcos Marcionilo. São Paulo: Parábola, 2002.

CASTILHO, A. T. de. **Nova Gramática do Português Brasileiro**. São Paulo: Contexto, 2012.

CHOMSKY, N. **Syntactic Structures**. The Hague/Paris: Mouton, 1957.

COSTA, Alessandra de Azevedo *et al.* Verbos Existenciais: ter/haver. **Revista Virtual de Estudos da Linguagem**, v. 9, n. 17, p. 367-381, 2011. Disponível em: <http://www.revel.inf.br/files/artigos/revel_17_verbos_existenciais.pdf>. Acesso em: 02 fev. 2014.

CUNHA, C.; CINTRA, L. F. L. **Nova gramática do português contemporâneo**. 5. ed. Rio de Janeiro: Lexikon, 2008.

DUTRA, C. **Ter e haver na norma culta de Salvador**. 2000. Dissertação (Mestrado em Letras) – Universidade Federal da Bahia, Salvador, 2000.

ECKERT, P.; MCCONNELL-GINET, S. **Language and gender**. New York: Cambridge University Press, 2003.

FRANCHI, C.; NEGRÃO, E.; VIOTTI, E. Sobre a gramática das orações impessoais com ter/haver. **Revista DELTA.**, São Paulo, vol.14, n. especial, p. 105-131, 1998. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-44501998000300009>. Acesso em: 03 mar. 2015.

FREIRE, F. J. **Reflexões sobre a Língua Portuguesa**. Lisboa: Typographia da Sociedade Propagadora dos Conhecimentos Uteis, 1842.

GUIMARÃES, T. de A. A. S. **Tu é doido, macho!**: a variação das formas de tratamento no falar de fortaleza. 2014. 237 f. Dissertação (Mestrado em Linguística Aplicada) – Programa

de Pós-Graduação em Linguística Aplicada, Universidade Estadual do Ceará, Fortaleza, 2014. Disponível em: <http://www.uece.br/posla/dmdocuments/DISSERTACAO_TATIANE.pdf>. Acesso em 02 jan. 2014.

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

LABOV, W. **The Social Motivation of a Sound Change**. *Word* 19. p. 273-309, 1963.

_____. **The Social Stratification of English in New York City**. 2nd ed. New York: Cambridge University Press, 2006 [1966].

_____. **A Study of Nonstandard English**. Washington, DC: National Council of Teachers of English, 1969.

_____. **Where does the linguistic variable stop?: A response to Beatriz Lavandera**. Sociolinguistics Working Paper. Texas, n. 44, p. 1-16, 1978.

_____. **Principles of linguistic change: internal factors**. Oxford: Backwell Publishers, 1994.

_____. **Principios del cambio lingüístico: factores sociales**. Tradução de Pedro M. Butragueño. Madrid: Editorial Gredos, 2006 [2001].

_____. **Padrões Sociolinguísticos**. Tradução de Marcos Bagno, Maria M. P. Scherre, Caroline R. Cardoso. São Paulo: Parábola Editorial, 2008 [1972]. 389 p. Título original: Sociolinguistic Patterns.

_____. **Principles of linguistic change: cognitive and cultural factors**. Oxford: Backwell Publishers, 2010.

LAVANDERA, B. **Where does the linguistic variable stop?** *Language in Society*, n. 7, p. 171-182, 1978.

LOBATO, A. J. dos R. **Arte da Grammatica da Lingua Portugueza**. Lisboa: Regia Officina Typografica, 1770.

MARTELOTTA, M. E. **Mudança Linguística: uma abordagem baseada no uso**. v. 1. São Paulo: Cortez, 2011.

MARTINS, Luciene; CALLOU, Dinah. Mudança em tempo aparente e em tempo real: construções ter/haver existenciais. In: ENCONTRO DO CÍRCULO DE ESTUDOS LINGÜÍSTICOS DO SUL, 5., 2002, Curitiba-PR. **Anais Eletrônicos...** Curitiba: Mídia Curitibana, 2003. p. 820-825. Disponível em:< <http://celsul.org.br/Encontros/05/pdf/114.pdf>>. Acesso em: 12 set. 2013.

MATTOS e SILVA, R. V.; MACHADO FILHO, A. V. L. Variação ter/haver. In: LOBO, Tânia; OLIVEIRA, Klebson (Orgs.). **África à vista: dez estudos sobre o português escrito por africanos no Brasil do século XIX**. Salvador: EDUFBA, 2009. cap. 10. p. 338-351. Disponível em: <<http://static.scielo.org/scielobooks/48/pdf/lobo-9788523208882.pdf>>. Acesso em: 05 jan. 2014.

MEYERHOFF, M. **Introducing Sociolinguistics**. New York: Routledge, 2006.

MOLLICA, M. C. Relevância das variáveis não linguísticas. In: MOLLICA, M. C.; BRAGA, M. L (orgs.). **Introdução à Sociolinguística: o tratamento da variação**. 4. ed. São Paulo: Contexto, 2012. p. 27-31.

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

OLIVEIRA, Solange Mendes. A variação das formas verbais *ter* e *haver* em textos escritos do século XIX. In: ENCONTRO DO CÍRCULO DE ESTUDOS LINGUÍSTICOS DO SUL, 9., 2010, Palhoça-SC. **Anais Eletrônicos...** Palhoça, 2011, p. 01-11. Disponível em: <<http://www.celsul.org.br/Encontros/09/artigos/Solange%20Oliveira.pdf>>. Acesso em: 18 set. 2014.

PERINI, M. A. **Gramática descritiva do português**. São Paulo: Ática, 1996.

_____. **Estudos de gramática descritiva: as valências verbais**. São Paulo: Parábola Editorial, 2008.

_____. **Gramática do português brasileiro**. São Paulo: Parábola Editorial, 2010.

RAMOS, Conceição de Maria de Araújo; SANTOS, Wendel Silva dos. Estruturas existenciais: a variação *ter/haver* no corpus constituído para ALiMA. **Cadernos de Pesquisa**, São Luís, v. 19, n. especial, jul. 2012. Disponível em: <<http://www.periodicoeletronicos.ufma.br/index.php/cadernosdepesquisa/article/viewFile/1069/757>>. Acesso em: 22 maio 2014.

RIBEIRO, P. R. O.; SOARES, M. S.; LACERDA, P. F. A. C. A realização da noção de existência no “mineirês”: um estudo da variação dos verbos *ter*, *haver* e *existir*. **Revista Signótica**, Goiânia, vol. 25, n. 2, p. 535-561, 2013. Disponível em: <http://www.revistas.ufg.br/index.php/sig/article/view/19192/15795> . Acesso em: 14 nov. 2013.

RODRIGUES, Ana Germana Pontes. **Râmo rê se rai dá certo: o enfraquecimento da fricativa /v/ no falar de Fortaleza**. 2013. 170 f. Dissertação (Mestrado em Linguística Aplicada) – Programa de Pós-Graduação em Linguística Aplicada, Universidade Estadual do Ceará, Fortaleza, 2013. Disponível em: <<http://www.uece.br/posla/dmdocuments/Anagermanapontesrodrigues>>. Acesso em: 26 dez. 2014.

SAUSSURE, F. de. **Curso de Linguística Geral**. Tradução de Antônio Chelini, José Paulo Paes e Izidoro Blikstein. 27. ed. São Paulo: Cultrix, 2006 [1916]. Título original: Cours de linguistique générale.

SILVA, R. N. A. da. **Variação *ter/haver* na fala pessoense**. 2001. 106 f. Dissertação (Mestrado em Letras) – Universidade Federal da Paraíba, João Pessoa, 2001.

TARALLO, F. A **Pesquisa Sociolinguística**. São Paulo: Ática, 2003.

VITÓRIO, E. Ter/haver existenciais na escrita de alunos de 5ª a 6ª séries do Ensino Fundamental da cidade de Maracanaú/CE. **Revista Virtual de Estudos da Linguagem**, v. 5, n. 9, p. 01-17, ago. 2007. Disponível em: < http://www.revel.inf.br/files/artigos/revel_9_ter_haver_existenciais.pdf >. Acesso em: 08 maio 2015.

_____. **Ter/haver existenciais na escrita de alunos dos ensinos Fundamental e Médio da cidade de Maceió/AL**. 2008. 120f. Dissertação (Mestrado em Linguística) – Universidade Federal de Alagoas, 2008.

_____. Aquisição e variação dos verbos ter e haver existenciais no PB. **Veredas online** – aemática, p.53-63 – PPG linguística/UFJF, Juiz de Fora, 2010.

_____. A alternância ter/haver existenciais na fala maceioense. **Revista Interdisciplinar**, ano VI, v. 14, 2011, p. 77-85.

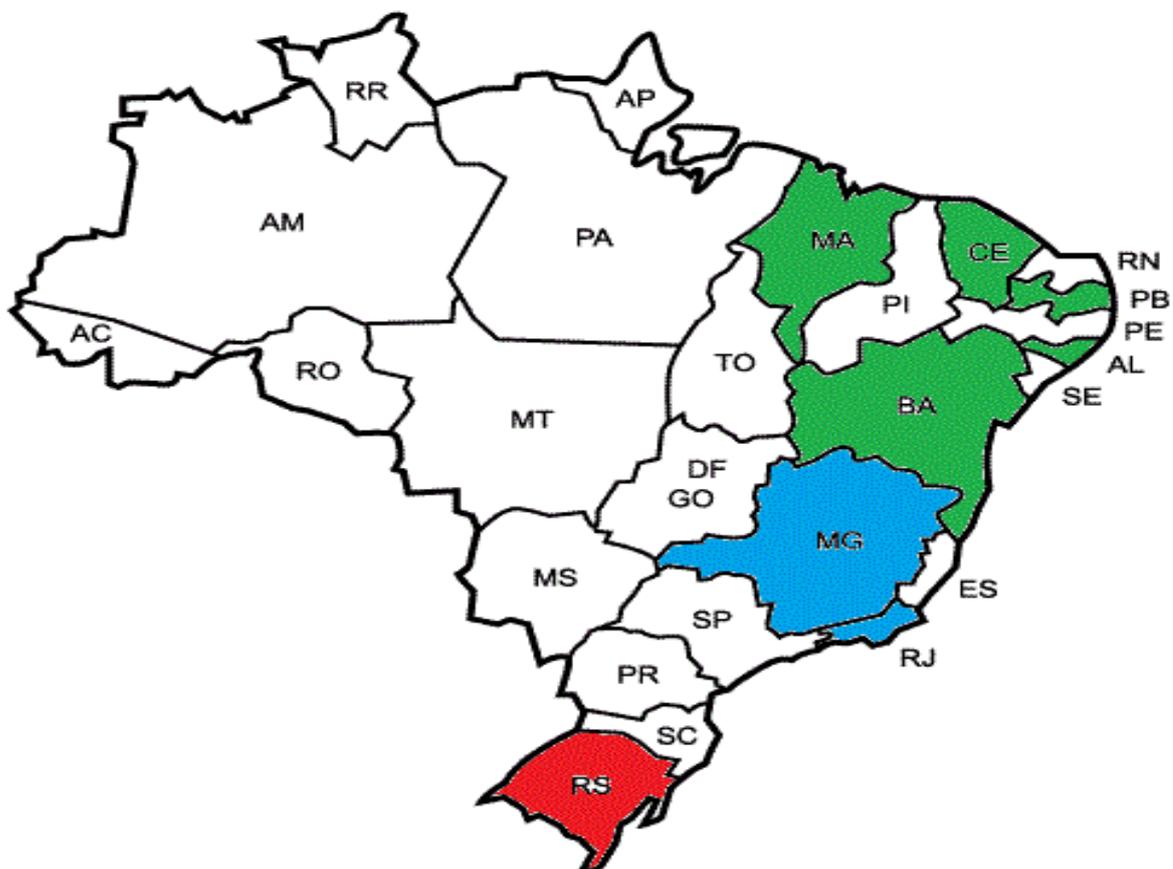
_____. **Ter/haver existenciais na fala alagoana**: variação estável ou mudança em progresso. 2012. 152f. Tese (Doutorado em Linguística) – Universidade Federal de Alagoas. 2012.

_____. **Construções existenciais com os verbos ter e haver na fala e na escrita**: uma análise comparativa. 2013. 26f. (Pós-doutorado Júnior em Letras Vernáculas) – Universidade Federal do Rio de Janeiro. 26f. 2013.

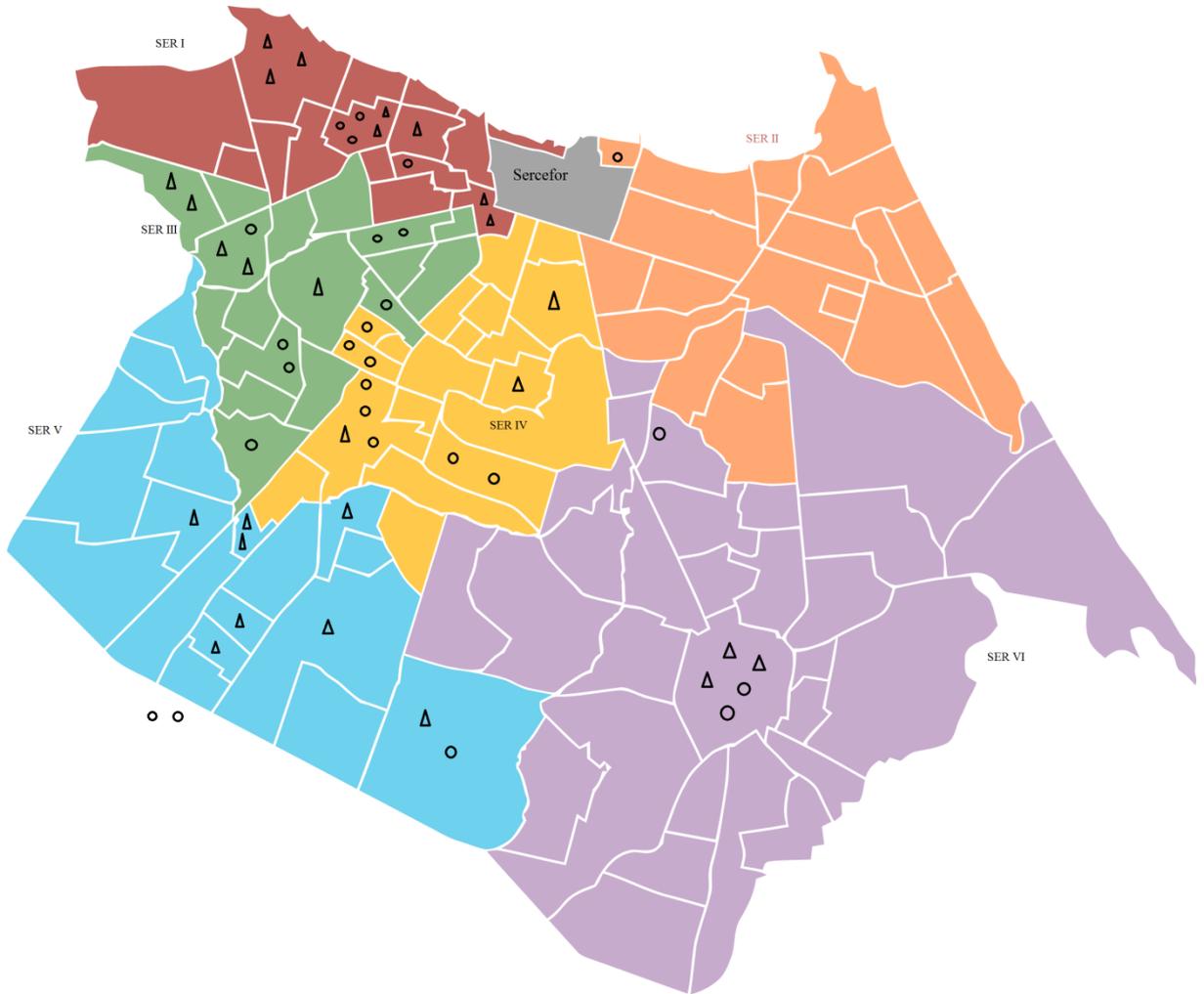
VOTRE, S. J. Relevância da variável escolaridade. In: MOLLICA, M. C.; BRAGA, M. L (orgs.). **Introdução à Sociolinguística**: o tratamento da variação. 4. ed. São Paulo: Contexto, 2012.

WEINREICH, U.; LABOV, W.; HERZOG, M. Empirical foundations for a theory of language change. In: Lehmann, W.; Malkiel, Y. **Directions for historical linguistics**. Austin: University of Texas Press, 1968. Disponível em: <<http://www.utexas.edu/cola/centers/lrc/books/hist05.html>> Acesso em 20 jul. 2014.

APÊNDICES

APÊNDICE A – Mapa dos estudos variacionistas sobre *haver/ter* e sobre *haver/ter/existir*

APÊNDICE B – Distribuição dos falantes por bairros (os triângulos representam os homens e os círculos representam as mulheres).



APÊNDICE C - Chave de codificação

VARIÁVEL DEPENDENTE		CÓDIGOS
	haver	h
	existir	e
	ter	t
VARIÁVEIS INDEPENDENTES	FATORES	CÓDIGOS
Sexo		
	masculino	m
	feminino	f
Faixa etária		
	15 a 25 anos	1
	26 a 49 anos	2
	a partir de 50 anos	3
Escolaridade		
	0 a 4 anos	0
	5 a 8 anos	5
	9 a 11 anos	9
Animacidade do SN		
	[+humano]	+
	[-humano]	-
Presença x ausência da variante na fala do interlocutor		
	presença	ç
	ausência	o
Presença x ausência de elementos à esquerda do verbo		
	presença	*
	ausência	#
Posição do SN em relação ao verbo		
	SN à direita do verbo	d
	SN à esquerda do verbo	b
Peso do SN		

SN simples	z
SN complexo	x
SN nulo	o
Núcleo do SN nulo	n
Forma verbal	
presente do indicativo	r
presente do subjuntivo	j
pretérito perfeito do indicativo	p
pretérito imperfeito do indicativo	m
pretérito imperfeito do subjuntivo	e
futuro do pretérito do indicativo	a
futuro do subjuntivo	y
infinitivo	i
gerúndio	g
Repetição do verbo no mesmo enunciado	
sem repetição	w
com repetição	u
Concordância entre o verbo e o SN	
Verbo singular → SN singular	~
Verbo singular → SN plural	<
Verbo plural → SN plural	=