



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE HUMANIDADES
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA

MARIA VANDERLÚCIA SOUSA TABOSA

**A VARIAÇÃO NA CONCORDÂNCIA NOMINAL DE NÚMERO NO FALAR DO
CARIRI CEARENSE**

FORTALEZA – CE

2016

MARIA VANDERLÚCIA SOUSA TABOSA

A VARIAÇÃO NA CONCORDÂNCIA NOMINAL DE NÚMERO NO FALAR DO
CARIRI CEARENSE

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Linguística da Universidade Federal do Ceará, como requisito para obtenção do título de Mestre em Linguística. Área de concentração: Linguística.

Orientadora: Prof^a. Dr^a. Hebe Macedo de Carvalho

FORTALEZA – CE

2016

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

T117v Tabosa, Maria Vanderlúcia Sousa.
A variação na concordância nominal de número no falar do Cariri cearense / Maria Vanderlúcia Sousa
Tabosa. – 2016.
95 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Humanidades, Programa de Pós-Graduação em Linguística, Fortaleza, 2016.
Orientação: Prof. Dr. Hebe Macedo de Carvalho.

1. Sociolinguística Variacionista. 2. Concordância Nominal. 3. Cariri cearense. I. Título.

CDD 410

MARIA VANDERLÚCIA SOUSA TABOSA

A VARIAÇÃO NA CONCORDÂNCIA NOMINAL DE NÚMERO NO FALAR DO
CARIRI CEARENSE

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Linguística da Universidade Federal do Ceará, como requisito para obtenção do título de Mestre em Linguística. Área de concentração: Linguística.

Aprovada em ____ / ____ / ____

BANCA EXAMINADORA

Prof^a. Dr^a. Hebe Macedo de Carvalho (orientadora)
Universidade Federal do Ceará (UFC)

Prof^a. Dr^a. Maria Elias Soares
Universidade Federal do Ceará (UFC)

Prof^a. Dr^a. Aluíza Alves de Araújo
Universidade Estadual do Ceará (UECE)

A Deus, aos meus pais, Raimundo e Célia (*in memoriam*), a Ednaldo (meu esposo), aos meus filhos Pedro Arthur e Pedro Matheus.

AGRADECIMENTOS

A Deus, pelo dom da vida e por iluminar o meu caminho.

À minha família: meus pais, Raimundo e Célia; meu esposo Ednaldo; meus filhos Pedro Arthur e Pedro Matheus; meus irmãos, sobrinhos e sobrinhas que estiveram sempre por perto dando o apoio necessário.

À Professora Dra. Hebe Macedo de Carvalho, pelas orientações seguras e produtivas, pelo incentivo constante, pela tranquilidade e sabedoria com as quais sempre me transmitiu os conhecimentos, pela disposição e zelo diante das minhas limitações e, principalmente, pela compreensão e amizade.

À Universidade Federal do Ceará e aos professores e funcionários do Programa de Pós-graduação em Linguística.

À Secretaria de Educação do Estado do Ceará – SEDUC –, pelo incentivo.

À CAPES, pelo incentivo financeiro.

À Professora Maria Elias Soares e Aluíza Alves de Araújo, pelas contribuições acerca da pesquisa.

Aos amigos que Deus permite colocar em nosso caminho: Adriana Pinho, Adriana Cidralle, Meyssa, Nair, Rose, Susana, Jorge Carvalho, Ana Paula Martins, Cátia Lemos, Sávio, Rogéria e todos da turma do Mestrado Infantil V, e muitos outros que, de alguma forma, colaboraram na elaboração desse trabalho.

A todos aqueles que colaboraram, com a elaboração deste trabalho, meus sinceros agradecimentos!

RESUMO

Nesta pesquisa, analisamos a variável concordância nominal de número entre os constituintes do sintagma nominal (SN) no falar da região do Cariri cearense à luz dos pressupostos teórico-metodológicos da Teoria da Variação e Mudança Linguística (LABOV, WEINREICH, HERZOG, 2006; LABOV, 1978, 1994; 2001; 2003). Nosso objetivo geral foi investigar em que medida os fatores linguísticos *posição dos elementos no sintagma, classe e posição em relação ao núcleo e à posição nuclear; classe gramatical do sintagma nominal e processos morfofonológicos de formação de plural; tonicidade das sílabas dos itens lexicais singulares e marcas precedentes de plural no âmbito do sintagma nominal*, e fatores extralingüísticos, como *sexo; anos de escolaridade e faixa etária* condicionam o uso de concordância nominal no falar dos municípios que formam a região CRAJUBAR – Crato, Juazeiro do Norte e Barbalha – na região do Cariri cearense. Para análise quantitativa, coletamos 3304 dados de entrevistas realizadas com 24 informantes nesses três municípios. As entrevistas estão presentes no *corpus O Português falado no Ceará* do banco de dados do PROFALA . Os dados foram submetidos à análise estatística no programa GoldVarb X. No tocante aos nove grupos de fatores considerados inicialmente, seis grupos foram selecionados como significativos: *posição dos elementos no sintagma; classe e posição em relação ao núcleo e posição nuclear; processos morfofonológicos de formação de plural; marcas precedentes de plural no âmbito do sintagma nominal*; e os fatores extralingüísticos: *sexo e faixa etária*. Quanto aos resultados obtidos, identificamos um alto índice de marcação, totalizando um percentual de 76,9 %. As mulheres apresentaram uma probabilidade de marcação maior que a dos homens, e os informantes na faixa etária de 15 a 25 anos e a partir de 50 anos apresentaram índices maiores que os dos informantes com faixa etária de 26 a 49 anos.

Palavras-chave: Sociolinguística Variacionista. Concordância Nominal. CRAJUBAR.

ABSTRACT

This work analyzes the nominal agreement of number between the constituents of the nominal phrase (NP) in Portuguese spoken in region of Cariri, in the state of Ceará, Brazil, in light of the theoretical-methodological assumptions of the Theory of Variation and Linguistic Change (LABOV, WEINREICH, HERZOG, 2006; LABOV 1978, 1994, 2001). Its general objective is to investigate to what extent the linguistic factors *position of the elements in the phrase, class and position in relation to nucleus and nuclear position; grammatical class of the nominal phrase and morpho-phonological processes of formation of plural; tonicity of the syllables of the singular lexical items and precedent marks of plural within the scope of the nominal phrase*, and extralinguistic factors *sex; years of schooling and age range* condition the use of nominal agreement in the speech of the towns that make up the CRAJUBAR region – Crato, Juazeiro do Norte, and Barbalha – in the region of Cariri, in the south of Ceará, Brazil. For quantitative analysis, it was collected 3304 data from interviews with 24 informants in these three towns. The interviews are present in the corpus PROFALA (The Portuguese spoken in Ceará). The data were submitted to statistical analysis in the GoldVarb X program. For the nine groups of factors initially considered, six groups were selected as significant: position of the elements in the nominal phrase, class and position in relation to the nucleus and nuclear position; morpho-phonological processes of plural formation; precedent marks of plural in the scope of the nominal phrase and the extralinguistic factors; sex; and age group. Regarding the results obtained, it was identified a high marking index totaling a percentage of 76.9%. Women had a marking probability higher than men, and informants between the ages of 15 and 25 years and 50 years old presented rates higher than the informants with ages ranging from 26 to 49 years.

Keywords: Variationist Sociolinguistics. Nominal agreement. CRAJUBAR.

LISTA DE FIGURAS

Figura 1 - Distribuição das pesquisas acerca da concordância nominal no Brasil	38
Figura 2 - Região do Cariri cearense	50
Figura 3 - Mapa do território do Cariri (Municípios da Região do Cariri cearense).....	51
Figura 4 - Adaptação ao mapa da Região Metropolitana do Cariri.....	54

LISTA DE TABELAS

Tabela 1 – Frequência da “presença de marcas explícitas de plural /s/ nos constituintes do SN” segundo a variável <i>posição linear</i>	71
Tabela 2 – Frequência da “presença marcas explícitas de plural nos constituintes do SN” segundo a variável <i>classe gramatical</i>	72
Tabela 3 – Frequência e peso relativo da “presença de marcas explícitas de plural” segundo a <i>posição e classe gramatical em relação ao núcleo</i>	74
Tabela 4 – Frequência e peso relativo da “presença de marcas explícitas de plural” segundo a variável <i>marcas precedentes</i>	76
Tabela 5 – Frequência e peso relativo da “presença de marcas explícitas de plural” segundo a variável <i>processos morfofonológicos de formação de plural</i>	79
Tabela 6 – Frequência e peso relativo da “presença de marcas explícitas de plural” segundo a variável <i>faixa etária</i>	80
Tabela 7 – Frequência e peso relativo da “presença de marcas explícitas de plural” segundo a variável <i>sexo</i>	83
Tabela 8 – Frequência da “presença de marcas explícitas de plural” segundo o cruzamento das variáveis <i>sexo vs faixa etária</i>	85
Tabela 9 – Frequência da “presença de marcas explícitas de plural” segundo o cruzamento das variáveis <i>sexo vs anos de escolaridade</i>	83

LISTA DE QUADROS

Quadro 1 - Padrões de mudança do indivíduo e da comunidade.....	28
Quadro 2 - Estratificação social dos informantes do CRAJUBAR.....	47
Quadro 3 - Características sociais dos 24 falantes do CRAJUBAR.....	48
Quadro 4 - Os sintagmas nominais considerados.....	54
Quadro 5 - Sintagmas excluídos	54
Quadro 6 - Posição linear do elemento no SN.....	57
Quadro 7 - Classe Gramatical do elemento sob análise	58
Quadro 8 - Classe e posição em relação ao núcleo.....	59
Quadro 9 - Marcas precedentes em função da posição	60
Quadro 10 - Processos morfofonológicos de formação de plural	61
Quadro 11 - Tonicidade dos itens lexicais singulares	61

SUMÁRIO

1	INTRODUÇÃO	12
2	FUNDAMENTAÇÃO TEÓRICA	18
2.1	A Sociolinguística Variacionista	18
2.2	A concordância nominal de número no português do Brasil segundo a GT e sob a concepção variacionista	32
3	PROCEDIMENTOS METODOLÓGICOS	47
3.1	O banco de dados PROFALA	47
3.2	Perfil dos Informantes da amostra	48
3.3	A região do Cariri cearense	50
3.3.1	<i>Formação Histórica</i>	50
3.3.2	<i>CRAJUBAR</i>	52
3.4	Caracterização dos dados	55
3.5	O tratamento estatístico dos dados	56
3.6	A variável dependente e as variáveis linguísticas	57
3.6.1	<i>As variáveis linguísticas</i>	58
3.6.1.1	<i>Três variáveis linguísticas: posição linear no SN, classe gramatical do elemento sob análise e classe e posição em relação ao núcleo</i>	58
3.6.1.2	<i>Marcas precedentes</i>	61
3.6.1.3	<i>Processos morfofonológicos de formação de plural</i>	62
3.6.1.4	<i>Tonicidade dos itens lexicais singulares</i>	63
3.7	As variáveis independentes extralinguísticas	63
3.7.1	<i>Sexo</i>	63
3.7.2	<i>Anos de Escolaridade</i>	65
3.7.3	<i>Faixa etária</i>	65
4	ANÁLISE E DISCUSSÃO DOS RESULTADOS	67
4.1	A Variação na concordância nominal de número na região CRAJUBAR	67
4.1.1	<i>Variáveis Linguísticas</i>	70
4.1.1.1	<i>Posição linear</i>	70
4.1.1.2	<i>Posição e classe gramatical em relação ao núcleo</i>	73
4.1.1.3	<i>Marcas precedentes</i>	76
4.1.1.4	<i>Processos morfofonológicos de formação de plural</i>	78
4.1.2	<i>Rodada sem as variáveis posição linear e classe gramatical</i>	81

4.1.3	<i>Variáveis Sociais</i>	83
4.1.3.1	<i>Faixa etária</i>	83
4.1.3.2	<i>Sexo</i>	84
4.1.4	<i>Cruzamentos entre fatores sociais</i>	85
5	CONSIDERAÇÕES FINAIS	88
	REFERÊNCIAS	91

1 INTRODUÇÃO

Em visita ao Museu da Língua Portuguesa na Estação da Luz, São Paulo, assisti a um vídeo na voz de Fernanda Montenegro. Nele, um trecho chamou minha atenção: “Os encontros e desencontros entre essas culturas e falares criaram uma língua única e original e que continua a se reinventar todos os dias pelas ruas e praças do país nos seus ritmos e ritos nos poemas, todos os dias”. Levando em consideração as expressões “*uma língua única*”, e “*... e que continua a se reinventar todos os dias pelas ruas e praças do país...*”, percebemos que a primeira vai de encontro a um dos princípios labovianos – não há falantes com estilo único, todos nós usamos mais ou menos variações a depender dos condicionamentos como relações entre falantes e interlocutores e contexto social. Quanto à segunda, teríamos, então, uma língua que sofre variação, que muda por causa de fatores linguísticos e sociais. E isso é fato! Mas, fazer uso dessas variações é compreensível e aceitável até que ponto? Com esse questionamento, introduzo minha justificativa.

Por muitas vezes, em conversa com outros professores, talvez pela ocasião informal, representava em minha fala o falar de meu pai – analfabeto proveniente das cheias do Aracati –, a de minha mãe – alfabetizada, que parou de estudar aos quatorze anos quando casou – e também a dos vizinhos – residentes na periferia de Fortaleza e, em sua maioria, analfabetos. Esse era o falar que não estava prescrito nas Gramáticas Tradicionais, (doravante GT), e não era o esperado para uma professora de Língua Portuguesa, mas insistia e se fazia presente vez em quando em meu discurso. Por conseguinte, esse falar “inadequado” era estigmatizado e reprovado por alguns interlocutores. Segundo Scherre (2008), o uso mais frequente de construções sem concordância leva a determinadas observações: “fulano não sabe falar português”, “empobrece a língua portuguesa”, entre outras. Pressupõe-se que falar português sem concordância de número – circunscrevendo a discussão ao tema da pesquisa – é não “saber falar” a Língua Portuguesa, ideia bastante generalizada.

Tal visão sobre o falar se deve ao fato de termos, como pressuposto, que falar ou escrever corretamente é seguir as regras prescritas na GT. Regras relevantes, porém, passíveis de contestação, visto que a língua se manifesta de forma heterogênea. Assim, este trabalho justifica-se, inicialmente, para responder a uma inquietação particular: no tocante à concordância nominal de número, quais motivações levam os falantes, especificamente do Cariri cearense, a não fazer uso de regras prescritas na GT? Quanto a essa problemática, a

Sociolinguística muito tem contribuído para desmistificar essas escolhas presentes na língua e utilizadas por diversos falantes.

Na verdade, em meu senso comum, buscava uma resposta: De quem é a culpa? Meus questionamentos iniciais encontram eco em Bagno (2009, p. 166), para o autor,

[...] existe na cultura ocidental uma velhíssima “tradição de queixa” contra a mudança linguística. Em cada momento dessa tradição, as pessoas tentam atribuir a “culpa” da mudança linguística a alguém ou a algum grupo social: ora são os jovens (“não sabem mais falar a língua só usam gíria” etc.), ora são os professores (que não são mais tão bons como antigamente” etc.), ora são os escritores (“que deixam contaminar pelos vícios da linguagem relaxada de hoje”), ora são os meios de comunicação modernos, ora é a “mistura de raças” resultante dos contatos de povos de “cultura superior” com povos de “cultura primitiva” (ou nenhuma cultura, melhor dizendo...), ora são as “babás, cozinheiras e engraxates” (como escreveu, num primor de preconceito social, o célebre gramático brasileiro Napoleão Mendes de Almeida). Enfim, não faltam candidatos para ocupar o lugar do réu no julgamento do “crime” da mudança linguística.

Acreditando ser a variação linguística condicionada e inerente ao falante, nossa pesquisa buscou os fatores linguísticos e extralinguísticos que favorecem “esse meu falar inadequado”. Encontramos na Teoria da Variação e Mudança (LABOV, 1972; 2001) os subsídios para tal empreendimento.

Desde o início dos anos 1960, com o advento da Sociolinguística, pesquisas com o propósito de descrever a heterogeneidade da língua intensificam-se. No entanto, foi com William Labov e seus colaboradores – (WEINREICH; LABOV; HERZOG, 2006) – que os trabalhos sobre a língua ganharam um novo modelo teórico-metodológico, capaz de descrever a heterogeneidade da língua e desvelar a influência dos fatores sociais que a condicionam.

O modelo de análise, proposto por Labov, é reconhecido como Teoria da Variação e Mudança ou Sociolinguística Quantitativa, por utilizar números e dar tratamento estatístico aos dados coletados. (TARALLO, 2007). A Teoria da Variação tem por objeto de estudo a língua falada – o vernáculo. Esse modelo teórico-metodológico traz em seu cerne a concepção de língua heterogênea, dinâmica e passível de mudanças decorrentes não só de fatores linguísticos, mas também de fatores extralinguísticos.

Segundo Mollica e Braga (2003), a Sociolinguística considera, em especial, como objeto de estudo exatamente a variação linguística. Sob essa perspectiva de estudo, a língua possui um caráter heterogêneo; tem função social, comunicativa e é fator importante na identificação dos grupos e na demarcação de diferenças sociais na comunidade. Sendo esse dinamismo inerente à língua, encontramos, no falar, formas variantes, ou seja, duas ou mais maneiras de manifestar o mesmo significado.

Feitas essas considerações acerca do arcabouço da Teoria da Variação e Mudança, passamos a enumerar vários trabalhos, produzidos a partir de diferentes estados do Brasil, cujo foco é a variável em estudo: concordância nominal de número no Português brasileiro. A relação seguinte, além de destacar importantes contribuições tem por objetivo demonstrar que no estado do Ceará precisamente na região do Cariri cearense não foi produzida pesquisa acerca deste fenômeno – a variável concordância nominal de número entre os constituintes do sintagma nominal (SN) no falar da região do Cariri cearense.

As primeiras pesquisas foram feitas por Braga e Scherre (1976). Em seguida, Braga (1977), Scherre (1978, 1988), Ponte (1979), Nina (1980), Guy (1981), Fernandes (1996), Carvalho, H., (1997), Carvalho, R., (1997), Lopes (2001), L. Andrade (2003), P. Andrade (2003), Santos (2010), Schneider (2012), Martins (2013), Guimarães (2014), Meira (2015), para citar apenas alguns.

Louvamos e reconhecemos a importante contribuição que todas as pesquisas, concluídas até então, trouxeram aos estudos sobre a descrição do uso de concordância nominal de número. No entanto, sabemos que o processo de construção do conhecimento científico não se dá em curto prazo, nem os resultados obtidos a cada pesquisa produzida podem ser vistos como absolutos.

Sendo um país de extensão continental, que apresenta uma diversidade linguística surpreendente em suas regiões, existem muitas localidades no Brasil que são campos férteis para a investigação linguística. Há muito que se fazer em termos de descrição do português falado no Brasil.

Respaldados nos pressupostos teóricos da Sociolinguística Quantitativa e confrontados pela necessidade de contribuir com os estudos de descrição do uso das variantes de concordância nominal de número, sentimo-nos impelidos a realizar nossa pesquisa, e, assim, vemo-nos justificados em fazê-lo. Afinal, não consta nas pesquisas, até então produzidas, estudos referentes à concordância nominal de número no falar do Cariri cearense, à luz da Teoria da Variação e Mudança ou Teoria Laboviana.

Tal lacuna nos conduziu ao presente estudo, que tem como objetivo responder a três questionamentos, a saber: I) Em que medida os fatores linguísticos: *posição da classe gramatical em relação ao núcleo e posição nuclear; classe gramatical do sintagma nominal e processos morfofonológicos de formação de plural; e tonicidade das sílabas dos itens lexicais singulares e marcas precedentes de plural no âmbito do sintagma nominal* condicionam o uso de concordância nominal na região CRAJUBAR - Crato, Juazeiro e Barbalha -, pertencente ao Cariri cearense? II) O sexo, a faixa etária e o nível de escolaridade condicionam os usos do

fenômeno em estudo? III) A análise dos dados relacionados à faixa etária dos falantes do CRAJUBAR revela uma variação estável ou demonstra existir tendência à mudança em progresso?

Essas questões da pesquisa estão paralelamente articuladas aos objetivos específicos deste trabalho, quais sejam: I) investigar a atuação dos grupos de fatores linguísticos no âmbito do sintagma nominal como possíveis condicionadores da variação da concordância nominal de número, no falar da região CRAJUBAR, do Cariri cearense; II) analisar o efeito das variáveis sociais: sexo, faixa etária e anos de escolaridade sobre o fenômeno em estudo; III) observar, com base na análise dos dados, a partir da faixa etária dos falantes, se a variável concordância nominal de número, na região do CRAJUBAR - Crato, Juazeiro e Barbalha -, reflete um processo de variação estável ou apresenta tendência à mudança em progresso.

Na busca pela resolução dos problemas de pesquisa implicados por esses objetivos, apresentamos as hipóteses que serão verificadas no desenvolvimento de nosso trabalho. Quanto aos fatores linguísticos no âmbito do sintagma nominal, concernentes ao objetivo I, supusemos que os fatores linguísticos *posição e classe gramatical em relação ao núcleo e posição nuclear, classe gramatical do sintagma nominal e processos morfofonológicos de formação de plural e tonicidade das sílabas dos itens lexicais singulares e marcas precedentes de plural* atuavam, no âmbito do sintagma nominal, como possíveis condicionadores da variação da concordância nominal de número, no falar da região CRAJUBAR - Crato, Juazeiro e Barbalha, doravante CRAJUBAR.

No que toca ao objetivo II, conjecturamos que, quanto ao *sexo*, mulheres são mais favoráveis à presença de marcas explícitas de plural /s/. Em relação à *faixa etária*, julgamos que os falantes do CRAJUBAR de 26 a 49 anos e os falantes com mais de 50 anos empregam, com mais intensidade, as regras de concordância nominal de número em razão de outros fatores como empregabilidade. No que diz respeito ao fator *anos de escolaridade*, supomos que quanto maior o tempo de escolaridade, mais frequente é o emprego da presença de marcas explícitas de plural /s/ pelos falantes do CRAJUBAR.

A hipótese que levantamos acerca do objetivo III é a de que o uso das regras de concordância nominal pelos falantes do CRAJUBAR seria considerado um fenômeno de variação estável.

A realização desta pesquisa se justifica também pelo fato de que, em quase todas as investigações apresentadas neste trabalho sobre o fenômeno analisado por nós, os fatores

avaliados, geralmente, são semelhantes; no entanto, alguns resultados se manifestam de maneira diferenciada de acordo com a *comunidade de fala* pesquisada. Assim, faz-se necessário desenvolver nosso estudo para identificarmos semelhanças e diferenças a outras localidades. Os índices obtidos, também, podem ser úteis como fonte de pesquisa para, em outra oportunidade, compararmos os fatores linguísticos e extralinguísticos com a metrópole Fortaleza – capital do Estado do Ceará.

Por fim, a realização desta pesquisa faz-se necessária para, também, contribuir com a educação sistematizada – a sala de aula propriamente dita – visto que, no que concerne ao ensino da concordância nominal, encontramos, ainda, um enorme fosso entre o que está prescrito nas gramáticas e manuais didáticos que utilizamos como parâmetro e as reais e diferentes manifestações presentes no falar cotidiano dos brasileiros. Os estudos realizados sobre as diversas regiões do país têm mostrado que a concordância nominal de número é um fenômeno variável, motivado por fatores de ordem interna e externa e que sofre pressões de uso formal e informal. O resultado dessa e das demais pesquisas nesse campo precisam adentrar o âmbito escolar e fundamentar nosso discurso no tocante à descrição gramatical.

A fim de analisarmos o comportamento da variável dependente em estudo – *presença da marca explícita de plural /s/ no falar do CRAJUBAR-*, coletamos os 3304 dados das entrevistas realizadas por 24 informantes da região CRAJUBAR ocorridas no período de 1996 a 2001, encontrando-se presentes no *corpus O Português falado no Ceará*, do banco de dados do PROFALA. Após a seleção, os dados foram submetidos à análise estatística no programa GoldVarb X.

Esta pesquisa encontra-se dividida em cinco capítulos. O primeiro capítulo traz a introdução. O segundo capítulo trata da fundamentação teórica, do estado da arte. Nele, apresentamos os pressupostos teórico-metodológicos da Teoria da Variação e Mudança e fazemos uma breve referência a alguns estudos realizados na área da sociolinguística variacionista e a visão segundo a tradição gramatical, tendo o português do Brasil como objeto. Esses estudos contemplam, sob algum aspecto, o nosso objeto de estudo – regra variável da concordância nominal de número no sintagma nominal. No terceiro capítulo, abordamos os procedimentos metodológicos desta pesquisa: esclarecemos o que vem a ser o banco de dados – PROFALA-; aprofundamos explicações acerca do perfil dos informantes, da história e da geografia da Região do Cariri cearense e da região CRAJUBAR; discutimos sobre variável dependente, variáveis independentes e explanamos o tratamento estatístico ao qual submetemos os dados coletados. No quarto capítulo, apresentamos a análise e a

discussão acerca dos resultados das variáveis linguísticas e extralinguísticas. Por último, apresentamos as nossas considerações finais. Sintetizamos as conclusões finais a que chegamos em nosso trabalho, apontando as contribuições fornecidas para novas pesquisas e avaliando os pontos de dificuldade enfrentados ao longo do processo de realização de nossa pesquisa.

Por fim, esperamos, com a conclusão deste trabalho, ter contribuído com o desenvolvimento de estudos no campo da descrição da língua portuguesa na região do Cariri cearense.

2 A SOCIOLINGUÍSTICA VARIACIONISTA

Este capítulo está subdividido em duas seções. Na primeira, apresentamos os pressupostos teórico-metodológicos da Teoria da Variação e Mudança – arcabouço teórico em que nosso trabalho está inserido. Na segunda seção, apresentamos alguns estudos realizados, no Português do Brasil, referentes ao nosso objeto de estudo – *concordância nominal de número no sintagma nominal* com base na Teoria da Variação e na tradição gramatical.

2.1 Teoria da Variação e Mudança

Nesta seção, abordamos a Sociolinguística laboviana; elencamos seus principais fundamentos e destacamos, principalmente, a relação entre língua e sociedade, bem como o caráter heterogêneo da língua.

O interesse pelos estudos acerca da linguagem remonta a tempos bem distantes. As razões e os interesses em busca de conhecimentos sobre essa capacidade humana foram e são diversos ao longo da história. Fiorin (2002), fazendo um breve painel desse percurso, traz o século IV a. C. como o período no qual os hindus, por motivos religiosos, e o gramático Panini, descreveram e produziram modelos de análise da língua dos hindus. Fiorin cita também os gregos Platão e Aristóteles. O primeiro buscava uma análise da estrutura linguística, já Aristóteles tentava uma análise da estrutura linguística. O latino Varrão tentava definir a gramática como ciência e como arte. O italiano Ambrósio Calepino, no século XVI, produziu o mais antigo dicionário poliglota. Outros feitos foram acrescentados, demonstrando o interesse pela linguagem; em 1660, a *Gramática de Port Royal* de Lancelot e Arnaud é concebida como parâmetro a inúmeras gramáticas produzidas no século XVII. No século XIX, estudos direcionados a um raciocínio mais abstrato da linguagem tornam-se mais raros, por conseguinte, as pesquisas direcionam seu foco ao estudo comparativo dos falares. Nesse período, ressalta Fiorin (2002, p. 12): “O estudo comparado das línguas vai evidenciar o fato de que as línguas se transformam com o tempo, independentemente da vontade dos homens, seguindo uma necessidade própria da língua e manifestando-se de forma regular”. O reconhecimento de que a mudança da língua é inerente ao homem será de grande importância aos estudos variacionistas no início dos anos de 1960. Vale destacar, nesse período, o surgimento da Linguística Histórica com Franz Bopp em 1816. No século XIX, a descoberta

de que as mudanças observadas na escrita teriam acontecido inicialmente na fala leva os estudiosos a priorizarem também a fala.

É no início do século XX, mais especificamente em 1916, com o lançamento da obra *Curso de Linguística Geral*, que Saussure delimita e define o objeto de estudo da linguística – a *língua*. Por conseguinte, a Linguística é reconhecida enquanto ciência. Considerado um dos maiores expoentes do *Estruturalismo*, Saussure é referência importante aos estudos linguísticos até então. A partir da década de 1960, nos Estados Unidos, Noam Chomsky introduz à Linguística o *Gerativismo*. Em seu livro *Estruturas Sintáticas*, publicado em 1957, Chomsky assegurava que o papel fundamental da linguística é descrever com cunho científico o conhecimento linguístico dos falantes, suas palavras, frases e discursos. De acordo com Kenedy (2013, p. 18)

A abordagem de Chomsky foi revolucionária para a época, pois até a metade do século passado, a linguística ocupava-se quase exclusivamente da dimensão social e histórica da linguagem humana. A partir das ideias de Chomsky, os linguistas passaram a não mais apenas descrever a estrutura das línguas, mas também a procurar explicações para como a mente humana era capaz de adquirir e processar essas estruturas. Com Chomsky, a morada da linguagem, a das línguas naturais, passou a ser a mente dos indivíduos.

Diferente de Saussure, que considera a língua uma instituição social, Chomsky considera a linguagem uma faculdade mental e geneticamente transmitida pela espécie. No entanto, nenhum dos dois nega o lado social da língua, apenas não o priorizam como seu objeto científico.

É diante desse contexto de estudos de viés *estruturalista* ou *gerativista* que nasce a *Sociolinguística* em 1960 nos Estados Unidos, tendo como arcabouço as pesquisas de William Labov, seu principal mentor. Inserindo o estudo da língua ao contexto social, Labov percebe que a língua não deve ser analisada como uma estrutura autônoma, independente dos contextos reais de produção.

A Sociolinguística possui muitas áreas de interesse, um campo vasto de estudos capaz de descrever diversas formas de abordagem da linguagem. Nesse campo de pesquisa, encontramos estudos relativos ao contato entre as línguas, ao surgimento e à extinção linguística, o multilinguismo, a variação e a mudança linguística, estes dois últimos aspectos – variação e mudança – fundamentais para o fenômeno em estudo.

Quanto ao conceito de Sociolinguística, Hudson (1980, p.1) afirma que podemos defini-la como o estudo da língua em relação à sociedade. Esse conceito nos parece ser o cerne da Sociolinguística. No entanto, para os dias atuais e diante da diversidade de estudos

nesta área, ou seja, diante da área que hoje abrange a Sociolinguística, de acordo com Meyerhoff (2006, p. 1), acrescentamos:

[...] a Sociolinguística é um campo muito vasto e pode ser usado para descrever diversas formas de estudar a linguagem. Se uma pessoa, por exemplo, descreve uma situação engraçada como “kicksin” eu sei que ela vem ou passou um bom tempo no Caribe, região que fala a língua inglesa. Estou me baseando no conhecimento sociolinguístico para inferir isso¹.

Ratificando o dito por Meyerhoff (2006), Camacho (2006, p. 49) afirma: “dizer que a sociolinguística trata da relação entre língua e sociedade é fazer uma afirmação correta, mas, ao mesmo tempo, excessivamente simplificadora”. Destaca o autor que, nas últimas três décadas, é crescente o número de estudos da linguagem em uso no contexto social. Segundo ele, sob o rótulo “Sociolinguística”, trabalhos com diferentes enfoques são produzidos. Cabe ao pesquisador delimitar seu trabalho. Assim, já seguindo essa orientação, nossa pesquisa está inserida no arcabouço da Teoria da Variação e Mudança ou Sociolinguística Quantitativa (LABOV, 2008; 1994; 2001; WEINREICH; LABOV e HERZOG, 2006; GUY; ZILLES, 2007).

O termo Sociolinguística, em princípio, sofreu certa resistência da parte de Labov por vários anos. Para o autor, era difícil conceber uma teoria ou prática linguística que não fosse social. Em outros termos, para Labov, a Sociolinguística era a própria Linguística. Essa visão, no entanto, não era compartilhada por todos os pesquisadores nos anos 1960.

Labov, enquanto estudante, inicia seus estudos linguísticos em 1961. Os primeiros projetos concebidos por ele foram *Ensaio em linguística experimental*, aplicados em espaços reais, tendo a língua em uso na sociedade. Observando o comportamento dos falantes em seu mundo cotidiano, Labov começa a perceber que seus princípios empíricos não se encaixavam na linguística em voga naquele contexto, pois encontrava diversos obstáculos ideológicos ao estudo da língua em relação com a sociedade (LABOV, 2008, p. 13). O primeiro obstáculo surge diante da opção de Saussure por separar estudos diacrônicos e sincrônicos. O segundo decorre de Bloomfield defender a regularidade da mudança sonora contra qualquer evidência irregular do presente, declarando serem apenas casos de empréstimo dialetal (BLOOMFIELD, 1933, p. 364 *apud* LABOV, 2008, p. 14). Por fim, considerada por Labov a mais importante restrição aos seus princípios, a variação livre, em tese, não podia ser condicionada. Segundo o postulado básico da linguística – em Bloomfield (1933) –, formas

¹ “Sociolinguistics is a very broad field, and it can be used to describe many different ways of studying language. (...) For example, if a speaker describes a funny or amusing situation as ‘kicksin’, I know they are from, or have spent a good deal of time in, the English-speaking Caribbean. I am drawing on sociolinguistic (social and linguistic) knowledge to draw this inference.”

linguísticas em variação livre se permutam indiferentemente a condicionantes. Se uma ou outra forma linguística ocorria num momento particular, isso era linguisticamente irrelevante. Essa forma de avaliar o fenômeno linguístico excluía a avaliação social das variantes linguísticas, que, nesse contexto, estavam fora de consideração (LABOV, 2008, p. 14).

Nesse contexto de tantos empecilhos ideológicos, a Teoria da Variação e Mudança talvez não se firmasse se Labov não tivesse encontrado, na Columbia University, o professor Uriel Weinreich. Foi sob a orientação de Weinreich que Labov defendeu a tese intitulada *The social history of sound change on the island Of Martha's Vineyard*, na qual apresenta um fenômeno de mudança fonética mediante dados da fala dos habitantes da ilha de Martha's Vineyard.

A ilha Martha's Vineyard, município de Dukes, localiza-se no litoral do estado Norte Americano de Massachussets). É caracterizada por ser um elegante local de veraneio e por receber aproximadamente 40 mil turistas durante os meses de verão. Na época do estudo, em 1962, segundo o *U.S Bureau of the Census of Population: 1960*, a ilha contava com 5.563 habitantes aproximadamente. Martha's Vineyard contava com a vantagem de ser separada do continente, aproximadamente cinco quilômetros, no Oceano Atlântico. Uma ilha social e geograficamente complexa a ponto de oferecer um vasto campo à diferenciação do comportamento linguístico. Como suporte à investigação, Labov contou com os registros do *Linguistic Atlas of New England*. Segundo Labov, mesmo sendo conhecida entre os linguistas como uma área conservadora do inglês americano, pronunciadores do *r*, e preservadores de traços arcaicos e relíquias lexicais do inglês seiscentista, fortalecem essa natureza arcaica da tradição vineyardense. Por conter essa riqueza arcaica, Labov percebeu outra variável muito singular: as diferenças na altura do primeiro elemento dos ditongos /ay/ e /aw/. E foi com essa variável que buscou compreender a estrutura interna do inglês vineyardense – diferenças sistemáticas já existentes bem como as mudanças linguísticas que ocorriam na ilha. Para o feito, Labov produziu uma sequência metodológica: entrevistas, escalas de medição, distribuição dos falantes por idade e localização geográfica e outras técnicas, as quais com o decorrer do tempo foram sendo aperfeiçoadas. (LABOV, 2008, p. 22 - 33).

Após as pesquisas a respeito do inglês falado na ilha de Martha's Vineyard, Labov produziu outros estudos, como a estratificação social do inglês falado em Nova Iorque (1966), estudo sobre o inglês vernacular dos adolescentes negros do Harlem e estudos sociolinguísticos da Filadélfia. Em 1968, Weinreich, Labov e Herzog publicaram a obra *Fundamentos empíricos para uma teoria da linguística*, um clássico dos estudos em

linguística histórica pela relevância de suas reflexões para a área, sendo um marco definidor de caminhos sobre a mudança em comunidades linguísticas contemporâneas (cf. a apresentação feita por Faraco na publicação da tradução do livro para o português, p. 10). Nessa obra, os autores fazem uma leitura crítica das tradições neogramáticas, estruturalistas e gerativistas e, com base nesses estudos, sistematizaram um conjunto de princípios para o estudo da mudança linguística, tomando por princípio a concepção de língua como fenômeno caracterizado pela heterogeneidade ordenada.

Weinreich, Labov e Herzog (2006, p. 36) (doravante WHL) asseguram que o domínio de um falante nativo (*nativelike command*) de estruturas heterogêneas não tem relação com multidialetalismo nem com o “mero” desempenho, mas é parte da competência linguística monolíngue. Afirmam ainda que numa língua, utilizada em uma comunidade real, a ausência de heterogeneidade estruturada é que seria disfuncional. Essa afirmativa só reitera a concepção de língua da Teoria Laboviana: heterogênea, dinâmica, relacionada com a sociedade e passível de mudanças decorrentes não só de fatores internos – inerentes à língua, linguísticos –, mas também por motivações extralinguísticas.

Portanto, surge, na década de 1960, a Teoria da Variação e Mudança ou Sociolinguística Quantitativa, desenvolvida por William Labov, Uriel Weinreich, e Marvin Herzog, cujo objeto de estudo é a língua falada – o vernáculo. De acordo com Tarallo (2007, p. 19), a língua falada a que nos referimos é o veículo linguístico de comunicação usado em situações naturais de interação social, do tipo comunicação face a face. Refere-se aos momentos em que o mínimo de atenção é prestado à língua. O vernáculo constitui, portanto, o material básico para a análise sociolinguística.

A Teoria da Variação e Mudança ou Sociolinguística Quantitativa acrescenta muito à forma como, cada um a seu modo, Saussure e Chomsky concebem a *língua*. No que concerne ao estruturalismo, Labov ressalta que, mesmo reconhecendo a parte social da língua, a maioria dos linguistas que adotam o pensamento de Saussure não leva em consideração a vida social. Segundo Saussure,

langue “est la partie sociale du langage... elle n'existe qu'en vertu d'une sorte de contrat passé entre les membres de la communauté” [“é a parte social da linguagem... ela não existe fora de um tipo de contrato estabelecido entre os membros da comunidade”] (1962: 321). Por esta razão, a Escola de Genebra saussuriana é frequentemente mencionada como a escola “social” da linguística. Saussure concebia a linguística como uma parte de “*une Science qui étudie le vie des signes au sein de la vie sociale* [uma ciência que estuda a vida dos signos no seio da vida social] (LABOV, 2008, p. 217).

Para Saussure (2001, p. 22), a língua não constitui, pois, uma função do falante: é o produto que o indivíduo registra passivamente; não supõe jamais premeditação. Ele prioriza, portanto, a língua, não a fala, e a concebe de forma homogênea. Por sua vez, para Chomsky, o objeto dos estudos linguísticos é a competência linguística do falante-ouvinte ideal, pertencente a uma comunidade linguisticamente homogênea (TARALLO, 2007, p. 6).

Nessa concepção de língua heterogênea, não podemos deixar de dedicar uma atenção ao conceito de variação. A Teoria da Variação e Mudança rompe com modelos teóricos, vigentes na época, contrários ao reconhecimento da língua como sistema heterogêneo ordenado, passível de sistematização, que, por conseguinte, sofre variação. Uma variação não livre, mas condicionada por fatores linguísticos e extralinguísticos – já que não se pode compreender o desenvolvimento de uma mudança linguística sem considerar a vida social da comunidade. Nesse contexto, Labov introduz no modelo de heterogeneidade ordenada o conceito de variável linguística, visto que, em toda comunidade de fala, são frequentes as formas linguísticas em variação. Essas formas em variação, define Labov (1978), são dois enunciados que se referem ao mesmo estado de coisas e que possuem o mesmo valor de verdade pois constituem variantes linguísticas de uma mesma variável².

Quanto às variáveis, podem ser variáveis dependentes e independentes. Em nosso estudo, temos por variável dependente *a presença da marca explícita de plural /s/* e *a ausência de marca explícita de plural*. Nessa variável, as variantes são as formas que co-ocorrem: *a presença ou a ausência da regra de concordância nominal de número*. Enquanto variação condicionada, as ocorrências são influenciadas por variáveis independentes: fatores linguísticos e extralinguísticos.

Ainda acerca da definição de variáveis dependente e independente, Guy e Zilles (2007) afirmam que, geralmente, os sociolinguistas concebem a taxa de uso de um traço linguístico como de algum modo causada, influenciada, ou determinada por características sociais, utilizando como exemplo a relativização em inglês, com a qual se deseja estabelecer uma associação entre duas dimensões diferentes: a escolha dos relativos e a classe social dos falantes – ambas variáveis. Assim, a escolha do pronome relativo é chamada de *variável dependente*, enquanto a classe social do falante é chamada de *variável independente*.

O conceito de variante linguística proposto por Labov – dois enunciados referindo-se ao mesmo estado de coisas e com mesmo significado – a princípio, foi aceito sem

² “Though formal linguistics recognizes the existence of expressive and affective information, these are in practice subordinated to what Buhler (1934) called “representational meaning” or what I will call “states of affairs.” To be more precise, I would like to say that two utterances that refer to the same state of affairs have the same truth-value, ...”

muitas restrições às variáveis fonológicas. No entanto, fora do nível fonológico, a literatura cita, com relevância, o questionamento feito por Lavandera (1978) sobre a não aplicabilidade da análise da variação a fenômenos morfológicos, lexicais e principalmente sintáticos.

A discussão entre Labov (1978) e Lavandera (1978) surge acerca de um trabalho produzido por Weiner e Labov (1983) sobre um fenômeno de variação sintática - estruturas ativa e passiva do inglês. A autora defende que toda construção sintática possui seu próprio significado referencial. Sendo assim, é impossível haver variação nas construções apresentadas por Weiner e Labov.

Acho inadequado, no estado atual das pesquisas sociolinguísticas, estender a outros níveis de análise variacionista a noção de variável sociolinguística, que originalmente foi desenvolvido com base em dados fonológicos, aos estudos quantitativos de variação que lidam com variação morfológica, sintática e lexical. Estes necessitam de uma teoria mais consistente no tocante ao conceito de significado. Embora a análise de variação em fonologia através da definição de variáveis fonológicas possa contribuir para uma melhor compreensão aos tipos de informação que as diferenças na forma possam apresentar, a extensão da noção de variável a variações não-fonológicas podem, em muitos casos, não ser viáveis.³ (Tradução nossa)

Em resposta a Lavandera, Labov amplia a noção de *significado* para *estado de coisas*. Considerando “estado de coisas” como *significado representacional*. Neste contexto, percebe-se que mesmo as variantes apresentando sutis diferenças de sentido serão consideradas com o mesmo valor representacional.

Encontrada nos níveis fonológico, morfológico, sintático e discursivo, a ocorrência de variação linguística não implica necessariamente ocorrência de mudança na língua. Considerando a heterogeneidade sistemática e ordenada da língua e as conclusões acerca das observações feitas em suas pesquisas, WLH (2006) apresentam cinco questões pertinentes à teoria da mudança e lançam cinco problemas a serem resolvidos. Considerados como princípios relativos aos fundamentos empíricos para uma teoria da mudança linguística – *o problema dos fatores condicionantes; o problema da transição; o problema do encaixamento; o problema da avaliação e o problema da implementação*.

Sobre *fatores condicionantes*, WLH (2006) sugerem que se faz necessário determinar o conjunto de mudanças possíveis e de condições possíveis para a mudança que

³ I want to show that it is inadequate at the current state of sociolinguistic research to extend to other levels of analysis of variation the notion of sociolinguistic variable originally developed on the basis of phonological data. The quantitative studies of variation which deal with morphological, syntactic and lexical alternation suffer from the lack of an articulated theory of meanings. While the analysis of variation in phonology by defining phonological variables can be accepted as contributing to a better understanding of the kinds of information that differences in form may be conveying, the parallel extension of the notion of variable to non-phonological variation may in many cases be unrevealing. (LAVANDERA, 1978, p.171)

pode ocorrer numa dada estrutura – fatores linguísticos e sociais que favorecem ou restringem a mudança linguística. Tentamos observar esse problema em nosso estudo, ao controlarmos grupos de fatores linguísticos e sociais. Há ainda um ponto que precisamos discutir nesta seção: como ocorre a *transição*, isto é, como e por quais caminhos a língua muda?

O trabalho de Labov, realizado na ilha de Martha's Vineyard sobre essa transição ou transferência de traços de falante para outro, parece ocorrer por meio de falantes bidialetais ou, mais geralmente, falantes com sistemas heterogêneos caracterizados pela diferenciação ordenada. A mudança se dá à medida que um falante aprende uma forma alternativa; durante o tempo em que as duas formas existem em contato dentro de sua competência e quando uma das formas se torna obsoleta. A transferência parece ocorrer entre grupos de pares de faixas etárias levemente diferentes; todas as evidências empíricas reunidas até agora indicam que as crianças não preservam as características dialetais de seus pais, mas sim as do grupo de pares que domina seus anos pré-adolescentes.

No tocante ao *problema do encaixamento*, as restrições ou condicionadores internos e externos ajudam a compreender como ocorre o *encaixamento* na língua e na sociedade. O *encaixamento* pode ser observado quando estudos detectam uma correlação entre a estrutura social (grupo socioeconômico, idade, sexo, escolaridade, etnia, localização geográfica) e o fenômeno de mudança.

WLH afirmam que as mudanças linguísticas em observação devem ser vistas como encaixadas no sistema linguístico como um todo. Classificam o *encaixamento* de acordo com sua natureza e extensão: *encaixamento na estrutura linguística* e *encaixamento na estrutura social*.

Quanto à *avaliação*, Labov (2008, p. 193) afirma que o problema da *avaliação* é encontrar os correlatos subjetivos (ou latentes) das mudanças objetivas (ou manifestas) que foram observadas,

a teoria da mudança linguística deve estabelecer empiricamente os correlatos subjetivos dos diversos estratos e variáveis numa estrutura heterogênea. Estes correlatos subjetivos das avaliações não podem ser deduzidos a partir do lugar das variáveis dentro da estrutura linguística. Além disso, o nível de consciência social é uma propriedade importante da mudança linguística que tem de ser determinada diretamente (WLH, 2006, p. 124).

Sobre o *problema da implementação*, a dificuldade está em descobrir o número de fatores que restringiram ou favoreceram a mudança linguística. Os autores acrescentam também que, uma vez que a mudança linguística está encaixada na estrutura linguística, ela é gradualmente generalizada a outros elementos do sistema. Essa generalização não ocorre de

forma abrupta e, ao finalizar, novos grupos sociais entram na comunidade de fala. Por fim, uma das mudanças secundárias se torna primária (WHL, 2006, p. 124).

Preocupados com os métodos para explicação da mudança, WLH sintetizaram sete princípios gerais para o estudo da mudança linguística, a saber:

1. A mudança linguística não deve ser identificada como deriva aleatória procedente da variação inerente na fala. A mudança linguística começa quando a generalização de uma alternância particular num dado subgrupo da comunidade de fala toma uma direção e assume o caráter de uma diferenciação ordenada.
2. A associação entre estrutura e homogeneidade é uma ilusão. A estrutura linguística inclui a diferenciação ordenada dos falantes e dos estilos através de regras que governam a variação na comunidade de fala; o domínio do falante nativo sobre a língua inclui o controle destas estruturas homogêneas.
3. Nem toda variabilidade e heterogeneidade na estrutura linguística implica mudança; mas toda mudança implica variabilidade e heterogeneidade.
4. A generalização da mudança linguística através da estrutura linguística não é uniforme nem instantânea; ela envolve a covariação de mudanças associadas durante substanciais períodos de tempo, e está refletida na difusão de isoglossas por áreas do espaço geográfico.
5. As gramáticas em que ocorre a mudança linguística são gramáticas de comunidades de fala. Como as estruturas variáveis contidas na língua são determinadas por funções sociais, os idioletos não oferecem a base para gramáticas autônomas ou internamente consistentes.
6. A mudança linguística é transmitida dentro da comunidade como um todo; não está confinada a etapas discretas dentro da família. Quaisquer discontinuidades encontradas na mudança linguística são os produtos das discontinuidades específicas da comunidade, mais do que os produtos inevitáveis do lapso geracional entre pais e filhos.
7. Fatores linguísticos e sociais estão intimamente interrelacionados no desenvolvimento da mudança linguística. Explicações confinadas a um ou outro aspecto, não importa quão bem construídas, falharão em explicar o rico volume de regularidades que pode ser observado nos estudos empíricos do comportamento linguístico. (WHL, 2006, p. 125-126)

Sobre o terceiro princípio, em que nem toda variabilidade e heterogeneidade na estrutura linguística implica mudança, mas toda mudança envolve, obrigatoriamente, variabilidade e heterogeneidade, pode-se dizer que isso ocorre porque as variantes linguísticas podem estar em processo de mudança como também permanecerem estáveis. Cabe aos diferentes estudos sociolinguísticos investigar a estabilidade e a instabilidade das variantes. Para obtermos resultados sobre esse caminho seguido pela variação, podemos aplicar estudos em *tempo real* e em *tempo aparente*.

Segundo Tarallo (2007, p. 65), “um recorte transversal da amostra sincrônica em função da faixa etária dos informantes chama-se *tempo aparente*”. Sobre essa observação, Paiva e Duarte (2003, p. 179) apresentam dificuldades ao tempo aparente, são elas: “a própria validade da hipótese clássica acerca da aquisição da linguagem; a outra está no fato de que correlações sistemáticas com a variável idade não são, muitas vezes, índices conclusivos de uma mudança em progresso”. Para os autores, somente o estudo da mudança no *tempo real* –

o qual envolve comparação entre a forma como as pessoas falam em um ponto no tempo e a forma como eles falam em uma década, ou geração – pode fornecer evidências que permitam interpretar de maneira mais segura correlações etárias significativas. Portanto, nos estudos sobre *tempo real* e *tempo aparente*, podemos evidenciar como a variação desempenha um papel importante na mudança da língua. Sobre as relações entre *tempo real* e *tempo aparente*, Labov (1994) acrescenta:

1. Se o comportamento dos indivíduos é estável ao longo de suas vidas, e a comunidade permanece no mesmo nível, não há variação para analisar, e nós temos a *estabilidade*;
2. Se os indivíduos mudam o seu comportamento linguístico em toda sua vida, mas a comunidade como um todo não muda, o padrão pode ser caracterizado como uma *age grading* – gradação etária;
3. Se falantes individuais entram na comunidade com uma frequência característica para uma variável particular, mantida ao longo de suas vidas; mas com aumentos regulares nos valores adotados por indivíduos, muitas vezes motivados por gerações, levam a mudança linguística para a comunidade, temos a mudança geracional. Essa é a típica mudança de som e alteração morfológica;
4. O inverso desse último padrão é a mudança comum, quando todos os membros da comunidade alteram suas frequências em conjunto, ou adquirem novas formas simultaneamente.⁴

A interpretação dos dados em tempo real, elaborados a partir de estudos de *painéis* ou estudos de *tendência*, requer um modelo subjacente de como os indivíduos mudam ou não mudam, de como as comunidades mudam ou não mudam ao longo do tempo, e do que pode resultar das combinações dessas possibilidades. As combinações mais simples produzem quatro padrões distintos apresentados no Quadro 1⁵:

Quadro 1 – Padrões de mudança do indivíduo e da comunidade

	Indivíduo	Comunidade
1. Estabilidade	Estável	Estável
2 Gradação Etária	Instável	Estável
3 Mudança Geracional	Estável	Instável
4 Mudança Comunitária	Instável	Instável

⁴ (1) If the behavior of individuals is stable throughout their lifetimes, and the community remains at the same level, there is no variation to analyze, and we have *stability*; (...) If individuals change their linguistic behavior throughout their lifetimes, but the community as a whole does not change, the pattern can be characterized as one of *age-grading*. (...) (3) *generational change* is the normal type of linguistic change that we have been considering so far – most typical of sound change and morphological change. Individual speakers enter the community with a characteristic frequency for a particular variable, maintained throughout their lifetimes; but regular increases in the values adopted by individuals, often incremented by generations, lead to linguistic change for the community. (4) The converse of this pattern is *comunal change*, where all members of the community alter their frequencies together, or acquire new forms simultaneously. (LABOV. 1994. P.83-84)

⁵ “The interpretation of real-time data, drawn from panel studies or trend studies, requires an underlying model of how individuals change or do not change during their lives, how communities change or do not change over time, and what may result from combinations of these possibilities. The simplest combinations produce four distinct patterns.” (LABOV, 1994, p. 83)

Fonte: Labov (1994, p. 83).

Meyerhoff (2006, p. 24) apresenta alguns fatores que motivam a variação linguística. Segundo a autora, é possível identificar um pequeno grupo de motivos recorrentes em análises sociolinguísticas. A variação com que as pessoas usam a língua é muitas vezes atribuída às quatro motivações seguintes:

- 1) desejo de demonstrar como você se adequa com algumas pessoas e como são diferentes umas das outras;
- 2) desejo de fazer coisas que tenham valor na comunidade (e associar-se a elas);
- 3) desejo de não fazer coisas que são mal vistas na comunidade (e de não ter as pessoas recriminando-lhe); e
- 4) desejo de observar como os outros estão orientando quanto às questões 1 e 3.

Destacamos alguns pontos relevantes no tocante à *mudança linguística*. A seguir adentraremos um espaço muito complexo e escorregadio, o conceito de *comunidade de fala*. De acordo com Monteiro (2000), uma das questões mais complexas na teoria laboviana refere-se ao conceito de *comunidade*, visto que não é fácil estabelecer os seus limites geográficos ou sociais, conforme podemos verificar no exemplo abaixo:

*Em 2003, recém aprovada no Curso de Letras da UFC, iniciando a disciplina Introdução à Linguística, após conversas iniciais, pergunta a aluna:
- Professora, a senhora é do Cariri, Pernambuco, ou Paraíba? De Fortaleza a senhora não é! Sua fala é diferente da nossa. (Aluna – curso de Letras da Universidade Federal do Ceará)*

Analisando o ocorrido, a aluna não reconhece o falar da professora como o compartilhado em sua cidade. No entanto, apresenta três localidades possíveis. Localidades delimitadas e reconhecidas política e geograficamente. Mas, quanto à *comunidade de fala*? Como podemos delimitar em nossas pesquisas sociolinguísticas uma comunidade de fala?

Concebendo a língua heterogênea, passível de mudança e relacionada ao contexto social da *comunidade de fala*, Labov apresenta que:

A comunidade de fala não é definida por nenhuma concordância marcada pelo 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, p. 150).

Percebemos que Labov elenca como mais importante para caracterizar uma *comunidade de fala*, doravante CF, a homogeneidade dos padrões normativos da língua, os quais são compartilhados pelo grupo. Assim, considera a consciência dos falantes no tocante aos seus valores e atitudes referentes à língua em uso.

Guy (2000) acrescenta três características à definição de *comunidade de fala*:

- i) **características linguísticas compartilhadas**; isto é, palavras, sons ou construções gramaticais que são usadas na comunidade, mas não o são fora dela.
- ii) **densidade de comunicação interna relativamente alta**; isto é, as pessoas normalmente falam com mais frequência com outras que estão dentro do grupo do que com aquelas que estão fora dele.
- iii) **normas compartilhadas**; isto é, atitudes em comum sobre o uso da língua, normas em comum sobre a direção da variação estilística, avaliações sociais em comum sobre variáveis linguísticas.

Para delimitar a região CRAJUBAR como *nossa comunidade de fala*, elencamos características linguísticas da região e as relacionamos às concepções apresentadas por Guy.

No tocante à primeira característica referente aos traços linguísticos compartilhados pelos falantes, Seraine (1987) cita, no plano linguístico, especificidades da região do Cariri. Vejamos o que o autor aponta: a) uma particularidade sintática: omissão do artigo definido antes dos nomes próprios de pessoas – “O livro é **de Joaquim**”–, diferente do que se ouve em Fortaleza – “O livro é **do Joaquim**”; b) no domínio morfológico, exemplifica com o feminino de nomes terminados em **-or**: *moradeira*, usado por moradora em outras partes do Ceará; c) no domínio fonológico, na pronúncia das oclusivas línguo-dentais /d/ e /t/ seguidas do fonema /i/ acontece uma rápida projeção da ponta da língua entre os incisivos superiores e os inferiores. Sob esse aspecto, acrescenta Carvalho, H.(2007, p.74): “como observador participante, verificamos que o falar do Cariri normalmente é bem marcado e facilmente reconhecido pelos falantes de Fortaleza, especificamente pela ausência da palatalização diante do /t/ e /d/”.

Quanto à densidade de comunicação e a utilização de normas compartilhadas, acreditamos que os fatores geográficos e socioeconômicos asseguram esse alto nível de interação e absorção dos falantes da região CRAJUBAR. Considerando que Crato, Juazeiro e Barbalha estão em constantes relações sociais e comerciais, acreditamos no compartilhamento das normas de fala entre os falantes desses municípios. Assim, elencamos em nossa pesquisa, a região CRAJUBAR como uma macro comunidade de fala.

Sobre as imbricações presentes entre Crato, Juazeiro e Barbalha, abriremos um tópico no terceiro capítulo.

O principal objetivo da Teoria da Variação e Mudança é analisar e sistematizar variantes linguísticas usadas por uma mesma comunidade de fala. Para sistematizar a

variação, a Sociolinguística laboviana faz uso de uma metodologia específica: seleção de informantes, levantamento de dados, descrição da variável e das variantes, análise de condicionamentos, encaixamento linguístico e social e posição histórica da variável.

A realização de análises quantitativas possibilita o estudo da variação linguística, permitindo, ao pesquisador, apreender sua sistematicidade, seu encaixamento linguístico e social e sua eventual relação com a mudança linguística. A variação linguística, entendida como alternância entre dois ou mais elementos linguísticos, por sua própria natureza, não pode ser adequadamente descrita e analisada em termos estritamente qualitativos. Antes do advento da metodologia de quantificação, a variação linguística era considerada secundária, aleatória ou mesmo impossível de ser cientificamente apreendida. O uso de métodos estatísticos, contudo, tem permitido demonstrar o quão central a variação pode ser para o entendimento de questões como identidade, solidariedade ao grupo local, comunidade de fala, prestígio e estigma, entre outras (GUY; ZILLES, 2007, p. 73).

Para compreender o porquê de alguém fazer uso “dessa” e não “daquela” forma de dizer alguma coisa, é importante coletar um grande volume de fala natural, e de maneira bem gravada, quer dizer, as técnicas elementares para contatar informantes e levá-los a falar livremente numa entrevista não devem ser ignoradas, visto que vários princípios do comportamento linguístico e social estão contidos nestas técnicas de procedimentos. Neste sentido, Labov, a partir de suas observações, elabora cinco axiomas metodológicos os quais foram sistematizados durante suas pesquisas de campo: alternância de estilo, a atenção, o vernáculo, a formalidade e os bons dados. Esses axiomas levam o pesquisador a um paradoxo metodológico – paradoxo do observador, a solução para esse paradoxo é um problema metodológico central (LABOV, 2008).

Vejamos mais detalhadamente a seguir cada um desses princípios apontados por Labov para orientar a pesquisa em Sociolinguística Variacionista.

Princípio 1 – “*não há falantes com um estilo único*”, todos nós usamos mais ou menos variações. Cada falante vai apresentar alguma variação de acordo com o contexto em que está inserido naquele momento. Esse falar é condicionado por: relações entre falante e interlocutor, particularmente, relações de poder e solidariedade; contexto social (casa, escola, igreja, vizinhança); tópico discursivo, nesse condicionamento, ressaltamos que quanto menor for a participação social do falante mais restritas serão suas escolhas (LABOV, 2003, p. 245).

Princípio 2 – “*todos os grupos apresentam as mesmas tendências de estilo*”, o julgamento do nível estilístico absoluto do falante não é baseado apenas em um falante, pois sabemos que, enquanto estamos perguntando e recebendo respostas, o falante está usando um estilo “consultivo”, relativamente “cuidadoso”, diferentemente daquele mais “casual” ou “íntimo” com o qual ele discute com os seus amigos, com o qual briga com sua família. Existem técnicas para a obtenção de discurso casual em uma situação de entrevista, mas a

abordagem mais sólida é observar o falante interagindo com os colegas que controlam seu discurso na vida cotidiana, quando o observador não está presente. Portanto, são regras que nascem primeiramente em um grupo social e se espalham para outros grupos⁶. (LABOV, 2003, p. 245).

Princípio 3 – *o comportamento sociolinguístico apresenta uma diferenciação social*, de modo que tal comportamento reflete um conjunto de normas, crenças ou atitudes subjetivas para com características particulares e da linguagem em geral. A estratificação regular daqueles que mais usam na fala casual, formas estigmatizadas reflete-se no fato de que são eles os que mais estigmatizam a fala dos outros (LABOV, 2003, p. 243)⁷. Labov cita, como exemplo, que

[...] os líderes do movimento nacionalista negro entre o povo negro não fazem uso da língua fora do padrão Negro Inglês em seus discursos públicos. Sua gramática é essencialmente padrão. Embora haja uma tendência crescente para usar fragmentos de língua vernácula em discursos públicos, a análise cuidadosa mostra que estes são elementos isolados: a gramática básica e fonologia utilizada é a da comunidade de classe média. [...] Por outro lado, os Philadelphians e os nova-iorquinos desprezam as suas próprias formas vernáculas. Em geral, este é um importante princípio sociolinguístico que aqueles que usam o mais alto grau de uma forma estigmatizada em seu próprio discurso informal são os que mais estigmatizam a fala dos outros (LABOV, 2003, p. 244)⁸.

Esse princípio tem consequências importantes para a situação de sala de aula. O professor de uma comunidade que usa a linguagem não padrão tem a vantagem de identificar e corrigir as formas não padrão de seus alunos; mas tem a desvantagem de reagir de modo extremo a essa linguagem.

Princípio 4 – *“há, em geral, diferenças entre os sexos”*. Mulheres são mais sensíveis à correção e tendem a usar formas de prestígio mais do que fazem os homens. Mas esta diferença não é independente da classe social (LABOV, 2003, p. 245).

Princípio 5 – *“o sistema linguístico mais consistente e regular de uma comunidade de fala é o do vernáculo básico”*, as implicações desse princípio para a análise diacrônica das variantes linguísticas são importantes. A partir dessa afirmação é fácil perceber

⁶ So far we have been considering stable sociolinguistic situations. Wherever the language is in the process of change, there is a tendency for new forms to be adapted first by one a social group and only gradually spread to others. The social value attributed to these forms is derived from the values associated with the groups which introduced them.

⁷ So far we have been considering stable sociolinguistic situations. Wherever the language is in the process of change, there is a tendency for new forms to be adapted first by one a social group and only gradually spread to others. The social value attributed to these forms is derived from the values associated with the groups which introduced them.

⁸ [...] Conversely, the Philadelphians and the New Yorkers both despise their own vernacular forms. In general, it is an important sociolinguistic principle that *those who use the highest degree of a stigmatized form in their own casual speech are quickest to stigmatize it in the speech of others.*

que é possível comparar fases distintas do vernáculo de uma comunidade pela comparação das falas de diferentes gerações (LABOV, 2003, p. 247).

Por fim, não devemos esquecer que, numa análise de pesquisa variacionista, os dados quantitativos são complementados pelos qualitativos. Isso significa que, para interpretarmos o que a distribuição das formas significa por meio dos grupos ou em contextos diferentes, precisamos de dados de uso dessa distribuição. Também é necessário sabermos quando e como podemos ir além dos números, a fim de avaliar a forma como esses modelos distribucionais estão sendo usados pelos falantes em um contexto social ou interacional específico⁹ (MEYERHOFF, 2006, p. 6). Cabe ao pesquisador ir além dos resultados estatísticos para compreender e descrever o fenômeno pesquisado.

2.2 A concordância nominal de número no português do Brasil segundo a GT e sob a concepção variacionista

Vários estudos produzidos comprovam que a *concordância nominal de número* é um fenômeno variável no português do Brasil e suas variantes apresentam-se, principalmente na fala, de diferentes maneiras com o mesmo significado referencial, conforme os exemplos a seguir apresentados na dissertação de mestrado de Scherre (1978, p. 66). As letras em parênteses ao lado dos exemplos correspondem à identificação do informante cujos dados foram utilizados por Scherre em sua pesquisa.

- a) /os garotus inglesis/ (GUI)
- b) /os **cara** entraram/ (SID)
- c) /umas cinco **colega minha...**/ (JOR)

Utilizadas por inúmeros e diferentes falantes do PB, essas variantes da marcação de plural, presentes nos exemplos b e c, se elencadas, não são consideradas nas gramáticas tradicionais como variantes de prestígio, visto que esses manuais prescrevem apenas uma variedade. Por conseguinte, usos como os apresentados em b e c geralmente são estigmatizados.

Vejam os que diz Bechara (2009, p. 543) sobre concordância nominal: “é a que se verifica em gênero e número entre o adjetivo e o pronome (adjetivo), o artigo, o numeral ou o particípio (palavras determinantes) e o substantivo ou pronome (palavras determinadas) a

⁹ “As we will see, quantitative data is complemented by what researchers call qualitative data. In practice, this means that in order to interpret what the distribution of forms means across different groups or in different contexts we need solid data on the distribution of forms; but we also have to know when and how to move beyond the numbers in order to evaluate the way in which those distributional patterns are being used by speakers in a particular social or interactional context.” (p. 239).

que se referem”. Em seguida, o autor, apresenta três formas em que a concordância nominal pode ser estabelecida.

- a) concordância de palavra para palavra: “*Os bons* exemplos dos pais são *as melhores lições* e a *melhor* herança para *os filhos*” [MM].
- b) concordância de palavra para sentido: “Sua Majestade *fidelíssima* foi contrariado pelos representantes diplomáticos”.
- c) Outros casos de concordância nominal: “Parou um momento e, olhando para um e outro lado, endireitou a carreira” (BECHARA, 2009, p. 544-547).

Segundo Brandão (2014), entre os gramáticos tradicionais da atualidade, Bechara é o que mais se detém no tocante à concordância nominal de número. Mesmo assim, as variantes apresentadas nos exemplos “b” e “c” por Scherre (1978, p. 66) não foram mencionadas por Bechara.

É fato que prescrever o contrário iria de encontro ao escrever e falar “corretamente”. Visto que, falar e/ou escrever corretamente pressupõe seguir as regras prescritas na GT. Regras importantes e necessárias, porém, passíveis de contestação, visto que a língua se manifesta de forma dinâmica entre seus usuários.

Em outros autores também encontramos semelhanças conceituais ao que diz Bechara (2009). Vejamos: “A *concordância nominal* se ocupa da relação entre os nomes, ou seja, entre as classes de palavras que compõem o chamado grupo nominal (substantivos, adjetivos, pronomes e numerais)” (CIPRO NETO; INFANTE, 2008); “*Concordância nominal* é a combinação, em gênero e número, do artigo, do adjetivo e do numeral com o substantivo a que se referem” (GIACOMOZZI; VALÉRIO; REDA, 2010); “Toda palavra variável que se refere ao **substantivo** concorda com ele em **gênero** (masculino/feminino) e **número** (singular/plural)” (AMARAL *et al.*, 2010).

Essas são normas de concordância nominal de número encontradas nas GT com prescrição dos usos “corretos” de concordância. Por conseguinte, espera-se que falantes escolarizados não utilizem variedades estigmatizadas. Com relação ao fenômeno em tela, o fator escolaridade tem sido testado amplamente para verificar o seu grau de influência sobre os falantes quanto à apropriação das normas de prestígio.

Na obra *Gramática texto: análises e construção de sentido*, as autoras (ABAURRE, L; ABAURRE, B; PONTARA; 2006) apresentam, mesmo de forma sutil, outras situações de concordância entre os elementos do sintagma nominal. Segundo as autoras, a

concordância nominal se estabelece entre núcleo de um sintagma nominal (em suas flexões de gênero e número) e todos os termos que o determinam.

Porém, fazendo uso da seguinte fala de Níquel (personagem de Fernando Gonsales), presente na tirinha *Níquel Náusea – Vá pentear macacos!*, “Gatinha **as coisa** fica brilhenta!”, Abaurre e Pontara (2006) lembram que: diante de falantes que não seguem as normas prescritas pela GT, devemos levar em consideração o contexto de ocorrência (2006, p. 479). Afirmando:

A análise das estruturas presentes na tira do Níquel Náusea nos permite constatar algo muito importante nas relações de concordância: elas podem ser organizadas a partir de princípios diferentes. Por esse motivo, quando ouvimos falantes de uma variedade estigmatizada dizerem algo como *As laranja tá madura*, não devemos concluir que essas pessoas estão cometendo erro gramatical. No sistema de concordância da variedade linguística que falam, a regra geral determina a flexão de número somente para o determinante do sintagma nominal (no caso, o artigo). Em termos do conteúdo informacional, uma fala como essa é equivalente à de um falante de outra variedade que diz “*As laranjas estão maduras*”. O importante, quando refletimos sobre formas diferentes das estabelecidas, pela gramática normativa é lembrar que elas devem ser consideradas em seus contextos de ocorrência. Em situações formais, espera-se o respeito à norma gramatical.

Sarmento (2012) também faz uma breve consideração acerca das variantes presentes na concordância nominal de número. Segundo a autora, as linguagens regional e coloquial, próprias da fala popular, são formas adequadas de expressão de acordo com a situação de uso. Nessas linguagens, a concordância nominal nem sempre segue a língua considerada padrão. O autor cita como exemplo: *Ela contou muitas história engraçada*.

Nesta abordagem, feitas em algumas gramáticas, constatamos que pouco ou quase nunca os autores comentam sobre as diferentes variáveis do fenômeno de *concordância nominal de número* descritas em várias pesquisas com o português brasileiro a luz da Teoria Variacionista.

Dessa forma, reconhecemos a importância e a validade das regras prescritas ao uso da língua. No entanto, existem outras variantes equivalentes, presentes durante os diferentes processos de interação entre os falantes, e não as considerar é não contar com as forças internas e externas que motivam ou restringem o falante a dizer de forma diferente a mesma mensagem.

Sobre a variável em estudo, Tarallo (2007) afirma que a marcação de plural encontra-se em estado de variação, apresentando duas variantes linguísticas correspondentes: (1) é a presença do segmento fônico /s/; e (2) é a ausência desse segmento, ou seja, a forma “zero”. O plural, no português, é marcado redundantemente ao longo do sintagma nominal: no determinante, no nome-núcleo e nos modificadores adjetivos.

Entre os elementos do sintagma nominal, comenta Bortoni-Ricardo (2004) que há uma tendência, no Português do Brasil, a não se fazer a concordância nominal. Diz, ainda, que muitos linguistas têm pesquisado esse fenômeno, sendo a professora Maria Marta Pereira Scherre (UnB/UFRJ) uma das pioneiras.

Acerca da temática *variação na concordância de número no português do Brasil*, diversos estudos já foram concluídos. Sob a ótica da Teoria da Variação e Mudança, tanto linguistas quanto dialetólogos reconhecem a variação e procuram explicá-la; identificando os fatores que favorecem ou refutam o uso de uma ou de outra variante.

Portanto, temos acesso a um número considerável de análise de dados que descrevem a heterogeneidade da língua. A partir dos anos de 1970, vários estudos sobre fenômenos variáveis foram produzidos no Brasil, utilizando-se de princípios teóricos e metodológicos da Teoria da Variação e Mudança.

Scherre (1994) destaca uma série de trabalhos pioneiros sob o arcabouço da Teoria da Variação. Lemle e Naro (1977), com estudo sobre concordância verbo/sujeito com dados de 20 falantes semiescolarizados da zona urbana do Rio de Janeiro. O banco de dados foi retirado do projeto “Competências Básicas do Português”, sob a coordenação da professora Miriam Lemle. Em seguida, dessa mesma amostra, Naro (1981) e Guy (1981) produzem seus trabalhos.

Após essas pesquisas iniciais, vários trabalhos foram produzidos descrevendo os falares e legitimando o fenômeno da variação no Brasil, por exemplo: Gryner (1977) pesquisa a *variação de concordância com verbos impessoais na cidade de Petrópolis* e trabalha com dados de falantes de Petrópolis; Mota (1979) descreve a fala de 34 adolescentes com 8 anos de escolarização moradores de bairros populares de Salvador; Nicolau (1984) estuda a *ausência de concordância verbal em português, utilizando-se de uma abordagem sociolinguística*; Rodrigues (1987) *pesquisa a concordância verbal no português popular em São Paulo* - analisa 40 falantes moradores em favelas de São Paulo de diferentes graus de escolarização; Graciosa (1991), em sua obra *Concordância verbal na fala culta carioca*, analisa dados de 18 universitários coletados do projeto Norma Urbana Culta (NURC) do Rio de Janeiro; Bortoni-Ricardo (1985), em *The urbanization of rural dialect speakers - A sociolinguistic study in Brazil* -, trabalha com dados de 80 falantes da zona rural da região do Alto Parnaíba em Minas Gerais radicados em Brasilândia; Scherre e Naro produzem três artigos (1999, 1992, 1993) com base em dados de 64 falantes cariocas com 1 a 11 anos de escolarização com o *Corpus Censo do Programa de Estudos sobre o Uso da Língua*.

Quanto aos trabalhos referentes ao fenômeno em estudo – *concordância nominal de número no Português brasileiro* (doravante PB) –, as primeiras pesquisas foram feitas por Braga e Scherre (1976), que analisaram dados de sete falantes residentes no Rio de Janeiro, falantes procedentes de regiões e classes sociais diferentes. Em seguida, Braga (1977) - *A Concordância de número no sintagma nominal no triângulo mineiro* - trabalha com o falar de sete informantes de classe média e baixa do Triângulo Mineiro. Em 1978, Scherre analisa dados de 10 falantes da zona urbana do Rio de Janeiro (seis semiescolarizados, alunos do MOBREAL¹⁰; três universitários; e um falante com 11 anos de escolarização); Ponte (1979) - *A concordância nominal de uma comunidade de Porto Alegre* - estudou o falar de Porto Alegre; Nina (1980) - *Concordância nominal/verbal do analfabeto na microrregião de Bragantina* - trabalhou com dados de 20 falantes da microrregião bragantina no Pará.

Scherre (1988), em sua tese de doutorado intitulada *Reanálise da Concordância Nominal em Português*, retoma o tema da concordância nominal de número e, desde então, seu trabalho tornou-se uma das referências mais importantes às inúmeras pesquisas em variação linguística na concordância nominal de número no PB.

O material linguístico analisado por Scherre (1988) foi extraído do banco de dados do *Corpus Censo* do PEUL. A amostra é constituída de 64 horas de fala, gravada com 64 falantes radicados no município do Rio de Janeiro. Essas 64 horas estão na forma de entrevistas, transcritas e armazenadas eletronicamente. Constitui-se de 48 falantes adultos (15-71 anos) e 16 crianças (7-14 anos), nas faixas etárias: 7-14, 15-25, 26-49 e 50-71.

Reconhecemos no trabalho de Scherre (1988) um marco para uma nova etapa das pesquisas sobre variação da concordância nominal de número no PB. Assim, elencaremos algumas teses e dissertações produzidas a partir do estudo de Scherre (1988). Para visualizarmos melhor as localidades em que pesquisas já foram produzidas, faremos a distribuição por regiões geográficas, de acordo com a localidade em que os dados analisados foram extraídos. Quanto aos grupos de fatores e resultados desses trabalhos, apresentaremos em momentos específicos, ao longo de nossas análises, visando correlacioná-los com nossas observações.

REGIÃO NORTE

¹⁰ Programa criado em 1970 pelo governo federal com objetivo de erradicar o analfabetismo do Brasil em dez anos. O Mobral propunha a alfabetização funcional de jovens e adultos, visando “conduzir a pessoa humana a adquirir técnicas de leitura, escrita e cálculo como meio de integrá-la a sua comunidade, permitindo melhores condições de vida”. O programa foi extinto em 1985 e substituído pelo Projeto Educar.

1. Rio Branco – AC (CARVALHO, R., 1997) – trata da descrição e análise da concordância nominal de número tendo como amostra de fala a Língua Portuguesa da classe baixa da comunidade urbana de Rio Branco (Estado do Acre), sob a perspectiva da Sociolinguística Variacionista, aliada a alguns aspectos do Funcionalismo Linguístico. Segundo a autora, embora exista uma polêmica, no Brasil, acerca da estratificação em classes sociais, para classificação em classe baixa foram considerados os seguintes fatores: analfabetos e 1º Grau (Ensino Fundamental); estudos em escola pública; trabalhar ou estudar; morar em bairros periféricos; renda de 1 a 2 salários e possuir TV, geladeira, aparelho de som e outros.

2. Alto Solimões – AM (MARTINS, 2013) – trata da análise e descrição da concordância nominal de número no falar dos habitantes do Alto Solimões (AM), à luz da Teoria da Variação e Mudança e da Dialetoologia.

REGIÃO NORDESTE

3. João Pessoa – PB (CARVALHO, H., 1997) – trata da análise e descrição da concordância nominal de número do PB falado em João Pessoa. O material linguístico analisado por Carvalho, H., foi extraído do Projeto Variação Linguística no Estado da Paraíba (VALPB).

4. Salvador – BA (LOPES, 2001) – trata da análise e descrição da concordância de número no sintagma nominal nas falas popular e culta atuais de Salvador.

5. Helvécia – BA (P. ANDRADE, 2003) – Um fragmento da constituição sócio-histórica do português do Brasil: variação na concordância nominal de número em um dialeto afro-brasileiro. Analisa as variantes na Concordância Nominal (CN), no dialeto de Helvécia (BA).

6. Vitória da Conquista – BA (GUIMARÃES, 2014) – trata da análise e descrição da concordância nominal de número no português popular do Brasil: estudo de variação e mudança no vernáculo conquistense. Analisou inquéritos coletados do *corpus* Português Popular de Vitória da Conquista (PPVC).

7. Vitória da Conquista – BA (MEIRA, 2015) – trata da análise e descrição da concordância nominal de número. Um estudo comparativo entre os dados extraídos dos *corpora* PPVC e PCVC (*Corpus* do Português Popular de Vitória da Conquista e *Corpus* do Português Culto de Vitória da Conquista, respectivamente). Os dados referentes ao Português Popular é uma análise de Guimarães (2014).

REGIÃO SUDESTE

8. Rio de Janeiro – RJ (SCHERRE, 1988) – Em *Reanálise da Concordância Nominal em Português*, Scherre descreve e analisa a indicação de pluralidade sob uma perspectiva teórica que une aspectos da Teoria da Variação Linguística Laboviana ou Sociolinguística Quantitativa e aspectos da Teoria Funcionalista.

9. Pedro Leopoldo – MG (SANTOS, 2010) – analisa a variável linguística constituída pela ausência e pela presença de concordância entre os elementos flexionáveis do sintagma nominal (SN) na fala de Pedro Leopoldo – Minas Gerais.

REGIÃO SUL

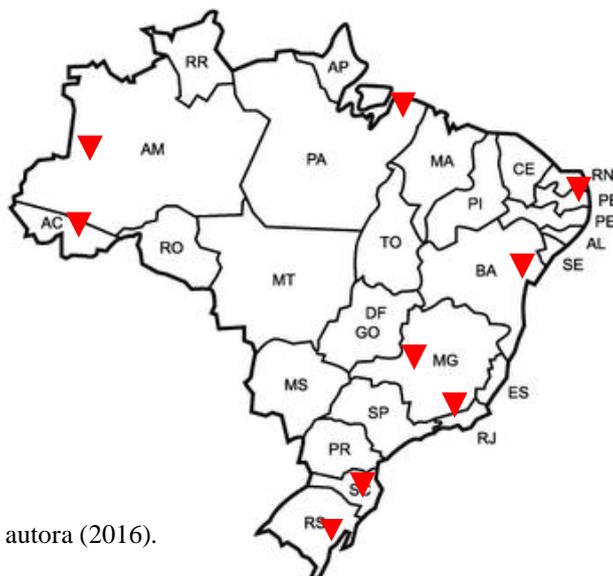
10. SC/RS/PR (FERNANDES, 1996) – analisa o comportamento da concordância de número plural no português do Brasil. O *corpus* analisado é do Projeto Variação Linguística na Região Sul do Brasil (VARISUL).

11. Tubarão-SC e São Borja-RS (L. ANDRADE, 2003) – analisa as similaridades e diferenças no comportamento do fenômeno de concordância nominal de número nas cidades da região Sul do Brasil: São Borja (cidade gaúcha de fronteira com a Argentina) e Tubarão (cidade litorânea catarinense).

12. Novo Hamburgo – RS (SCHNEIDER, 2012) – analisa e descreve o desenvolvimento da concordância nominal de número na produção oral de crianças monolíngues adquirindo o português falado no Brasil.

Podemos constatar que a variável *concordância nominal de número* no Português Brasileiro está sendo descrita em várias regiões do Brasil. Na figura 1, identificamos as cidades brasileiras onde essa variável foi descrita. Lembramos que identificamos somente os estados dos trabalhos elencados acima de acordo com as regiões brasileiras. É importante ressaltar que não destacamos todos os trabalhos produzidos a partir de Scherre (1988). No entanto, a amostra apresentada já sinaliza que muitos trabalhos já foram produzidos nesta linha. No entanto, existem vários lugares ainda a serem descritos, por exemplo, o Ceará.

Figura 1 – Distribuição das pesquisas apresentadas acerca da concordância nominal no Brasil na perspectiva variacionista. A partir de 1988, quando foi divulgado o estudo de Scherre.



Fonte: Elaborada pela autora (2016).

Os trabalhos sobre concordância nominal de número já apresentados são fundamentais para este estudo, visto que alguns grupos de fatores linguísticos e extralinguísticos controlados serão replicados aos dados do falar da região CRAJUBAR. Quanto aos resultados referentes a esses grupos de fatores, também serão utilizados como referência aos encontrados neste trabalho.

Detalharemos, a seguir, as variáveis independentes linguísticas e extralinguísticas controladas dos trabalhos mencionados a partir de Scherre (1988). Com objetivo de comprovar o interesse pelo fenômeno em estudo, demonstramos os trabalhos por ordem crescente do ano da pesquisa.

Iniciamos com a pesquisa de Scherre (1988) pelas razões já citadas anteriormente. O material linguístico analisado foi extraído do banco de dados do *Corpus* Censo do PEUL. A amostra é constituída de 64 horas de fala, gravada com 64 falantes radicados no município do Rio de Janeiro. Essas 64 horas estão na forma de entrevistas transcritas e armazenadas eletronicamente e ocorreram em dois momentos: em 1982 – 1984 foram gravadas 48 entrevistas com 48 falantes de 15 a 71 anos. Já em 1983 – 1985 gravaram-se 16 entrevistas com 16 falantes de 7 a 14 anos.

A primeira fase da coleta composta de 48 falantes foi estratificada segundo três características sociais:

- i. **Anos de escolarização:** 1 a 4 anos (antigo primário); 5 a 9 anos (antigo ginásial); e 9 a 11 anos (antigo colegial). Hoje encontramos outras

nomenclaturas correspondentes aos anos de escolarização: Fundamental I; Fundamental II e Ensino Médio.

ii. Sexo: feminino (24 mulheres) e masculino (24 homens).

iii. Faixa etária: 15 a 25 anos (adolescentes e adultos jovens que se iniciam no mercado de trabalho); 26 a 49 anos (adultos maduros) e 50 anos em diante (próximos à aposentadoria).

A segunda fase da coleta de dados complementa a primeira quanto à faixa etária.

Assim estratificada:

i. Anos de escolarização: 1 a 4 anos (8 falantes) e 5 a 8 anos (8 falantes)

ii. Sexo: feminino (8 meninas) e masculino (8 meninos)

iii. Faixa etária: 16 falantes de 7 a 14 anos.

As crianças nasceram e foram criadas no Rio de Janeiro e todas estudavam em escolas públicas. Após a coleta de todos os dados relevantes e, excluídos alguns dados, Scherre analisou 7.193 sintagmas nominais – 6.027 produzidos por adultos e 1.166 pelas crianças – sob dois tipos: uma análise *atomística*, que consiste na observação da relação entre os elementos dos sintagmas nominais e a *não atomística* com observação da relação entre os sintagmas nominais no plano oracional.

Quanto às variáveis discutidas na primeira análise – atomística – foram: *processos morfofonológicos de formação de plural, tonicidade dos itens lexicais singulares, número de sílabas dos itens lexicais singulares, posição linear do elemento no SN, Classe gramatical do elemento nominal, marcas precedentes ao elemento nominal analisado, contexto fonético seguinte ao elemento nominal sob análise/fonológico, função sintática do SN, animacidade dos substantivos, grau dos substantivos e dos adjetivos e formalidade dos substantivos e adjetivos.*

Após a codificação de todos os elementos relevantes em função das variáveis citadas, a autora obteve 13.229 constituintes, sendo 11.086 extraídos dos SNs produzidos pelos adultos e 2.143 dos SNs produzidos pelas crianças.

A seguir, alguns resultados da pesquisa de Scherre (1988), do tipo *atomística*, que são pertinentes ao estudo do nosso fenômeno.

Com referência a *processos morfofonológicos de formação do plural* de todos os dados dos adultos, evidencia mais marcas de concordância nos itens lexicais que apresentam mais diferença de material fônico entre as suas respectivas formas singulares e plurais. Os

plurais duplos (tijolo/tijolos) favorecem-nas com 0.86 e os regulares inibem-nas com 0.24. Scherre retoma a apreciação dessa variável em uma 2ª e 3ª análise.

Sobre as três variáveis: *posição linear*, *classe gramatical* e *marcas precedentes*, segundo Scherre (1988), todos os trabalhos realizados, até então, apresentam uma conclusão unânime: a primeira posição do SN é a mais marcada, num índice probabilístico nunca inferior a 0,70. As demais posições evidenciam um índice baixo de marcas. A pesquisa segue apresentando diversas relações entre classe gramatical e posição.

Quanto à variável *marcas precedentes* em função da posição, Scherre observa que a ausência de qualquer marca fora do SN provoca um alto percentual de marcas no elemento nominal da primeira posição, assim como o zero na primeira posição provoca a sua presença no constituinte da segunda, e isso sinaliza para a confirmação da hipótese funcionalista de Kiparsky (1972 *apud* Scherre 1988), visto que nesse caso há uma tendência de marcação para indicar a ideia de pluralidade. A autora nos apresenta, ao longo de seu trabalho, muitas conclusões relevantes e que serão retomadas no decorrer de nossa pesquisa.

Fernandes (1996), em sua análise, controlou as seguintes variáveis linguísticas: posição linear dos elementos no SN, classe gramatical dos elementos, relação com o núcleo do SN, marcas precedentes, processo morfofonológico de formação do plural, tonicidade dos itens, contexto seguinte, grau dos substantivos e adjetivos, animacidade dos substantivos e adjetivos variáveis extralinguísticas: idade, níveis de escolarização, sexo, etnia, níveis de formalidade. Semelhante a outros trabalhos, a primeira posição é a que mais retém a marca formal de plural e não é ocupada só por determinantes, visto que foram observados adjetivo e substantivo ocupando a primeira posição, apresentando marcas de plural bastante altas – 0.61.

Continuando nossa apresentação da variável dependente em concordância nominal de número, vejamos o trabalho de Carvalho, H., (1997). A pesquisa traz uma análise da fala de 60 informantes nascidos em João Pessoa e que compõem o Projeto VALPB, estratificados nas seguintes características sociais: 60 informantes estratificados por sexo; faixa etária: 15 - 25, 26 - 49, 26 - 49 e mais de 50 anos; escolarização: nenhum ano, 1 - 4 anos, 5 - 8 anos, 9 - 11anos e mais de 11 anos.

Adotando a perspectiva *atomística*, Carvalho H. analisa seus dados (8505) à luz das hipóteses e resultados de Scherre (1988), com o propósito de estabelecer comparações entre as duas amostras. As variáveis linguísticas consideradas foram: *posição linear ao elemento analisado*, *classe gramatical do elemento sob análise*, *classe e posição em relação ao núcleo ou relação com o núcleo*, *marcas precedentes ao elemento analisado em função da*

posição, tonicidade do item lexical singular, processos morfofonológicos de formação de plural, contexto fonético-fonológico seguinte.

Compartilhando do resultado de Scherre (1988), no tocante à *posição linear* do elemento no SN, constata Carvalho, H., (1997) que a primeira posição detém o percentual mais alto de aplicação da regra - 96%. Submetida ao programa VARB2000 que executa uma interação entre todas as variáveis e elenca a mais relevante estatisticamente, a variável posição foi eliminada. No tocante à classe gramatical, os resultados da pesquisadora mostram que as classes gramaticais que funcionam como determinantes (quantificador, possessivo, indefinido, artigo e demonstrativo) apresentam mais probabilidade de serem marcadas do que os substantivos e adjetivos.

Ao realizar o cruzamento das variáveis: posição, classe gramatical e classe e posição em relação ao núcleo, a variável classe e posição em relação ao núcleo mostrou-se estatisticamente relevante.

No que diz respeito à variável *marcas precedentes* em função da posição Carvalho H., (1997) conclui que os resultados se mostram compatíveis com os do Rio de Janeiro, confirmando assim, o princípio do Processamento do Paralelismo Formal. Semelhante ao trabalho de Scherre (1988), a autora constata que duas ou mais marcas seguidas favorecem mais a retenção de plural do que apenas uma marca precedendo o segmento analisado ou do que a presença de pelo menos um –s antecedente não mediado por zero. Carvalho H. conclui que os dados de João Pessoa apresentam mais semelhanças do que diferenças em relação à amostra do Rio de Janeiro.

Carvalho R., (1997) descreve em sua pesquisa a fala de informantes de classe baixa da cidade de Rio Branco. A pesquisa apresenta as seguintes variáveis linguísticas: *processos de formação de plural, tonicidade, número sílabas dos itens lexicais singulares, posição linear, classe gramatical, marcas precedentes, e contexto fonético/fonológico*. No tocante às variáveis sociais, temos *grau de escolarização, sexo, grau de formalismo do discurso e faixa etária*. Os resultados da pesquisa indicam os homens mais favoráveis ao uso das marcas de plural. Esse resultado levou a autora observar nas fichas dos informantes a função que eles exerciam, levando-a concluir que esses resultados podem estar relacionados à *ocupação dos informantes*.

Lopes, em 2001, acrescenta aos estudos sob o prisma da Teoria da Variação outro estudo sobre concordância nominal de número no Brasil. A autora utilizou como amostra inquiridos do Projeto Norma Urbana Culta/70 (NURC/70), com informantes de nível

universitário e parte do NURC/90 e os inquéritos do Programa de Estudos sobre o Português Popular de Salvador, o PEPP. A amostra da pesquisa é composta por setenta e quatro inquéritos, constituídos da seguinte forma: 12 informantes de escolaridade superior, todos do NURC/70, controlados nos fatores sexo e faixa etária: 25 – 35anos, 45 – 55 anos e acima de 55 anos. Já da década de 1990, analisou 48 representantes da fala popular e da fala culta: os representantes da fala popular são todos do PEPP, em dois níveis de escolaridade: vinte e quatro de cada nível: 1 a 5 anos (Ensino Fundamental), com 11anos – Ensino Médio).

Os representantes do nível de Escolaridade Superior são todos informantes do projeto NURC/90, num total de 18 com faixa etária: 25 – 35 anos, 45 – 55 anos e acima de 55anos.

A Pesquisa de Lopes apresenta as seguintes variáveis linguísticas: *processos de formação de plural, tonicidade, posição linear, posição em relação ao núcleo, grau dos substantivos e adjetivos, marcas precedentes, e contexto fonológico*. No tocante às variáveis sociais tradicionais - *escolaridade, gênero e faixa etária*, Lopes (2001) acrescenta as variáveis *tempo e etnia*.

Andrade (2003) acrescenta aos estudos sociolinguísticos outra pesquisa sobre o fenômeno aqui analisado. A pesquisadora coletou seus dados dos informantes de duas cidades: Tubarão (SC) e São Borja (RS). Os dados destes municípios foram retirados do Banco de Dados de fala do Projeto VARSUL – Variação Linguística Urbana da Região Sul –, e os de Tubarão (SC) constituem também amostras de textos orais, cedidas pelo PROCOTEXTOS/AMUREL (Projeto de Coleta de Textos de informantes da AMUREL).

Foram coletados um total de 3.511 dados, 1896 de Tubarão e 1615 de São Borja, estratificados em 24 informantes: 12 de Tubarão e 12 de São Borja.

As variáveis linguísticas estudadas por L. Andrade foram: *posição dos elementos no sintagma nominal, classe gramatical dos elementos, relação com o núcleo do sintagma nominal, marcas precedentes, processo morfofonológico de formação do plural, tonicidade dos itens, grau dos substantivos e adjetivos*. As variáveis extralinguísticas foram: *idade, nível de escolaridade, sexo e cidade*.

Santos (2010) analisa um *corpus* constituído por dados extraídos de 27 (vinte e sete) entrevistas realizadas, utilizando-se de uma amostra composta por pessoas de três diferentes grupos sociais (classes alta, média e baixa) e distribuídas em três faixas etárias: jovem - de 17 a 23 anos; adulto - de 40 a 47 anos; idoso - acima de 60 anos; níveis de escolaridade: Ensino Fundamental, Ensino Médio e Ensino Superior. Das entrevistas com os

27 informantes, a pesquisadora coletou 1.297 sintagmas nominais. Após codificação, 1.461 dados foram submetidos a uma análise quantitativa com a utilização do programa VARBRUL.

Quanto às variáveis linguísticas, a pesquisadora controlou: 1. *Elemento nuclear do SN: posição (primeira, segunda, terceira)* 2. *Elemento nuclear do SN: classe gramatical (substantivo, não-substantivo)* 3. *Elemento não-nuclear do SN: posição (primeira, segunda, terceira)* 4. *Elemento não-nuclear do SN: classe gramatical (adjetivo, artigo, demonstrativo, possessivo, quantificador, indefinido)* 5. *Elemento não-nuclear do SN: presença e ausência de flexão de plural*

Por sua vez, quanto às variáveis sociais: *faixa etária; sexo; escolaridade; grupo social dos informantes: C (baixo), B (médio), A (alto).*

Os resultados obtidos nessa análise mostraram que a ACN (ausência de concordância nominal de número) ocorre em 759 casos, ou seja, 52% dos dados analisados. Concluem também que a variável em estudo não representa um caso de mudança em progresso, mas caracteriza-se como um caso de variável estável.

Schneider, em 2012, traz resultados bastante singulares. O material linguístico analisado pela pesquisadora parte de um *corpus* obtido por meio da gravação audiovisual da fala de 30 crianças, em uma amostra estratificada por gênero – meninos e meninas – e faixa etária – 3 a 4 anos de idade; 4,1 a 5 anos de idade; e 5,1 a 6 anos de idade. Todas as crianças pertencem a famílias de escolaridade média e, principalmente, superior, de classe social média-alta. A metodologia de análise teve como base o Programa GoldVarb X.

Quanto às variáveis linguísticas independentes desse trabalho, temos as seguintes: *processos morfofonológicos de formação do plural, tonicidade do item lexical singular, posição linear dos elementos no SN, classe gramatical, contexto fonológico/fonético seguinte, grau dos substantivos e adjetivos, posição do elemento em relação ao núcleo do SN.*

Martins, em 2013, desenvolveu sua pesquisa em concordância nominal de número à luz da teoria variacionista. Seus dados foram coletados das entrevistas com 57 informantes em cinco das nove localidades pertencentes à microrregião do Alto Solimões (São Paulo de Olivença, Santo Antônio do Içá, Tocantins, Jutai e Fonte Boa). Foram coletados 4.458 sintagmas nominais dos cinco municípios investigados. Após a categorização de cada elemento dos sintagmas, finalizou com 7.270 dados submetidos ao programa estatístico GoldVarb X 2001.

Com o propósito de contribuir para o conhecimento das áreas dialetais brasileiras por meio de um registro sistematizado do falar amazonense, Martins (2013) controlou em sua pesquisa as seguintes variáveis independentes linguísticas: *posição em relação ao núcleo/núcleo, posição linear, classe gramatical, processos morfofonológicos de formação de plural e tonicidade dos itens lexicais, marcas precedentes, contexto fonético-fonológico subsequente e características dos itens lexicais*. E as variáveis independentes extralinguísticas: *idade, escolaridade, sexo/gênero, diatopia, ocupação, mobilidade e localismo*.

Levando em consideração a rodada estatística sem a variável classe gramatical, todas se mostraram significativas na regra de funcionamento da concordância nominal de número na microrregião do Alto Solimões.

Guimarães, em 2014, fez sua análise sob uma abordagem atomística, utilizando o *corpus* do Português Popular de Vitória da Conquista (PPVC) - constituído pelo Grupo de Pesquisa em Linguística Histórica e Sociofuncionalismo (UESB). O *corpus* é composto de 24 inquiridos, dentre os quais, a autora selecionou 12 informantes. Foram coletadas do *corpus* do PPVC 2.979 ocorrências pluralizáveis submetidas ao Programa Estatístico GoldVarb X 2001. Os resultados apresentaram 57,3% marcas de concordância, enquanto 42,7% não apresentaram marcas de concordância entre os elementos formadores do sintagma nominal.

A pesquisadora controlou as seguintes variáveis independentes linguísticas: *posição linear do constituinte; posição do constituinte com referência ao núcleo do SN; classe gramatical do constituinte; e saliência fônica*. No tocante aos condicionamentos sociais ou variáveis sociais, Guimarães (2014) controlou os seguintes fatores: *faixa etária; sexo (ou gênero); estada fora da comunidade; nível de letramento; e exposição à mídia*.

Meira, em 2015, envolvendo a variável dependente *concordância nominal de número*, teve por objetivo realizar um estudo comparativo entre as duas modalidades do português brasileiro – português popular e culto de Vitória da Conquista. Para o estudo, utilizou 24 (vinte e quatro) entrevistas - 12 (doze) de falantes do português culto, analisadas pelo autor e 12 (doze) de falantes do português popular, analisadas por Guimarães (2014). A pesquisa apresenta a seguinte estratificação:

- i. Sexo:** masculino e feminino.
- ii. Faixa etária:** faixa I (20 a 35 anos); faixa II (36 a 50 anos).
- iii. Nível de escolaridade:** 5 anos de escolarização (falantes do português popular) mais de 11 anos de escolarização (falantes do português culto).

O pesquisador controlou inicialmente as seguintes variáveis linguísticas independentes: *posição linear do constituinte*, *posição do constituinte*, *saliência fônica*. Como variáveis sociais: *faixa etária*, *sexo* e *escolaridade*.

No tocante à variável posição linear do constituinte, os resultados de Meira (2015) não deixam dúvidas de que a primeira posição tende a favorecer a retenção de marca de plural.

Constatamos que muitos trabalhos foram produzidos nessas últimas quatro décadas sobre a variável *concordância nominal de número no português brasileiro* à luz da Teoria da Variação Laboviana. Ressaltamos que foram elencadas somente algumas pesquisas, e o critério de escolha além do período foi o acesso a elas. Averiguamos que algumas variáveis estão presentes em vários trabalhos e que muitos dos resultados são semelhantes e outros divergem. No entanto, todos trazem sua contribuição à descrição da língua portuguesa nesse país continental, o Brasil.

3 PROCEDIMENTOS METODOLÓGICOS

Neste capítulo, tecemos considerações sobre os procedimentos metodológicos da análise dos dados. Na primeira seção, apresentamos o banco de dados do PROFALA, a coleta dos dados, o perfil dos informantes e um breve panorama sócio-histórico da região do Cariri cearense. Na segunda seção, apresentamos a caracterização dos dados, a variável dependente *presença da marca explícita de plural /s/ vs ausência da marca de plural /Ø/* e as variáveis linguísticas e sociais. Por fim, abordaremos o tratamento estatístico a que foram submetidos os dados.

3.1 O banco de dados PROFALA

Para esta pesquisa, utilizamos o banco de dados do PROFALA.¹¹ Os dados de fala foram extraídos das entrevistas presentes no *corpus O Português falado no Ceará*, especificamente, da Região Metropolitana do Cariri cearense.¹² O *corpus* foi coletado em cidades da Região do Cariri, especialmente, Crato, Juazeiro, Barbalha, Nova Olinda, Várzea Alegre, Altaneira, Mauriti, Caririaçu e Brejo Santo. O PROFALA tem por objetivo geral disponibilizar um banco de dados sobre o português falado no Ceará, que torne acessível à análise descritiva sob a perspectiva dos aspectos fonético-lexicais, morfossintáticos, semânticos e pragmáticos da fala cearense, numa visão sociolinguística e discursiva, além de outras fases da história da língua. Ressaltamos que para efeito deste estudo serão considerados as cidades do Crato, Juazeiro e Barbalha.

O banco de dados do PROFALA – coletado no início da década de 1990 até 2001, aproximadamente – está em forma de entrevistas realizadas com cada informante, contendo questões sobre juízos pessoais do informante, tais como, política, saúde e educação.

Segundo Labov (2008), o método básico para a obtenção de uma grande quantidade de dados confiáveis da fala de uma pessoa é a entrevista individual gravada, sendo

¹¹ O Projeto Variação e Processamento da Fala e do Discurso: Análises e Aplicações – PROFALA – é coordenado pelas Professoras Maria Elias Soares e Maria do Socorro Aragão, ora em desenvolvimento no Programa de Pós-Graduação em Linguística da Universidade Federal do Ceará (UFC), em parceria com o Programa de Pós-Graduação em Teleinformática, também da UFC. O Grupo tem por objetivo geral a implantação de um sistema baseado em tecnologia da informação para análises e aplicações à língua falada ao discurso.

¹² A Região Metropolitana do Cariri cearense possui nove municípios, com 5460km²; de acordo com o Instituto Brasileiro de Geografia e Estatística (IBGE, 2014), a população da região é de 590.209 habitantes e tem como área de influência a região da divisa entre o Ceará e os estados de Pernambuco, Paraíba e Piauí.

ela uma fala pública – monitorada e controlada em resposta à presença de um observador externo.

O corpus *O Português falado no Ceará* está composto por entrevistas com 190 falantes – homens e mulheres – distribuídos em três faixas etárias: 15-25 anos; 26-49 anos e 50 anos ou mais. Encontramos também uma distribuição geográfica dos informantes em zona rural e zona urbana. Há cinco níveis de escolarização: nenhuma escolarização - 0; 1-4 anos; 5-8 anos; 9-11 anos; e 11 anos de escolarização em diante. Todas as entrevistas perfazem uma média de 85 horas de fala transcritas, armazenadas e disponíveis ao público pelo site: www.profala.ufc.br.

3.2 Perfil dos informantes da amostra

De acordo com Guy; Zilles (2007), em estudos linguísticos de comunidades de fala, um procedimento de amostra aleatória nem sempre é o ideal, mas qualquer que seja o método escolhido o importante é contemplar a representatividade. Isso significa dizer que a amostra deve ter significância para que possam estimar com certo grau de segurança o universo de onde a amostra foi retirada. Assim, analisamos os sintagmas presentes nas entrevistas com 24 informantes, perfazendo aproximadamente 12 horas de fala, com média de 30 minutos por gravação. A gravação de menor tempo (NBS – 38) teve 9 minutos, aproximadamente, e a maior (JBX – 28), com 60 minutos. Nossa estratificação apresenta-se da seguinte forma: 24 inquéritos de doze homens e doze mulheres da região CRAJUBAR (Crato, Juazeiro e Barbalha) distribuídos em três faixas etárias – seguindo as faixas etárias presentes no PROFALA. Quanto ao nível de escolarização, analisamos dois níveis: de 1 a 8 anos de escolarização e de 9 a 11 anos de escolarização.

Quadro 2 – Estratificação social dos informantes do CRAJUBAR

Faixa Etária	15 – 25		26 – 49		50 ou mais	
	H	M	H	M	H	M
1 a 8 anos de escolarização	2	2	2	2	2	2
9 a 11 anos de escolarização	2	2	2	2	2	2
24 Informantes						

Fonte: Elaborado pela autora (2016).

Apresentamos a seguir, no quadro 3, as características sociais dos falantes, com as respectivas abreviaturas de identificação e numeração apresentadas no *corpus* (APSN-19), sexo, escolaridade e faixa etária. Os códigos, no parêntese, servem como referência aos exemplos apresentados ao longo da pesquisa.

Quadro 3 – Características sociais dos 24 falantes do CRAJUBAR – todos pertencentes a zona urbana.

Identificação	Sexo	Escolarização	Faixa etária	Ocupação
APSN – 19	M	1 a 8 anos	15 a 25 anos	Estudante
JACS – 17	M	1 a 8 anos	15 a 25 anos	Pintor
JBX – 28	M	9 a 11 anos	15 a 25 anos	Estudante
ARSS – 20	M	9 a 11 anos	15 a 25 anos	Balconista
IGA – 10	M	1 a 8 anos	26 a 49 anos	Agricultor
JEBB – 11	M	1 a 8 anos	26 a 49 anos	Aux. de manutenção
MBS – 12	M	9 a 11 anos	26 a 49 anos	Agente de saúde
AMA – 63	M	9 a 11anos	26 a 49 anos	Atendente de farmácia
VCM – 87	M	1 a 8 anos	50 em diante	Agricultor
LF – 119	M	1 a 8 anos	50 em diante	Operador de sistema de abastecimento
AT – 60	M	9 a 11anos	50 em diante	Comerciante
JWP – 85	M	9 a 11 anos	50 em diante	Agricultor
NBS – 38	F	1 a 8 anos	15 a 25 anos	Estudante
MJNS – 42	F	1 a 8 anos	15 a 25 anos	Doméstica
CVSPA -179	F	9 a 11anos	15 a 25 anos	Comerciante
MSMA – 125	F	9 a 11anos	15 a 25 anos	Doméstica
MIS – 71	F	1 a 8 anos	26 a 49 anos	Doméstica
MISS – 165	F	1 a 8 anos	26 a 49 anos	Doméstica
MGT – 166	F	9 a 11 anos	26 a 49 anos	Agente comunitária
MJS – 72	F	9 a 11 anos	26 a 49 anos	Professora
ADS – 97	F	1 a 8 anos	50 em diante	Dona de casa
ALA – 98	F	1 a 8 anos	50 em diante	Enfermeira
MCSR – 112	F	9 a 11 anos	50 em diante	Doméstica
MTMT – 99	F	9 a 11 anos	50 em diante	Professora aposentada

Fonte: Elaborado pela autora (2016).

No tocante à composição do quadro 3 – características dos informantes – esclarecemos que, no Crato havia apenas três informantes do sexo feminino, com escolarização de 1 - 8 anos e faixa etária de 15 - 25 anos. Com o objetivo de completar uma célula, acrescentamos o único informante – *MSMA-125* – que fugia apenas ao fator faixa etária apresentando uma diferença apenas de um ano, pois ele se encontrava com 26 anos. Consideramos que esta mínima diferença etária não acarretará danos à confiabilidade dos resultados.

3.3 A região do Cariri cearense

A figura que segue ilustra o espaço geográfico onde fica situada a região do Cariri cearense. Como podemos observar, essa região está localizada no interior do Ceará, distante aproximadamente 550 quilômetros da capital Fortaleza e tem como área de influência a região sul do Ceará e a região da divisa entre Ceará e Pernambuco.

Figura 2 – Região do Cariri cearense



Fonte: Wikipédia.org/Região do Cariri (2015).

3.3.1 Formação histórica

De acordo com o Plano Territorial do Desenvolvimento Rural e Sustentável (2010), as terras localizadas no sopé da Chapada do Araripe eram habitadas pelos índios *kariris*, antes da chegada dos portugueses ao interior brasileiro, durante o século XVII. Os

integrantes das caravanas, militares e religiosos, mantiveram os primeiros contatos com os nativos, estudaram toda a região do *Kariri*, catequizaram os indígenas e os agruparam em aldeamentos ou missões.

Os resultados desses contatos e descobrimentos desencadearam notícias de que, na região, havia ouro em abundância, ocasionando uma verdadeira corrida para os sertões brasileiros. Famílias oriundas de Portugal sonhavam com as riquezas de terras inexploradas e com a esperança de encontrar o minério que as levaria a aumentar o seu patrimônio material, além de aumentar o seu prestígio pessoal com a corte portuguesa.

A busca do metal precioso, nas ribanceiras do Rio Salgado, trouxe para a região do Cariri a colonização e, como consequência, a doação de sesmarias. Isso permitiu o surgimento de lugarejos e vilas, que mais tarde se tornariam municípios, cada um com um processo distinto de formação.

O mapa seguinte demonstra os municípios da região do Cariri cearense os quais serão apresentados, em seguida, inseridos em seu microterritório.

Figura 3 – Mapa do território do Cariri (Municípios da Região do Cariri cearense)



Fonte: Adaptado do IPECE (2009).

O Território da Cidadania do Cariri corresponde a uma área de 16.350,40 km², localizada na região sul do estado do Ceará, zona semiárida, tendo como limites ao sul, o estado de Pernambuco; a oeste, o estado do Piauí; a leste, o estado da Paraíba e ao norte, os municípios de Aiuaba, Saboeiro, Jucás, Cariús, Cedro, Lavras da Mangabeira e Ipaumirim. O território abrange 28 (vinte e oito) municípios, divididos em três microterritórios, a saber:

- *microterritório Cariri Central*, com uma área de aproximadamente 5.099,7km², formado pelos municípios de Abaiara, Barbalha, Caririaçu, Crato, Farias Brito, Granjeiro, Jardim, Juazeiro do Norte, Missão Velha e Várzea Alegre;
- *microterritório Cariri Leste*, com uma área de 4.656,1km², formado pelos municípios de Aurora, Barro, Brejo Santo, Jati, Mauriti, Milagres, Penaforte e Porteiras;
- *microterritório Cariri Oeste*, com uma área de aproximadamente 5.186,1km², é formado pelos municípios de Altaneira, Antonina do Norte, Araripe, Assaré, Campos Sales, Nova Olinda, Potengi, Salitre e Santana do Cariri.

O *corpus* coletado – *O Português falado no Ceará* –, que serve de base para coleta de dados para nosso estudo, contemplou o microterritório Cariri Central (Barbalha, Caririaçu, Crato, Juazeiro do Norte, Várzea Alegre), o microterritório Cariri Leste (Mauriti, Brejo Santo) e o microterritório Cariri Oeste (Nova Olinda e Altaneira), cobrindo, assim, a região do estado denominada de Cariri. Ressaltando que os dados coletados para nossa pesquisa estão limitados às cidades denominadas CRAJUBAR – Crato, Juazeiro e Barbalha. A seguir, apresentamos essas cidades com mais detalhes por contemplarem a amostra deste estudo.

3.3.2 CRAJUBAR

O processo de formação sócio-político e territorial do Vale do Cariri, doravante Cariri cearense, remonta ao século XIX. Segundo Queiroz (2014), a primeira tentativa de formação de uma unidade federada na zona sob a influência de Crato data de 1828. Em 1839, o senador do Império, José Martiniano de Alencar, apresentou para a Assembleia Geral Legislativa o projeto de criação de uma nova Província que se denominaria “Província do Cariri Novo” cuja capital, por sua vez, seria a Vila de Crato. Um século depois da iniciativa do Senador Alencar, o Cariri volta à cena na tentativa de ser um Estado da Federação. Neste

período, 1957, o Cariri cearense possuía uma população de 751.000 habitantes, uma área de 96.560 km².

De acordo com Queiroz (2014), as justificativas para essa junção se baseavam na condição geoambiental da região e de prosperidade em meio ao sertão árido do Nordeste. A partir dos anos 1970, a região começava a desenvolver seu processo de urbanização, com ocupação mais intensiva das atividades urbanas nos territórios dos municípios de Crato, Juazeiro do Norte e Barbalha, isso aconteceu pela relação de proximidade geográfica, interdependência e complementaridade entre esses municípios. Por conseguinte, viabilizou a formação do aglomerado urbano CRAJUBAR. Sem data oficial de quando o nome passou a ser utilizado, vale ressaltar que essa aglomeração resultou no processo de conurbação das três principais aglomerações urbanas do Cariri, legitimadas pelo espaço físico, crescimento demográfico das três cidades e a integração socioeconômica do CRAJUBAR.

Crato abrigava excelentes estabelecimentos de educação básica, uma Escola Agrotécnica Federal e a Universidade Regional do Cariri (URCA) – única instituição de ensino superior da região até o final dos anos 1990 –, além de clubes de lazer e diversão. Barbalha, no mesmo período, apontava como um próspero centro industrial, com destaque para a sede da maior usina de açúcar e álcool da região, Usina Manoel Costa Filho, em operação até o princípio dos anos 2000. Sedia a principal unidade de saúde da região, o Hospital e Maternidade São Vicente de Paulo. Já Juazeiro do Norte se destacava desde os anos 1960 como o maior empório comercial da região. Singularidades que proporcionam a essas cidades uma unidade.

Nessa concepção de unidade, por força da Lei Complementar Estadual, aprovada pela Assembleia Estadual do Ceará e sancionada em ato público na cidade de Crato, em 29 de junho de 2009, pelo Governador do Estado do Ceará, foi criada a Região Metropolitana do Cariri, compreendida pelos três municípios polos do CRAJUBAR (Crato, Juazeiro do Norte e Barbalha) e mais seis municípios limítrofes dessa aglomeração urbana, a saber, Caririaçu, Farias Brito, Jardim, Missão Velha, Nova Olinda e Santana do Cariri.

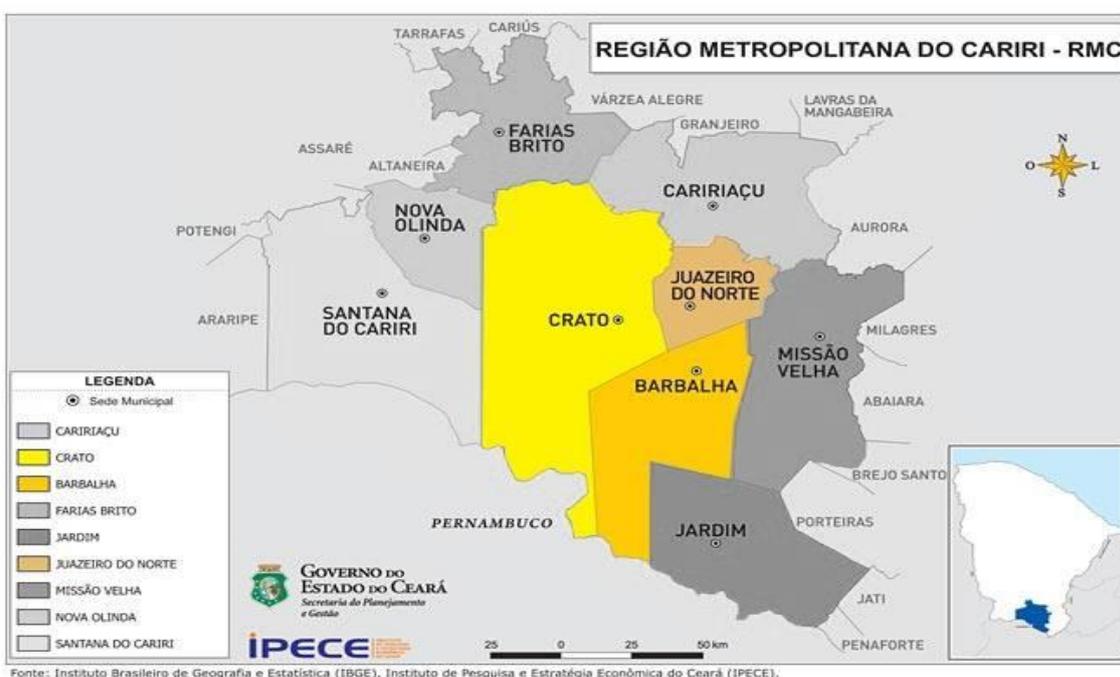
Esta decisão tornou-se realidade através da Lei Complementar N° 78, de 29 de junho de 2009 (DOE, 03/07/2009). Ressaltamos que essa nova realidade começou a ser sinalizada a partir da segunda metade dos anos 2000.

Na última década, a região do Cariri, especialmente a região CRAJUBAR, tornou-se um espaço de grande referência para empreendimentos públicos e privados, além de outras

realizações, antes, exclusivas aos grandes centros metropolitanos como Fortaleza. (QUEIROZ, 2014)

A Região Metropolitana do Cariri (RMC) é composta por nove cidades: Crato, Juazeiro, Barbalha, Cariri, Farias Brito, Jardim, Missão Velha, Nova Olinda e Santana no Cariri. Por essas singularidades apresentadas e especialmente pelas características defendidas no Cap. 2 – Fundamentação Teórica –, a região CRAJUBAR – Crato, Juazeiro e Barbalha – constituem as cidades que compõem a amostra de fala do Cariri em estudo conforme pode ser observada na figura 4.

Figura 4 – Adaptação ao mapa da Região Metropolitana do Cariri



Fonte: Instituto Brasileiro de Geografia e Estatística (IBGE), Instituto de Pesquisa e Estratégia Econômica do Ceará (IPECE).

Os dados coletados das cidades denominadas CRAJUBAR foram submetidos ao tratamento estatístico GoldVarb X. Acerca desse programa, detalharemos na quinta seção 3.5.

No que se refere à análise para a microrregião do Cariri cearense, apontam os autores que as cidades de Juazeiro do Norte, Crato e Barbalha possuem consecutivamente os três maiores aglomerados no que diz respeito à população total e às maiores densidades demográficas da microrregião do Cariri cearense. (FREIRE, FERREIRA; LIMA 2014, p.83). Esse dado põe em destaque essas três cidades elevando-as a uma condição urbana.

3.4 Caracterização dos dados

Com a finalidade de descrevermos o comportamento do fenômeno concordância nominal de número, coletamos todos os sintagmas nominais com marca formal ou explícita de plural e também aqueles que apresentam marca semântica de plural.

Quadro 4 – Os sintagmas nominais considerados

(1) Com marca formal ou explícita de plural em todos os elementos flexionáveis do SN.	“[...] vivia antes qual era <i>os seus costumes</i> ’ o que ele [...]” (JBX-28)
(2) em alguns dos elementos do SN	“[...] perco niuma missa’ <i>todos os domingo</i> eu vô [...]” (MCSR-112)
(3) em apenas um elemento do SN	“[...] devido <i>as energia negativa</i> que o marido penso [...]” (ADS-97)
(4) Sintagmas constituídos...	a) por expressões partitivas “ <i>a maioria dos homens</i> são machistas [...]” (JBX-28) “[...] <i>um dos homens mais verdadeiros</i> [...]” (JWP – 85)

Fonte: Elaborado pela autora (2016).

Os constituintes *a maioria dos; uma das; um dos*, para categorização, consideramos os constituintes como um bloco só. Para tal procedimento, analisamos o que diz Silva e Koch (1989):

[...] funcionam como pré-determinantes, certos tipos de expressões indefinidas – quantificadores universais (*todos os meus amigos, nenhum dos meus amigos*) ou partitivos (*alguns de meus amigos, muitos dos meus amigos, quatro dos meus amigos, a maioria dos meus amigos*). (p. 17-18)

Coletamos todos os sintagmas nominais presentes no falar dos 24 informantes expressos nas entrevistas. No entanto, foram excluídos os dados exemplificados no quadro abaixo.

Quadro 5 – Sintagmas excluídos

(1) A locução prepositiva <i>às vezes</i> e suas variações: <i>às veiz; às vez; às veze</i> .	“[...] casamento <i>às vez</i> dura dez ano [...]” (AMA – 63) “[...] todos os dias <i>às vezes</i> ele diz [...]” (MISS – 165)
(2) Sintagmas nominais que tivessem em seu núcleo nomes invariáveis.	“[...] na barriga <i>na:s costas</i> foi um na barriga [...]” (JBX-28)
(3) As locuções adjetivas e afins que constituem os sintagmas nominais	“[...] socialização é: é o: o <i>os grupo de de produçãozian</i> [...]” (MGT-166) “[...] fuma’ num tem <i>vícios de bebedeiras</i> e nem de jogo [...]” (MISS-165)

Fonte: Elaborado pela autora (2016).

Esclarecemos que houve certa dificuldade em encontrarmos, na literatura, uma indicação acerca da flexão de número dessas manifestações. Ressaltamos que somente os constituintes em negrito não foram analisados. Acreditamos que podem ser objeto de estudos posteriores.

“[...] socialização é: é o: o os grupo de **de produção**zian[...]”(MGT-166)

“[...] fuma’ num tem vícios **de bebedeiras** e nem de jogo [...]” (MISS-165)

3.5 O tratamento estatístico dos dados

No desenvolvimento desta pesquisa, utilizamos um método estatístico que permitiu transformar dados em informações quantitativas. A análise desses dados nos permitiu definir quais as diferentes ocorrências de indicação de pluralidade no falar do Cariri cearense, especificamente, no CRAJUBAR.

Objetivando maior compreensão do fenômeno de marcas de plural entre os elementos do sintagma nominal, fizemos nossa análise considerando todos os constituintes como unidade de análise, perspectiva chamada por Scherre (1988) de *atomística*.

Após estratificação do *corpus* e organização das células, e cientes dos grupos de fatores linguísticos e extralinguísticos que foram analisados – como possíveis condicionadores da regra variável de concordância nominal de número no falar do Cariri cearense –, coletamos no *corpus* escolhido dados da regra variável. Em seguida, fizemos a codificação dos dados e, a cada fator de análise, atribuímos um código. Os códigos foram digitados em arquivo de texto, o qual foi salvo dentro da pasta do GoldVarb X. Em seguida, realizamos as rodadas

Uma importante ferramenta quantitativa usada nos últimos anos é o programa de regra variável Varbrul – do inglês *Variable Rules Analysis*. Constitui-se de um conjunto de programas computacionais de análise multivariada, especificamente estruturado para investigar dados de variações sociolinguísticas. Análise *multivariada* significa que o programa é capaz de investigar situações em que o fenômeno em estudo é influenciado por múltiplas variáveis independentes ou grupo de fatores. (GUY; ZILLES, 2007, p.105).

O Varbrul também viabiliza ao pesquisador testar inúmeras hipóteses, sobre a natureza, tamanho e direção dos condicionamentos das variáveis independentes. Todas as versões do Varbrul trabalham com nível de significância estatística de 0.05, equivalente ao risco que se corre de rejeitar a hipótese nula. A hipótese nula postula que nenhuma das variáveis independentes construídas dá conta de explicar a variável analisada.

O GoldVarb X é a última versão do Varbrul para o ambiente *Windows*. O programa tem a vantagem de trabalhar com quantidades ilimitadas de fatores em cada variável independente e sem limites de células. No entanto, o GoldVarb X só realiza análise de pesos relativos de duas variantes (binominal). No que se refere ao peso relativo, ele não possui o módulo de três ou mais variantes (SCHERRE, 2012)

Existem muitos métodos analíticos quantitativos além do Varbrul, porém Guy e Zilles (2007) apresentam três vantagens que fazem do GoldVarb uma boa opção para os sociolinguistas. São elas: ele é dedicado à estruturação dos dados que encontramos na linguagem natural; muitas das células numa rodada típica do Varbrul não têm nenhum dado, ou se têm é apenas um. Esse programa tolera muito bem esses desvios; por fim, ele vem com rotinas que permitem recodificação e outros manuseios dos dados.

Em linhas gerais, o programa fornece resultados, em termos de frequência e probabilidade, dos grupos de fatores linguísticos e extralinguísticos que condicionam ou não o uso do fenômeno em estudo.

3.6 A variável dependente e as variáveis linguísticas

No início de nosso trabalho, já mencionamos que a variável linguística concordância nominal de número se encontra em estado de variação, apresentando entre os elementos do sintagma nominal diferentes possibilidades de presença ou não da marca de plural /s/ com o mesmo valor de verdade. Estas possibilidades não se manifestam aleatoriamente. O uso de uma ou outra variante pelos falantes está condicionado por fatores linguísticos e extralinguísticos.

Como já havíamos mencionado no PB a *concordância nominal de número* pode acontecer das seguintes formas:

- (1) Presença de marca explícita de plural /s/:

“[...] antes qual era *os seus costumes*’ o que ele fazia [...]” – padrão – (JBX-28)

- (2) Ausência de marca explícita de plural /s/

“[...] término de *cinco quilômetro*’ é razoável’ é melhor [...]” (ARSS- 20)

Foram controlados os seguintes grupos de fatores na pesquisa: *posição linear dos elementos no sintagma, classe e posição em relação ao núcleo, classe gramatical do elemento sob análise, marcas precedentes ao elemento analisado em função da posição, processos*

morf fonológicos de formação de plural, tonicidade dos itens lexicais singulares. A seguir iremos detalhá-los.

3.6.1 As variáveis linguísticas

Iniciamos com o detalhamento das variáveis linguísticas: *posição linear dos elementos no sintagma, classe gramatical do elemento sob análise e classe e posição em relação ao núcleo.*

3.6.1.1 Três variáveis linguísticas: posição linear dos elementos no sintagma, classe gramatical do elemento sob análise e classe e posição em relação ao núcleo

De acordo com Scherre (1988), a variável *posição linear* tem sido normalmente caracterizada em função do local que o constituinte analisado ocupa no SN no sentido estritamente linear. Afirma, ainda, haver uma conclusão uniforme, até a realização de sua tese, que a primeira posição é a que mais conserva a presença explícita de marca de plural /s/. Scherre (1988) acredita que essa conclusão acerca da primeira posição seja de cunho funcionalista, no sentido de Kiparsky. “Já é de nosso conhecimento que as *Condições de Distintividade* de Kiparsky estabelecem que há uma tendência para a informação semanticamente relevante ser retida na estrutura superficial” (KIPARSKY, 1972, p. 195 *apud* SCHERRE, 1988 p. 147).

A conclusão apontada por Scherre, quanto ao favorecimento de marcação de plural na primeira posição, perdura, até então, nos trabalhos Carvalho H., (1997); Carvalho R., (1997); Lopes (2001); Santos (2010); Martins (2013); Guimarães (2014) e Meira (2015).

O quadro 6 a seguir apresenta exemplos referentes a cada posição e em negrito, o constituinte analisado.

Quadro 6 – Posição linear do elemento no SN

Posições	Ilustrações
1ª Posição	“[...] o colégio’ se foi os professores’ mas eu ago:re [...]” (APSN-19)
2ª Posição	“[,,] ao terminá meus estudos eu [...]” (ARSS -20)
3ª Posição	“[...] longe’ né” com esses transporte fácil ’ aí ficô uma cidade [...]” (ADS – 97)
4ª Posição	“[...] a M.A. casô’ e os dois filho homem também vive. [...]” (ADS – 97)

5ª Posição	“[...] foram as principais coisas mais <i>importante</i> na [...]” (JBX -28)
------------	--

Fonte: Elaborado pela autora (2016).

No tocante à variável *classe gramatical*, Scherre (1988) destaca que a influência da classe gramatical sobre a variação de marcas de plural no sintagma nominal foi cuidadosamente estudada por Cedergren (1973) e Poplack (1980). De acordo com esses pesquisadores, os determinantes inibem a forma (0) zero, substantivos e adjetivos favorecem-na. Sobre a influência da variável *classe gramatical* em nosso fenômeno de estudo, Cedergren (1973) e Poplack (1980) não estão acordados em tudo. Aquela observa que os adjetivos têm mais chances de reterem a marca de plural do que os substantivos. Esta, ao contrário, verifica que os adjetivos têm menos chances de serem marcados do que os substantivos.

Scherre (1988), fazendo o cruzamento de *posição* e *classe gramatical*, conclui que os substantivos tendem a receber menos marcas de plural /s/ do que os adjetivos na segunda posição e a receber mais marcas de plural /s/ quando estão na terceira posição do SN. Conclui também que qualquer classe gramatical que esteja na primeira posição tende a receber mais marcas de plural /s/. Ao encontro dessa conclusão, Carvalho H.,(1997) observa que as classes gramaticais que funcionam como determinantes (quantificador, possessivo, indefinido, artigo e demonstrativo) apresentam mais probabilidade de favorecer a presença de marcas do que os substantivos e adjetivos. Nessa etapa, a autora testou a variável *classe gramatical do elemento analisado* isoladamente.

Ressaltamos que os advérbios que ocupam uma posição no sintagma, são controlados, no entanto, não são considerados como dado de análise, visto que não flexionam.

O quadro a seguir (7) apresenta a variável *classe gramatical* juntamente com os fatores elencados em nossa pesquisa.

Quadro 7 – Classe gramatical do elemento sob análise

Classes	Ilustrações
Substantivos	“[...] vinha aí só: <i>os animais</i> era quem [...]” (ERS - 16)
Categoria substantivada/ Pron. 3ª pessoa	“[...] inturmava com <i>as outras</i> ’ começo [...]” (MIS - 71) “[...] ajudo a <i>todos eles</i> ’ porque que a [...]” (ADS - 97)
Adjetivo	“[...] fazendo <i>esses serviços pesados</i> ’ [...]” (RES - 169)
Artigo	“[...] pessoal’ <i>os pulíticos</i> faz só inganá [...]” (VCM - 87)
Quantificador <i>Todos</i>	“[...] eu usei de <i>todos os truque</i> que a mulé pode [...]” (MCSR – 112)
Pronome Possessivo	“[...] estudá’ <i>nossos pais</i> eram care:ntes [...]” (LF - 119)

Pronome indefinido	“[...] normal’ e fiz outros curso de: de de [...]” (MGT – 166)
Pronome demonstrativo	“[...] aca:ba pra aquelas pessoas que [...]” (MBS - 12)
Adj. 2 – mesmo/próprio	“[...] minha nê ” e os próprios meu patrão que [...]” (AMA – 63) “[...] amigos mesmos ” (ARSS – 20)

Fonte: Elaborado pela autora.

A terminologia dos fatores utilizada em nossa pesquisa assemelha-se aos rótulos categoriais empregados por Scherre (1988). No entanto, em nossa coleta não codificamos nenhum item do tipo *indeterminado* (adjetivo 2), item coletado por Scherre (1988).

Acerca da variável *classe gramatical e posição em relação ao núcleo*, Martins (2013) observa que os resultados estatísticos do cruzamento das variáveis *posição linear* e *classe gramatical* revelaram que elementos não nucleares antepostos são mais marcados do que os elementos não nucleares pospostos. Percebemos, assim, a relevância de replicarmos essa variável para constataremos essas marcações nos constituintes antepostos e pospostos ao núcleo.

O quadro 8, a seguir, em sua categorização, assemelha-se aos fatores observados por Carvalho H., (1997). Esclarecemos que, nas ilustrações, o núcleo do SN consta sublinhado e o termo analisado em **negrito**.

Quadro 8 – Classe e posição em relação ao núcleo

Fatores	Ilustrações
Primeira posição anteposta ao núcleo	“[...] fazê as coisas <u>simples e justas da vida.</u> ” (JACS - 17)
Segunda posição anteposta ao núcleo	“[...] demais essas outras <u>crianças</u> num eram [...]” (MCA - 81)
Terceira posição anteposta ao núcleo	“[...] nóisi: e os meu ôtro <u>irmão</u> tia que ajuda [...]” (IGA - 10)
Núcleo na primeira posição	“[...]sufocados pur bibidas <u>alcoólicas</u> ’ pur [...]” (MISS - 165)
Núcleo na segunda posição	“[...]tem nem essas vontades <u>de passeá fora</u> [...]” (JGN - 106)
Núcleo na terceira posição	“[...] era os seus costumes ’ o que [...]” (JBX - 28)
Núcleo na quarta posição	“[...] com os seus próprios <u>filhos</u> ’ assim [...]” (JGN - 106)
Primeira posição posposta ao núcleo	“[...] não temos políticos <u>comprometidos</u> com [...]” (JPC - 15)
Segunda posição posposta ao núcleo	“[...] das nossas vontades <u>mais urgentes</u> [...]” (JBX - 28)

Fonte: Elaborado pela autora (2016).

Não encontramos em nossos dados nenhum constituinte ocupando a terceira posição posposta ao núcleo do SN.

3.6.1.2 Marcas precedentes

A variável *marcas precedentes* encontra-se replicada por vários pesquisadores que analisaram a variável concordância nominal de número no português brasileiro. No entanto, segundo Scherre (1988), há uma intensa diversificação na forma como aplicam essa variável. A autora exemplifica que Poplack observa o efeito das *marcas precedentes* em função da posição que o elemento ocupa no SN. Obtendo por conclusão que esta variável apresenta um efeito muito forte sobre o cancelamento de -s plural. Para a autora, “[...] *uma marca imediatamente precedente leva a mais marcas (quando s, Os ou ss precede o dado em questão, o cancelamento é desfavorecido) [...] e zeros levam a zeros, um efeito que é claramente estrutural*[...] (POPLACK, 1981, p. 61 apud SCHERRE, 1988, p. 168).

Scherre (1988) retoma essa variável fazendo analogia ao trabalho de Poplack e apresenta o estudo da variável *marcas precedentes* em nove categorias. Fernandes (1996), Carvalho H., (1997), Andrade (2003), Martins (2013), e outros, seguem a categorização utilizada por Scherre em sua tese, ao controlar a variável *marcas precedentes*. Essa categorização também é aplicada em nossa pesquisa. O Quadro 9 apresenta a variável *marcas precedentes* categorizada conforme proposto por Scherre (1988).

Destacamos que o elemento sob análise consta em **negrito**.

Quadro 9 – Marcas precedentes em função da posição

Fatores	Posição de análise	Ilustrações
1) Ausência	Primeira	“[...] novos <i>costumes</i> [...]” (JBX-28)
2) Zero na 1ª posição	Segunda	“ [...] na bicicletas ’ TAMÉM [...]” (JEBB -11)
3) Sintagma preposicionado		“[...] um bocado dos <i>filhos</i> [...]” (JWP - 85)
4) Numeral terminado em -s		“[...] tá com dois ano . [...]” (AMA – 63)
5) Numeral não terminado em -s		“[...] há cinco anos . ” [...] (JACS – 17)
6) Presença de marca a partir da 1ª posição	Segunda	“[...] todos esses <i>movimento</i> [...] (LF-119)
7) presença de duas ou mais marcas precedendo o elemento sob análise	Terceira	“[...]’ algumas coisas boas que [...]” (JBX – 28)

8) Mistura de marcas com marca precedendo o elemento sob análise	Terceira Quarta ou quinta	“[...] mas <i>os dois mais velhos</i> [...]” (179 – CVSPA)
9) Misturas de marcas com zero precedendo o elemento sob análise		“[...] <i>do meus pais</i> ’ começaram [...]” (MISS – 165) “[...] <i>os meu ôtro irmão</i> [...]” (IGA-10)

Fonte: Elaborado pela autora (2016).

3.6.1.3 Processos morfofonológicos de formação de plural

A introdução do Princípio da Saliência Fônica para explicar aspectos do funcionamento sincrônico de uma língua foi feita por Lemle e Naro em estudos realizados sobre o português do Brasil entre 1974 e 1976. Este princípio consiste em estabelecer que as formas mais salientes são mais prováveis de reterem a marca de plural do que as menos salientes e foi denominado *processos* por Scherre (1978). Braga e Scherre (1976) analisam esse princípio focalizando o fenômeno concordância de número entre os elementos do SN. Outras pesquisas seguindo essa mesma linha foram produzidas: Braga (1977); Scherre (1978); Ponte (1979); e Carvalho Nina (1980). Naro (1981), verificando de que forma o princípio atua na concordância verbal, estabelece duas dimensões no eixo da saliência fônica: diferenciação material fônica ou processos e tonicidade. Guy (1981) também faz uso dessas duas dimensões.

Scherre retoma o mesmo princípio, no entanto o faz observando a possibilidade de uma terceira dimensão: o número de sílabas. Em nossa pesquisa consideramos as variáveis processos morfofonológicos de formação de plural e tonicidade dos itens lexicais singulares. Observe-as exemplificadas nos quadros 10 e 11 a seguir.

Quadro 10 – Processos Morfofonológicos de Formação de Plural

Fatores	Ilustrações
1) Plural duplo com alternância vocálica, podendo haver ou não inserção de -s	“[...] <i>esse:s nossos idosos</i> [...]” (MBS-12)
2) Plural com alteração silábica nos itens terminados em -l, podendo haver ou não inserção de -s.	“[...] <i>trabalhos sociais</i> [...]” (JBX-28)
3) Plural nos itens terminados em -ão, que ao realizarem plural, sofrem alteração silábica quando da inserção do -s (ou apenas alterações silábicas)	“[...] <i>umas religiões</i> [...]” (ADS-97)
4) Plural nos itens terminados em -r, com inserção de -e ou -es.	“[...] <i>os nossos professores</i> [...]” (MTMT-99)
5) Plural dos itens terminados em -s, com inserção de -es.	“[...] ou <i>todos os meses</i> à missa [...]” (JWP -85)
6) Plural regular dos itens terminados em vogal	“[...] <i>dois irmãos</i> [...]” (NBS-38)

ou vogal + nasal com inserção de –s sem alteração morfonêmica.	
--	--

Fonte: Elaborado pela autora (2016).

3.6.1.4 Tonicidade dos itens lexicais singulares

O quadro 11 apresenta a variável tonicidade dos itens lexicais singulares.

Quadro 11 – Tonicidade dos itens lexicais singulares

Fatores	Ilustrações
1) Oxítonos e monossílabos de uso tônico	“[...] meus irmão [...]” (MGT-166)
2) Paroxítonos ou monossílabos de uso átono	“[...] as criança [...]” (ALA-98)
3) Proparoxítonos	“[...] os médico [...]” (MJNS-42)

Fonte: Elaborado pelo autor (2016).

3.7 As variáveis independentes extralinguísticas

Nesta seção, apresentamos as variáveis extralinguísticas controladas em nosso trabalho. Em seguida, apresentamos e discutimos acerca das variáveis – selecionadas pelo programa estatístico – que exercem influência sobre a variável concordância nominal de número na região do Cariri cearense, especificamente nas cidades do Crato, Juazeiro e Barbalha.

3.7.1 Sexo

No tocante à variável sexo, a primeira referência à relação entre variação linguística e ao fator sexo se encontra em Fischer, 1958 *apud* PAIVA, 2003, em um estudo intitulado *Influências sociais na escolha de variantes linguísticas*. Analisando a variação da pronúncia do sufixo “-ing”, o autor constata que a pronúncia velar (forma de prestígio) era mais frequente entre as mulheres. Outros estudos de cunho variacionista também apresentaram resultados favoráveis aos resultados de Fischer. Em níveis fonológico, morfossintático e semântico constata-se ser o sexo feminino quem mais utiliza a forma padrão. (PAIVA, 2003, p. 33-34). Labov, em suas pesquisas, traz dados que tratam da utilização da forma padrão pelas mulheres: “Para variáveis sociolinguísticas estáveis, as mulheres apresentam uma menor taxa de variantes estigmatizadas e maior taxa de variantes de prestígio do que os homens” (LABOV, 2001, p. 266).

No que diz respeito ao fenômeno concordância nominal de número, os resultados de trabalhos realizados com o português do Brasil, à luz da teoria da variação e mudança, também chegam à mesma conclusão, a exemplo os estudos de Scherre (1988), Carvalho H. (1997), Santos (2010), Guimarães (2014) e Meira (2015).

Sobre o comportamento das mulheres diante do uso das variáveis, Labov (2001) destaca a importância do fator social gênero aos fenômenos variáveis, apresentando o Paradoxo do Gênero, considerado inicialmente - *the linguistic conformity of women*: “Para variáveis sociolinguísticas estáveis, as mulheres apresentam uma menor taxa de variantes estigmatizadas e maior taxa de variantes de prestígio do que os homens” (LABOV, 2001, p. 266). Assim, as mulheres assumem mais as variáveis consideradas de prestígio. No entanto, as variáveis inovadoras, quando não estigmatizadas pela comunidade de fala, são deixadas de lado pelas mulheres, assumindo elas mais que os homens o uso dessa variável. Como a falante faz essas escolhas?

Contudo, Labov tenta solucionar esse *paradoxo*, argumentando que o problema continua, porque as mesmas pessoas às vezes são "conservadoras" e às vezes são "progressistas". Isso aparece de forma mais clara, segundo o autor, se substituímos estes termos por "conformismo" e "não-conformismo" - o *Paradoxo do Gênero* poderia ser reformulado, segundo Labov por *Paradoxo da Conformidade*, melhor dito, com o inverso de conformidade, o desvio (*deviation*), “mulheres desviam menos do que os homens das normas linguísticas quando os desvios são excessivamente prescritos, porém desviam mais do que os homens quando os desvios não são prescritos” (LABOV, 2001, p. 366 – 367).

Quanto à variável extralinguística *sexo*, o resultado de vários trabalhos realizados com o português do Brasil, à luz da teoria da variação e mudança, atribui às mulheres índices favoráveis à presença de marcas explícitas de plural /s/. Conclusão a exemplo dos estudos de Guimarães (2014), Carvalho, H; (1997), Meira (2015), Santos (2010) e Scherre (1988). Ressalte-se que na pesquisa de Carvalho R.,(1997), os homens mostraram-se mais favoráveis às presenças de marcas explícitas de plural.

3.7.2 Anos de escolaridade

No que se refere à variável escolaridade, segundo Votre (2003), os conteúdos programáticos referentes à concordância estão presentes em todos os níveis de ensino, numa ordem crescente de exigência proporcional aos anos de escolarização. Por conseguinte, quanto

maior o nível de escolarização do falante, maior a variável concordância. Enquanto variável social clássica e controlada em diversos trabalhos, o nível de escolarização se destaca como fator de influência significativa no uso da concordância nominal de número no português do Brasil.

Sendo uma variável social, o fator anos de escolaridade carrega em sua atuação sobre o uso da língua outras implicações. De acordo com Scherre (2005, p. 133):

A variação da concordância é parte inerente de nosso sistema linguístico (e é parte inerente também do francês, do inglês, do espanhol de Porto Rico, do crioulo de Cabo Verde e de outras tantas línguas), *mas a quantidade de variação, no Brasil, é marca social*. Inquestionavelmente, as pessoas mais escolarizadas, mais sensíveis às marcas de prestígio e que exercem profissões de trato público tendem a fazer mais concordância e, se não as fazem, são criticadas por nós, que também deixamos de fazer concordâncias verbais e nominais, de forma regular, quer queiramos quer não queiramos, quer reconhecamos quer não reconhecamos. (grifo da autora)

Acerca da importante influência da escolaridade nos processos variáveis, Naro e Scherre (2014) corroboram com a relevância do fator extralinguístico – ano de escolaridade. No entanto, levantam um ponto bastante intrigante: além do desenvolvimento educacional percebido da década de 1970 até então, os autores destacam o contato dos menos escolarizados com a mídia e isso vindo a influenciar no uso da norma padrão.

3.7.3 Faixa etária

É notório que existem diferenças linguísticas relacionadas à faixa etária do falante, principalmente em idades extremas ou entre a linguagem dos adolescentes e idosos. O controle da variável *faixa etária* é muito importante para a realização de estudos em tempo aparente.

Nesse caso, observando o comportamento linguístico do falante, relacionado a um fenômeno variável, as diversas faixas etárias, os estudos sócio-linguísticos podem ou não detectar processos de mudança linguística.

No tocante à variável social *faixa etária*, há alguns resultados obtidos nas pesquisas: Scherre (1988), a partir da análise em tempo aparente, sinaliza na fala dos cariocas um fenômeno de variação estável na fala do Rio de Janeiro; Carvalho H., (1997) indica que a faixa etária mediana 26 – 49 anos marca menos, constatando variação estável também na fala de João Pessoa; em Lopes (2001), pesquisando falantes da Bahia, encontra-se na faixa etária mais velha a única favorecedora do uso de presença da marca explícita de plural, pois é a que mais tende a marcar mais os elementos do sintagma nominal com o morfema de plural; por sua vez, em Guimarães (2014), tem-se que os falantes mais jovens tendem a apresentar mais as marcas

de concordância do que os falantes mais velhos, no português popular de Vitória da Conquista (PPVC).

Como podemos perceber, os resultados ora se assemelham ora divergem. Ou seja, não temos resultados que sigam somente em uma direção.

4 ANÁLISE E DISCUSSÃO DOS RESULTADOS

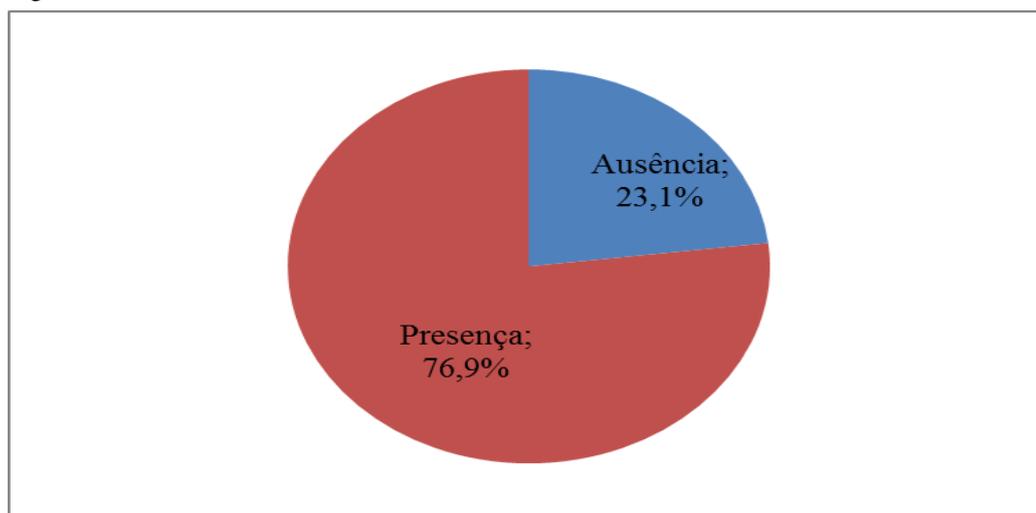
Neste capítulo, apresentamos e discutimos sobre os resultados obtidos acerca da variável dependente *presença ou ausência da marca de plural no SN*. Na primeira seção, analisamos os resultados dos grupos de fatores selecionados, após as rodadas do programa GoldVarb X, no que tange às variáveis linguísticas. Por sua vez, na segunda seção, discutimos os dados estatísticos das variáveis sociais, a fim de observarmos a influência desses fatores no uso da marcação de plural no SN, fenômeno em estudo. Na terceira seção, apresentaremos resultados obtidos na terceira rodada sem as variáveis *classe gramatical e posição linear*.

Nesta pesquisa, numa abordagem atomística, categorizamos 3304 constituintes dos sintagmas nominais coletados no falar dos 24 informantes da região CRAJUBAR. Os dados foram coletados na região do Cariri cearense. As entrevistas compõem o *corpus* do Profala e se encontram transcritas e armazenadas eletronicamente. Os 3304 constituintes foram submetidos ao programa estatístico Goldvarb X.

4.1 A Variação na concordância nominal de número na região CRAJUBAR

Iniciamos esta etapa de nossa análise examinando os resultados apresentados na primeira rodada do programa estatístico com os 3304 dados. Desses, 2541 representam a variante “*presença de marcas explícitas de plural /s/*”, equivalente a 76,9%, enquanto a variante “*ausência de marcas explícitas de plural /s/*” foi representada por 763 dados, equivalente a 23,1% dos dados. Diante disso, podemos observar o quanto os falantes da região do CRAJUBAR são favoráveis ao uso de *marcas de plural* (**gráfico 1**), em termos totais. Ressaltamos que esses resultados são gerais e contemplam, portanto, percentuais totais, sem levar em consideração a especificidade de cada constituinte do SN, bem como a atuação dos condicionamentos linguísticos e sociais. Durante este capítulo apresentamos os resultados por grupo de fatores.

Gráfico 1 – Distribuição da variável dependente presença ou ausência de marca de plural no SN da Região CRAJUBAR



Fonte: Elaborado pela autora (2016).

Outras pesquisas apresentaram para a variante *presença de marcação de plural* os seguintes resultados: Scherre (1988), 72% adultos e 65% crianças; Carvalho, H. (1997), 66%; Carvalho, R. (1997), 67%; Santos (2010), 52%; Martins (2013), 58%; Meira (2015), 89,9%. Todos apresentam valores percentuais sempre maiores que *ausência de marcas explícitas de plural /s/*". Contudo ressaltamos que resultados totais não dão conta das especificidades da concordância no SN.

Identificamos, nessa primeira rodada, quatro nocautes, três deles na variável *marcas precedentes em função da posição*. O primeiro nocaute apresentou marcação de plural em todos os constituintes e se referem a 17 casos de constituintes precedidos de expressões partitivas: “*a maioria dos/a maioria das*”, “*um dos/uma das*”. Os dezessete casos foram excluídos, por se tratarem de SN complexos partitivos com estruturas muito diferentes dos sintagmas nominais em estudo.

O segundo se refere ao fator *zero na primeira posição* (12 ocorrências). Os constituintes que precedem a primeira posição – zero na primeira posição - foram marcados, ou seja, apresentam marcas explícitas de presença de plural. Todas as ocorrências identificadas, a seguir, foram excluídas da rodada:

(1) “[...] continuidade **ao nossos estudo**’ hoje [...]” (165 – MISS)

(2) “[...] casa **do meus pais**’ começaram [...]” (165 – MISS)

(3) “[...] ôtos traz **na bicicletas**’ TAMÉM [...]” (11 – JEBB)

- (4) “[...] essas **pela essas casa** que tinha flores [...]” (179 – CVSPA)
- (5) “[...] porque **o:: meus irmão** também num [...]” (10 – IGA)
- (6) “[...] de conquistá **o nossos sonhos** [...]” (11 – JEBB)
- (7) “[...] temos **bom cantores**’ eu acho que [...]” (11 – JEBB)
- (8) “[...] temos **muito talentos**’ FALTA SÓ[...]
- (9) “[...] era **aquela duas músicas** e ele quando [...]” (99 – MTMT)
- (10) “[...] são **excelente crianças**’ eu se Deus [...]” (165 – MISS)
- (11) “[...] mensagem’ a **todo aqueles** que [...]” (165 – MISS)
- (12) “[...] essa **essa aulas** como é que chama.” (99 – MTMT)

Os exemplos (1), (2), (3) e (4), apresentam uma configuração estrutural **(Preposição + artigo) + Pronome + Núcleo**. Nesses casos, parece haver já cristalizada uma estrutura lexicalizada, estando o sintagma comportando-se como uma lexia complexa. Quanto aos dados “do meus pais”, “ao nossos estudos”, Scherre (1988, p. 48) encontra esse tipo de ocorrência em seu estudo e sugere que o falante pode estar analisando esta contração **(prep. + art)** como uma preposição, ou seja, “como uma categoria que não se flexiona e, portanto, não apresenta marcas de plural”.

Nos casos de “[...] *bom cantores* [...]”, exemplo (7), e “[...] *excelente crianças* [...]”, exemplo (10), não dispomos de outras ocorrências **(adjetivo + substantivo)** para analisarmos com mais detalhes.

O terceiro nocaute, presente na variável *marcas precedentes em função da posição*, refere-se a um caso de sintagma preposicionado “[...] *um bocado dos filhos* [...]”. Esse único item também foi excluído e também apresenta marca explícita de plural em *filhos*.

O quarto nocaute foi apresentado na variável *processos morfofonológicos de formação de plural*. Trata-se de 17 constituintes *de plural duplo com alternância vocálica, podendo haver ou não inserção de -s*. Todos eles apresentando “*presença de marcas explícitas de plural /s/*”. Abaixo, alguns exemplos:

- “[...] meus **esforços** [...]” (APSN – 19)
- “[...] **nôvos métodos** [...]” (APSN – 19)
- “[...] **novos sentimentos** [...]” (JBX – 28)
- “[...] **novos costumes** [...]” (JBX – 28)
- “[...] **novos tipos** [...]” (JBX – 28)

“[...] *esse:s nossos idosos* [...]” (MBS – 12)

“[...] *os estudiosos* [...]” (MBS – 12)

“[...] *puns porcos* [...]” (MJNS – 42)

“[...] *cuns fogos* [...]” (ADS – 97)

“[...] *os olhos* [...]” (ALA – 98)

Os 17 constituintes com *plural duplo com alternância vocálica, podendo haver ou não inserção de –s* apresentaram 100% de marcas explícitas de plural, resultado que se assemelha aos obtidos em Scherre (1988), 93%; Carvalho, H.(1997), 81%; Carvalho, R.(1997), 90%; Lopes (2001), 88%; Schneider (2012), 85%; Meira (2015), 79%. Percebemos, então, que esses constituintes com formação de plural irregular favorecem a aplicação de marcas explícitas de plural, por serem formas mais salientes. Esses 17 itens, presentes no quarto caso de nocaute, foram excluídos da rodada. Outros itens com formação de plural irregular também se mostraram favoráveis à marcação de plural e serão demonstrados na seção seguinte.

Assim, os nocautes foram solucionados com a exclusão dos 17 constituintes com *expressões partitivas*, 12 constituintes apresentando *zero na primeira posição do SN*, 1 constituinte com *sintagma preposicionado* e 17 constituintes com ocorrências de *plural duplo*.

4.1.1 Variáveis Linguísticas

Solucionados os nocautes, passamos para a análise das variáveis linguísticas selecionadas pelo programa estatístico numa segunda rodada, considerando-as isoladamente. Inicialmente, demonstraremos a frequência de marcas explícitas de plural nos constituintes do SN, segundo a variável *posição linear* e *classe gramatical* de acordo com os resultados obtidos na primeira rodada.

4.1.1.1 Posição linear

A seguir, utilizando os resultados obtidos na primeira rodada, apresentamos a frequência do fator *posição linear do elemento no sintagma nominal*, tomando como base a primeira rodada que realizamos com um total de 3304 constituintes, os quais se encontram demonstrados na tabela 1.

Tabela 1 – Frequência da “presença de marcas explícitas de plural nos constituintes do SN” segundo a variável *posição linear* do elemento no SN

Posição	Aplic./Total	Frequência
Primeira posição	1341/1365	98,2 %
Segunda posição	1022/1649	62 %
Terceira posição	163/266	61,3%
Quarta posição	11/18	61%
Quinta posição	4/6	66,7%
Total	2541/3304	76,9%

Fonte: Elaborado pela autora (2016)

Nos exemplos a seguir, o item sob análise encontra-se em negrito.

a) 1ª posição do elemento no SN.

(1) “[...] *os nossos governantes* [...]” (ARSS-20)

b) 2ª posição do elemento no SN.

(2) “[...] *os seus costumes* [...]” (JBX – 28)

c) 3ª posição do elemento no SN

(3) “[...] *umas novas ISCOLA*’ [...]” (IGA – 10)

d) 4ª posição do elemento no SN.

(4) “[...] *os meu ôtro irmão* [...]” (IGA – 10)

e) 5ª posição do elemento no SN

(5) “[...] *as principais coisas mais importante* [...]” (MISS – 165)

No tocante à primeira posição do SN – exemplo 1 “[...] *os nossos governantes* [...]” – podemos ratificar que os constituintes nessa posição apresentam consideravelmente marcas explícitas de plural. Os resultados representados na tabela 1 apresentam uma percentagem de 98,2% de marcas explícitas de plural na primeira posição. Esse resultado é compartilhado com os inúmeros trabalhos já concluídos nessa linha, visto que, segundo Scherre (1988), a primeira posição do SN é a mais marcada, num índice probabilístico nunca inferior a 0,70; e as demais posições evidenciam um índice baixo de marcas. Segunda posição – exemplo 2 “[...] *os seus costumes* [...]” –, terceira posição – exemplo 3 “[...] *umas novas ISCOLA*’ [...]” –, quarta posição – exemplo 4 “[...] *os meu ôtro irmão* [...]” – e quinta posição – exemplo 5 “[...] *as principais coisas mais importante* [...]” – não divergem em seus índices, apresentando uma média de 62,8 %.

Elencamos algumas análises de pesquisas anteriores, considerando a variável posição linear isoladamente. Os índices obtidos constataam o quanto são uniformes esses resultados referentes à primeira posição: Fernandes (1996): 97%; Carvalho, H. (1997): 96%;

Carvalho, R. (1997) 99%; Santos (2010): 88%; Martins (2013): 96%; Guimarães (2014): 99%; Meira (2015): 99,8%.

O comprovado favorecimento de marcas explícitas de plural aos constituintes que ocupam a primeira posição no SN não se deve somente à posição linear, outros fatores encontram-se imbricados nesses resultados, principalmente, a relação da posição dos constituintes com o núcleo. Estudos acerca da influência de outras variáveis relativos ao fato de a primeira posição no SN apresentar índices maiores de marcação se encontram, inicialmente, na tese de Scherre (1988). Na pesquisa, a autora faz uma reanálise da concordância nominal de número e nesse tópico especificamente propõe um fator cruzando *posição, classe gramatical e núcleo do SN*.

A partir de então, numa segunda rodada, resolvidos os nocautes – conforme apresentamos na seção 4.1 – o programa estatístico selecionou como significativos os seguintes grupos de fatores: *posição linear; posição e classe gramatical em relação ao núcleo; marcas precedentes em função da posição; processos morfofonológicos de formação de plural; sexo e faixa etária*.

É válido destacar a não seleção da variável *escolarização*, visto que, tratando-se especificamente do uso de concordância de número, é sabido e comprovado em vários trabalhos o peso significativo dessa variável no falar dos informantes. Outra variável não selecionada pelo programa estatístico foi *classe gramatical*.

Na tabela 2, apresentamos apenas as percentagens a título de esclarecimento acerca das classes gramaticais que mais favorecem a presença de marca explícita de plural /s/. Esclarecemos que estes resultados foram obtidos na primeira rodada, portanto constam todos os dados iniciais – 3304 constituintes. Vejamos, então, como ficou essa distribuição.

Tabela 2 – Frequência da “presença marcas explícitas de plural nos constituintes do SN” segundo a variável *classe gramatical*

Classes	Aplic./Total	Frequência
Artigo	873/888	98,3%
Quantificador <i>todos</i>	59/60	98,3%
Pronome demonstrativo	137/140	98%
Pronome possessivo	152/157	96,8%
Pronome indefinido	176/183	96,2%
Categoria substantivada/Pron. 3ª pessoa	40/60	66,7%
Adjetivo	96/156	61,5%
Substantivos	1006/1655	60,8%
Adjetivo 2 – mesmo/próprio	2/5	40%
Total	2541/3304	76,9%

Fonte: Elaborado pela autora (2016)

De acordo com os resultados demonstrados na tabela 2, podemos constatar que as classes gramaticais exercendo função de determinantes (artigo, quantificador, pronomes demonstrativos, possessivos e indefinidos) retêm índices significativos de marcação de plural, enquanto os substantivos e adjetivos retêm menor percentual de marcação de plural. Estamos cientes de que a análise feita com as variáveis *classe gramatical e posição linear* – isoladamente – não nos fornecem resultados a contento por entendermos, à luz de Scherre (1988) e outros estudos sobre a concordância nominal no Português do Brasil, que há um funcionamento morfossintático em que classe gramatical e posição da classe gramatical se entrecruzam ao núcleo do SN para o uso da marca ou da ausência de plural. Acerca dessas variáveis, Scherre (1994, p. 4) conclui: “[...] a melhor forma de se entender a variação na concordância no português do Brasil é através do cruzamento entre elas”.

Dessa forma, também controlamos em um só grupo de fatores classe e posição em relação ao núcleo. Esse foi o primeiro grupo de fatores selecionado pelo programa Goldvarb X em várias rodadas testes e na rodada seguinte, reforçando a importância que essa variável assume sob o fenômeno. Demonstramos, na tabela 3, os resultados estatísticos da variável *posição e classe gramatical em relação ao núcleo*.

4.1.1.2 *Posição e classe gramatical em relação ao núcleo*

Seguimos com os dados obtidos da variável *posição e classe gramatical em relação ao núcleo*, selecionada em primeiro lugar pelo programa estatístico como bastante significativa. Essa variável cruzada tem se mostrado relevante nos estudos sobre concordância nominal de número no português do Brasil.

O programa, ao selecioná-la em primeiro lugar, confirma a hipótese de que há, de fato, uma inter-relação entre a classe de palavra, sua posição no SN e núcleo do sintagma na aplicação de presença do morfema de plural no âmbito do SN.

Na tabela 3, apresentamos os resultados para o grupo de fatores *posição e classe gramatical em relação ao núcleo* obtidos nesta segunda rodada.

Tabela 3 – Frequência e peso relativo da “presença de marcas explícitas de plural” segundo a *posição e classe gramatical em relação ao núcleo*.

Fatores	Aplic./Total	Frequência da variante explícita	Peso relativo
Primeira posição anteposta ao núcleo	1276/1300	98,2%	0,70
Segunda posição anteposta ao núcleo	119/134	88,8%	0,71
Terceira posição anteposta ao núcleo	8/11	72,7%	0,85
Núcleo na primeira posição	57/60	95%	0,45
Núcleo na segunda posição	837/1431	58,5%	0,40
Núcleo na terceira posição	128/193	66,3%	0,55
Núcleo na quarta posição	6/11	54,5%	0,35
Primeira posição posposta ao núcleo	53/104	51,%	0,30
Segunda posição posposta ao núcleo	10/13	77 %	0,82
Total	2494/3257	76,6%	

Fonte: Elaborado pela autora (2016).

Acerca dos exemplos a seguir, o núcleo do SN consta sublinhado e o termo analisado em **negrito**.

a) 1ª posição anteposta ao núcleo

(1) “[...] *os professores* se dedicavam [...]” (AT – 60)

b) 2ª posição anteposta ao núcleo

(2) “[...] *os meus tio* moram [...]” (ADS – 97)

c) 3ª posição anteposta ao núcleo

(3) “[...] *e os próprios meu patrão* [...]” (AMA – 63)

d) Núcleo na 1ª posição

(4) “[...] *patrões* carrascos [...]” (AMA – 63)

e) Núcleo na 2ª posição

(5) “[...] *essas ropinha* [...]” (MCSR – 112)

f) Núcleo na 3ª posição

(6) “[...] *e os dois filho* homem [...]” (ADS – 97)

g) Núcleo na 4ª posição

(7) “[...] *os meu ôtro irmão* [...]” (IGA – 10)

h) 1ª posição posposta ao núcleo

(8) “[...] *esses transporte fácil*’_aí [...]” (ADS – 97)

i) 2ª posição posposta ao núcleo

(9) “[...] *coisas* boas’ *bunitas* [...]” (MISS – 125)

Podemos inferir, dos resultados expostos na tabela 3, que todos os constituintes antepostos ao núcleo se mostram favoráveis ao uso de marcação de plural com peso relativo de 0,70 para *primeira posição anteposta ao núcleo* – exemplo 1 “[...] os professores se dedicavam [...]” – de 0,71 para *segunda posição anteposta ao núcleo* – exemplo 2 “[...] os meus tio moram [...]” – e a *terceira posição anteposta ao núcleo* – exemplo 3 “[...] e os próprios meu patrão [...]” com peso 0,85.

Essa diferença apresentada na *terceira posição anteposta ao núcleo* em relação às primeira e segunda posições antepostas ao núcleo, somente no peso, acreditamos que se deve ao pequeno número de casos desse fator. Do total de 11 ocorrências, 8 casos com marcas explícitas de plural e 3 com ausência de marcas. As posições que estão antepostas ao núcleo favorecem a presença explícita do morfema de plural.

Os resultados a respeito da *variável posição e classe gramatical em relação ao núcleo* corroboram as pesquisas realizadas no Brasil, no sentido de que os estudos têm demonstrado que os constituintes antepostos ao núcleo favorecem a presença de marcas explícitas de plural no sintagma nominal.

Algumas pesquisas apresentaram, na primeira posição anteposta ao núcleo, os seguintes pesos relativos: Scherre (1988) – 0,87; Fernandes (1996) – 0,81; Carvalho, H. (1997) – 0,88; Lopes (2001) – 0,91; Andrade (2003) – 0,92; Martins (2013) – 0,78; Guimarães (2014) – 0,84. Esses resultados confirmam os índices encontrados em nosso estudo.

Quanto ao comportamento do núcleo, observamos que apresentam pesos relativos diferenciados, conforme sua posição ocupada no SN. O *núcleo na primeira posição* – exemplo 4 “[...] patrões carrascos [...]” – mostra uma disparidade entre a percentagem e o peso relativo. No entanto, apresenta-se mais marcado (0,45) que o *núcleo na segunda posição* – exemplo 5 “[...] essas ropinha [...]”, com (0,40). Este se mostra mais favorável a marcas explícitas de plural que o *núcleo na quarta posição* (0,35) – exemplo 7 “[...] os meu ôtro irmão [...]”. Observamos que o *núcleo na terceira posição* – exemplo 6 “[...] e os dois filho homem [...]” – apresenta peso relativo acima do ponto neutro (0,55) e se revela tão favorável quanto o *núcleo na primeira posição*.

Sobre os resultados dos constituintes pospostos ao núcleo, os que ocupam a *primeira posição posposta ao núcleo* – exemplo 8 “[...] esses transporte fácil’ aí [...]” – desfavorecem a presença de marca explícita de plural (0,30), mas na *segunda posição posposta ao núcleo* – exemplo 9 “[...] coisas boas’ bunitas [...]” –, o peso relativo (0,82) se

mostra semelhante aos constituintes antepostos ao núcleo. Mais uma vez, para justificar esse resultado, voltamos à quantidade dos constituintes que demonstram contagem bem inferior em relação aos constituintes de primeira posição antepostos ao núcleo.

Podemos concluir que os constituintes antepostos ao núcleo mostram-se mais favoráveis à marcação de plural no sintagma nominal do que os constituintes pospostos ao núcleo, os quais apresentam índices menores de marcação de plural, embora tenhamos identificado um peso relativo (0,82) no fator *segunda posição posposta ao núcleo*.

4.1.1.3 Marcas precedentes

A segunda variável selecionada pelo programa foi *faixa etária*, a qual será apresentada no tópico 4.1.2 destinado às *variáveis sociais*. Seguiremos apresentando os grupos de fatores linguísticos selecionados significativamente pelo programa Goldvarb X.

A variável *marcas precedentes* foi a terceira variável selecionada pelo programa estatístico Goldvab X em grau de significância. Acerca dessa variável, apresentamos nossos resultados na tabela 4 a seguir.

Tabela 4 – Frequência e peso relativo da “presença de marcas explícita de plural” segundo a variável *marcas precedentes*

Fatores	Posição de análise	Aplic./ Total	Frequência da variante explícita	Peso relativo
Presença de duas ou mais marcas explícitas de plural precedendo o elemento sob análise.	Terceira, Quarta	114/155	73,5 %	0,66
Presença de marcas explícitas na 1ª posição.	Segunda	813/1281	63,5%	0,52
Mistura de marcas com marca precedendo o elemento sob análise.	Terceira, Quarta	34/49	69,4%	0,50
Numeral não terminado em –s.	Segunda	122/200	61%	0,55
Numeral terminado em –s	Segunda	84/189	44,4%	0,41
Misturas de marcas com zero precedendo o elemento sob análise.	Terceira Quarta	1/33	3%	0.02
Total		1168/1907	61,2%	-

Fonte: Elaborado pela autora (2016).

Abaixo, o constituinte sob análise está sublinhado e os constituintes marcados estão em negrito.

a) presença de duas ou mais marcas explícitas de plural

(1) “[...] tirá ***todos os*** *documentos* [...]” (ALA – 98)

b) presença de marcas explícitas na 1ª posição

(2) “[...] assiste ***todos os*** *dias*’ mais [...]” (APSN – 19)

c) Mistura de marcas com marca precedendo o elemento sob análise

(3) “[...] ***as*** *pessoas* ***mais*** *carentes* [...]” (JBX-28)

d) Numeral não terminado em -s

(4) “[...] ***tem nove*** *anos* e [...]” (MGT – 166)

e) Numeral terminado em -s

(5) “[...] não’ tá com ***dois*** *ano* [...]” (AMA – 63)

f) Mistura de marcas com zero precedendo o elemento sob análise

(6) “[...] ***os*** *caso* *amorosos* [...]” (MCSR – 112)

Por meio da codificação da variável *marcas precedentes*, feita com base em Scherre (1988), temos os seguintes resultados apresentados na tabela 4.

Encontramos uma significativa influência no que diz respeito à presença de marcas explícitas de plural do fator *presença de duas ou mais marcas formais de plural precedendo o elemento sob análise* – exemplo 1 “[...] tirá ***todos os*** *documentos* [...]” – com peso 0,66, resultado que também encontramos em Scherre (1988), com peso relativo 0,60; Lopes (2001) - 0,63; e Carvalho H. (1987) - 0,67; constatando que a presença de duas ou mais marcas explícitas de plural precedendo o elemento sob análise contribui para o favorecimento de marcas. Nesse caso, aplica-se, também, aos dados do CRAJUBAR, o “achado” de Poplack (1981) *apud* Scherre (1988, p.168), afirmando que “[...] uma marca imediatamente precedente leva a mais marcas [...]”.

Acerca de *marcas que precedem o elemento sob análise*, outros fatores que seguem essa motivação é *presença de marcas explícitas na 1ª posição* – exemplo 2 “[...] assiste **todos os dias**’ mais [...]” – com peso relativo 0,52, mostrando-se semelhante ao resultado de Martins (2013) – 0,53; Scherre (1988) – 0,55; e Carvalho H. (1997) – 0,48 –, e *mistura de marcas com marca explícita precedendo o elemento sob análise* – exemplo 3 “[...] **as pessoas mais carentes** [...]” com peso relativo 0,50.

De acordo com os resultados, *presença de marcas explícitas na 1ª posição* e fator *presença de duas ou mais marcas formais de plural precedendo o elemento sob análise* demonstram que grande parte dos falantes do CRAJUBAR não opta por reter as marcas de plural quando os termos que o antecedem trazem essas marcas.

Quanto aos numerais, o valor semântico de plural imbricado nos numerais os torna constituintes motivadores à presença de marcas. É válido observar que os numerais não terminados em -s – exemplo 4 “[...] **tem nove anos** e [...]” – apresentam peso relativo (0,55) mais favorável a marcas do que o fator *presença de marcas formais na 1ª posição*. Sobre os numerais terminados em -s – exemplo 5 “[...] não’ tá com **dois ano** [...]” –, não foi significativo o bastante para favorecimento de marcas. Acreditamos que o falante do CRAJUBAR, no que se refere aos numerais, terminação com -s ou sem -s, não os considera como traço determinante para presença ou ausência de marcas explícitas de plural.

Já misturas de marcas com zero precedendo o elemento sob análise – exemplo 6 “[...] **os caso amorosos** [...]” – traz-nos peso relativo insignificante (0,02), apresentando marcas em apenas uma ocorrência no total de trinta e três encontradas e constatando que a *presença de marca explícita de plural antecedendo o constituinte sob análise* leva à *presença de marca*, enquanto a *ausência de marca antecedendo o constituinte* leva à *ausência de marca*.

4.1.1.4 Processos morfofonológicos de formação de plural.

Considerado pelo programa estatístico como o quarto grupo de fatores na ordem de significância, apresentamos, na tabela 5, os resultados da variável *processos morfofonológicos de formação de plural*.

Tabela 5 – Frequência e peso relativo da “presença de marcas explícitas de plural” segundo a variável de *processos morfofonológicos de formação de plural*

Fatores	Aplic./Total	Frequência da variante explícita	Peso relativo
1) Plural com alteração silábica nos itens terminados em -l, podendo haver ou não inserção de -s	51/62	82,3%	0,82
2) Plural nos itens terminados em -r, com inserção de -e ou -es	51/66	77,3%	0,70
3) Plural nos itens terminados em -ão, que, ao realizarem plural, sofrem alteração silábica quando da inserção do -s (ou apenas alterações silábicas)	35/56	62,5%	0,56
4) Plural regular dos itens terminados em vogal ou vogal + nasal com inserção de -s sem alteração morfofonêmica	2341/3024	77,4%	0,50
5) Plural dos itens terminados em -s, com inserção de -es	16/49	32,7%	0,31
Total	2494/3257	76,6 %	-

Fonte: Elaborado pela autora (2016)

Nos exemplos a seguir, os constituintes sob análise estão em negrito.

a) Plural com alteração silábica nos itens terminados em -l, podendo haver ou não inserção de -s

(1) “[...] *as famílias tradicionais* [...]” (MTMT – 99)

b) Plural nos itens terminados em -r, com inserção de -e ou -es

(2) “[...] têm *professores* [...]” (APSN – 19)

c) Plural nos itens terminados em -ão, que, ao realizarem plural, sofrem alteração silábica quando da inserção do -s

(3) “[...]de algumas *religiões* não [...]” (JBX – 28)

d) Plural regular dos itens terminados em vogal ou vogal + nasal com inserção de -s sem alteração morfofonêmica

(4) “[...] *aquelas mocinhas* [...]” (ADS – 97)

e) Plural dos itens terminados em -s, com inserção de -es.

(5) “[...] *logo seis meses* [...]” (AMA – 63)

Denominada por Scherre de *processos*, o princípio da *saliência fônica* consiste em estabelecer que as formas mais salientes estão mais favoráveis à marcação de plural do que as menos salientes. A autora, em sua tese, considera o princípio sob três dimensões: processos morfofonológicos de formação de plural, tonicidade dos itens lexicais singulares e número de sílabas dos itens lexicais singulares. Em nossa pesquisa, controlamos duas: *processos*

morf fonológicos de formação de plural e tonicidade da sílaba dos itens lexicais singulares como grupos de fatores separados.

Processos morf fonológicos de formação de plural foi considerada a quarta, em grau de significância, pelo programa estatístico, enquanto a variável *tonicidade dos itens lexicais* não foi selecionada.

Sobre a variável *processos morf fonológicos de formação de plural* em nossa pesquisa, apresentados na tabela 5, foram encontrados os seguintes resultados.

No fator *plural duplo com alternância vocálica, podendo haver ou não inserção de -s*, encontramos 17 ocorrências na rodada inicial e todos apresentaram marcas de plural em seus constituintes. Por conseguinte, o fator apresentou nocaute e resolvemos solucioná-lo com a exclusão das 17 ocorrências da rodada, usando o recurso nil do Goldvarb. Dito de outro modo, palavras que realizam o plural como “*novo*” – “*novos*”, cuja alternância vocálica é realizada na passagem singular-plural, apresentaram 100% de realização de marca de plural. Ressaltamos que esse resultado vai ao encontro de vários trabalhos (citamos algumas pesquisas na seção 4.1), os quais encontram nos constituintes mais salientes índices relevantes de marcação de plural.

Seguindo os índices do fator *plural duplo com alternância vocálica, podendo haver ou não inserção de -s*, os constituintes presentes no fator *plural com alteração silábica nos itens terminados em -l, podendo haver ou não inserção de -s* – exemplo 2 “[...] têm **professores** [...]” – mostram-se favorecedores da concordância nominal de plural. Destacaram-se com peso relativo bastante significativo (0,84); mesmo índice encontrado por Carvalho H. (1997) – (0,84) – e próximos ao encontrado por Martins (2013) – (0,90) – e Meira (2015) – (0,87) –, destoando do resultado de Scherre (1988) – (0,56).

Esses constituintes terminados em -l são seguidos pelos constituintes *plural nos itens terminados em -r, com inserção de -e ou -es* – exemplo 3 “[...] de algumas **religiões** não [...]” – com peso relativo 0,70, próximo ao encontrado por Carvalho H. (1997) – (0,79) –, mas destoando de Scherre (1988) – (0,48) – e de Guimarães (2014) – (0,53).

Quanto ao fator *plural regular dos itens terminados em vogal ou vogal + nasal com inserção de -s sem alteração morfofonêmica* – exemplo 4 “[...] aquelas **mocinhas** [...]” –, não se mostra como menos favorável às marcas de plural (0,50). No entanto, o peso relativo apresentado aproxima-se do último fator em nossa escala *plural dos itens terminados em -s, com inserção de -es* – exemplo 5 – com peso relativo (0,31).

Em nossos resultados, podemos constatar que, na variável *processos morf fonológicos de formação de plural*, a maioria dos constituintes com formação de plural

irregular apresenta-se com os índices maiores de marcação de plural, confirmando o princípio da *saliência fônica* de que há uma relação entre os constituintes mais salientes e a presença de marcas de plural nos sintagmas nominais dos falantes do CRAJUBAR.

4.1.2 Rodada sem as variáveis posição linear e classe gramatical

As variáveis selecionadas na 2ª rodada também foram nesta, o diferencial foi observado na ordem de significância de algumas variáveis. Ressaltamos que as variáveis *posição linear* e *classe gramatical* não estão sendo controladas nesta 3ª rodada. De acordo com a seleção do programa, por ordem de significância, temos as seguintes variáveis: *classe e posição em relação ao núcleo, faixa etária, marcas precedentes, processos morfofonológicos de formação de plural e sexo*.

Em todas as rodadas efetuadas na pesquisa, a variável *posição e classe gramatical em relação ao núcleo* foi selecionada em primeiro lugar pelo programa, visto que é a que melhor reproduz o fenômeno em estudo.

Vejam, então, na tabela 6, os resultados do grupo de fatores *classe e posição em relação ao núcleo*, que foi a primeira variável selecionada.

Tabela 6 - Frequência e peso relativo da “presença de marcas explícitas de plural” segundo a *posição e classe gramatical em relação ao núcleo*

Fatores	Aplic./Total	Frequência da variante explícita	Peso relativo
Primeira posição anteposta ao núcleo	1276/1300	98,2%	0,89
Segunda posição anteposta ao núcleo	119/134	88,8%	0,53
Terceira posição anteposta ao núcleo	8/11	72,7%	0,59
Núcleo na primeira posição	57/60	95%	0,69
Núcleo na segunda posição	837/1431	58,5%	0,20
Núcleo na terceira posição	128/193	66,3%	0,23
Núcleo na quarta posição	6/11	54,5%	0,33
Primeira posição posposta ao núcleo	53/104	51,%	0,20
Segunda posição posposta ao núcleo	10/13	77 %	0.50
Total	2494/3257	76,6%	-

Fonte: Elaborado pela autora (2016).

Acerca dos exemplos a seguir, o núcleo do SN consta sublinhado e o termo analisado em **negrito**.

a) 1ª posição anteposta ao núcleo

(1) – “[...] vem os adulto [...]” - (MBS – 12)

b) 2ª posição anteposta ao núcleo

(2) “[...] os meus filhos [...]” (MISS – 165)

c) 3ª posição anteposta ao núcleo

(3) “[...] os meu ôtro irmão [...]” (IGA - 10)

d) Núcleo na 1ª posição

(4) “[...] é: pessoas ruins [...]” (MJS - 72)

e) Núcleo na 2ª posição

(5) “[...] que essas criança [...]” (ALA - 98)

f) Núcleo na 3ª posição

(6) “[...] os nossos professores [...]” (MTMT - 99)

g) Núcleo na 4ª posição

(7) “[...] uns dos meus filho [...]” (ADS – 97)

h) 1ª posição posposta ao núcleo

(8) “[...] palestras educativa [...]” (MGT - 166)

i) 2ª posição posposta ao núcleo

(9) “[...] as pessoas mais carentes [...]” (JBX - 28)

Os pesos relativos apresentados na variável *posição e classe gramatical em relação ao núcleo*, nesta 3ª rodada, ratificam os resultados já apresentados, os quais demonstram o favorecimento as marcas explícitas de plural na *primeira posição anteposta ao núcleo* – exemplo 1 “[...] vem os adulto [...]” com peso relativo 0,88.

Os índices comprovam, também, que o *núcleo na primeira posição* – exemplo 4 “[...] é: pessoas ruins [...]” com peso relativo 0,69 – retém índices maiores de favorecimento às marcas explícitas de plural do que os constituintes da *segunda* – exemplo 2 “[...] os meus filhos [...]” – e da *terceira posição anteposta ao núcleo* – exemplo 3 “[...] os meu ôtro irmão [...]” com peso relativo 0,53 e 0,59, respectivamente. Dessa forma, observamos que os constituintes antepostos ao núcleo favorecem a marcação de plural. No entanto, os itens que

ocupam o núcleo na primeira posição também se mostram bastante favoráveis, visto que os índices obtidos ultrapassaram os valores da segunda e da terceira posições antes do núcleo, assim como as demais posições posposta ao núcleo.

As restantes variáveis selecionadas não sofreram alterações nos resultados. Portanto, não há necessidade de apresentá-los, visto que já foram comentados nas seções 4.1.1 e 4.1.2.

4.1.3 Variáveis Sociais

Quanto às variáveis sociais, devemos ressaltar que, das três analisadas pelo programa estatístico, foram selecionadas duas: *sexo e faixa etária*. Apresentamos, na tabela 6, os resultados obtidos acerca da variável *faixa etária*.

4.1.3.1 Faixa etária

Na tabela 7, demonstramos os resultados acerca da variável *faixa etária* que foi a segunda variável selecionada pelo programa estatístico.

Tabela 7 – Frequência e peso relativo da “presença de marcas explícitas de plural” segundo a variável *faixa etária*

Fatores	Aplic./Total	Frequência da variante explícita	Peso relativo
15 a 25 anos	755/848	89%	0.72
26 a 49 anos	920/1314	70%	0.40
50 em diante	819/1095	75%	0.45
Total	2494/3257	76,6 %	

Fonte: Elaborado pela autora (2016).

Em relação à variável *faixa etária*, nossa hipótese era a de que os falantes do CRAJUBAR na faixa etária de 26 a 49 anos e os falantes com mais de 50 anos tenderiam a usar mais marcas de plural no SN, em razão de outros fatores como o de ser essa a faixa etária que supostamente ocupa o mercado de trabalho. No entanto, pelos resultados obtidos representados na tabela 7, nossa hipótese foi refutada.

Podemos perceber que os jovens de 15 a 25 anos apresentam-se como os mais favoráveis ao uso da presença *de marcas explícitas de plural no SN* na região do CRAJUBAR

com peso relativo 0,72; os adultos, de 26 a 49 anos, apresentando-se desfavoráveis ao uso com peso relativo 0,40, assim como os informantes de 50 anos em diante, apresentando peso relativo 0,50 no ponto neutro.

Acerca dos dados, ratificamos que os mais jovens da amostra favorecem a presença de marcas e os falantes adultos e os falantes com mais de 50 mantêm um padrão semelhante de uso do fenômeno *concordância nominal de número entre os constituintes dos sintagmas nominais*.

4.1.3.2 Sexo

Apresentamos, na tabela 7, os resultados obtidos sobre a variável social *sexo*.

Tabela 7 – Frequência e peso relativo da “presença de marcas explícitas de plural” segundo a variável *sexo*

Fatores	Aplic./Total	Frequência da variante explícita	Peso relativo
Feminino	1208/1551	80%%	0,53
Masculino	1286/1706	75%	0,47
Total			

Fonte: Elaborado pela autora (2016).

No tocante à variável *sexo*, selecionada em sexto lugar pelo programa estatístico, nossa hipótese era a de que as mulheres demonstrassem ser mais favoráveis ao uso da *marcação de plural* no fenômeno da concordância nominal de número, no CRAJUBAR, do que os homens.

O que podemos inferir é que nossa hipótese foi confirmada, visto que as mulheres apresentam peso relativo 0,53, ainda que bem próximo do ponto neutro e os homens com peso relativo 0,47, mostrando-se abaixo do ponto de significância, porém com uma diferença numérica mínima com relação às mulheres.

Encontramos essa tendência favorável à retenção de *marcas de plural* pelas mulheres em Scherre (1988); Carvalho H. (1987), Santos (2010), Guimarães (2014) e Meira (2015).

Acerca desses resultados encontrados que evidenciam o uso por parte das mulheres, gostaríamos de destacar que das doze mulheres contidas em nossa amostra, nove exercem funções em casa, as demais trabalham exercendo profissões como agente comunitária, enfermeira, professora, costureira e estudante. Fizemos menção às atividades, por considerarmos que os papéis sociais desenvolvidos pelos falantes acarretam esse uso das

regras formais da concordância de número na região do CRAJUBAR, embora não tenhamos controlado a profissão como variável.

Após a apresentação dos resultados da segunda rodada, decidimos realizar o cruzamento entre as variáveis sociais *sexo vs. faixa etária* e *sexo vs. anos de escolaridade*. O cruzamento visa observarmos melhor o comportamento dos homens e das mulheres relacionando-os à escolaridade, ressaltando que essa variável não foi selecionada pelo programa Goldvarb X, por conseguinte, buscamos analisar sua sensibilidade ao cruzarmos com a variável sexo.

Quando na análise de fatores sociais **A** favorece a regra, e **B** favorece a regra, no entanto os dois juntos **A + B** desfavorecem a regra – isso raramente acontece com fatores linguísticos. É comum aos fatores sociais como sexo, idade e outros não se comportarem independentemente uns dos outros. Portanto, o pesquisador deve observar os resultados para os grupos mencionados fazendo cruzamentos usando o programa *Crosstab*. Esse tipo de análise com fatores sociais é possível aos usuários do Golvarb X e examina a possibilidade de interação entre eles. O programa executa a rodada fazendo o cruzamento entre dois grupos de fatores como, por exemplo: *sexo vs. escolaridade*. (GUY; ZILLES, 2007, p. 221).

4.1.4 Cruzamentos entre fatores sociais

O objetivo dos cruzamentos das variáveis *sexo vs. faixa etária* e *sexo vs. anos de escolaridade* é observarmos com mais detalhes o comportamento das mulheres e dos homens no processo de marcação de plural no SN. A tabela 8 nos traz os resultados obtidos do cruzamento das variáveis *sexo vs. faixa etária*.

Tabela 8 – Frequência da “presença de marcas explícitas de plural” segundo o cruzamento das variáveis *sexo vs. faixa etária*

Fatores	Masculino	Feminino	
Faixa etária	Aplic./Total Frequência	Aplic./Total Frequência	Total/Frequência
15 – 25 anos	531/582 = 91%	224/266 = 84%	755/848=89%
26 – 49 anos	485/775 = 63%	435/539 = 81%	920/1314=70%
50 em diante	270/349 = 77%	549/746 = 74%	819/1095=75%
Total	1286/1706 = 75%	1208/1551 = 78%	2494/3257 = 77%

Fonte: Elaborado pela autora (2016).

Ao observarmos os resultados desse cruzamento, percebemos que os jovens de 15 a 25 anos, tanto do sexo masculino (91%) quanto do feminino (84%), lideram as maiores percentagens favoráveis à marcação de plural na concordância nominal na região do CRAJUBAR. Os homens apresentam um *padrão curvilíneo*, os adultos – 26 a 49 anos – com 63% apresentam uma diminuição na retenção de marcas de plural e tornam a aumentar com os falantes de 50 anos em diante (77%).

Quanto às mulheres, apresentam um declínio gradativo à medida que aumenta a idade. Dessa forma, encontramos, dentro da comunidade de fala CRAJUBAR, mulheres de 15 a 25 anos com 84%, e de 26 a 49 anos com 81%, e de 50 anos em diante 74%. Então, de acordo com o cruzamento entre *sexo e faixa etária*, o comportamento das mulheres nos leva a uma perda de marcação de plural com o aumento da idade, não com o mesmo declínio, mas como acontece também com os homens.

Portanto, embora seja pequena a diferença entre as percentagens obtidas para homens e para mulheres, o resultado total confirma que as mulheres são mais favoráveis ao uso de *marcação de plural no SN* do que os homens.

Na tabela 9, destacaremos os resultados obtidos ao cruzarmos *sexo vs. anos de escolaridade*.

Tabela 9 – Frequência da “presença de marcas explícitas de plural” segundo o cruzamento das variáveis *sexo vs. anos de escolaridade*

Fatores	Masculino	Feminino	Aplic./Total/Frequência
	Aplic./Total/Frequência	Aplic./Total/Frequência	
1 a 8 anos	426/600 = 71%	746/926 = 81%	1172/1526 = 77%
9 a 11 anos	860/1106 = 78%	462/625 = 74%	1322/1731 = 76%
Total	1286/1706 = 75%	1208/1551 = 78%	2494/3257 = 77%

Fonte: Elaborado pela autora (2016).

Podemos perceber, na tabela 9, que o favorecimento à *marcação de plural no SN* encontra-se nas mulheres inclusas no período 1 a 8 anos de escolarização (81%), já os homens apresentam 71%. No período de 9 a 11 anos de escolaridade, são os homens que se mostram mais favoráveis à marcação de plural (78%) do que as mulheres (74%).

Observando os índices gerais dos resultados, as mulheres apresentam uma média maior de marcação de plural (78%) do que os homens (75%). O interessante é que o diferencial para retenção de plural encontra-se nos falantes que apresentam um tempo menor de escolaridade (1 a 8 anos).

De acordo com os resultados obtidos, é possível afirmar que os falantes mais jovens da amostra (15 a 25 anos) apresentam tendência de uso para mais presença de marca de plural no SN do que falantes com 9 – 11 anos de escolarização.

No tocante as variáveis sociais, os resultados obtidos nos mostram que o comportamento do fenômeno *presença ou ausência da marca de plural na região CRAJUBAR* não destoa muito das pesquisas apresentadas. Observamos que os grupos de fatores extralinguísticos apresentam-se significativos no uso de marcas de plural no SN, principalmente *sexo* e *faixa etária*, visto que a variável *escolaridade* não foi selecionada. Quanto à faixa etária, os jovens de 15 a 25 anos apresentam-se como os mais sensíveis ao uso da presença *de marcas explícitas de plural no SN* na região do CRAJUBAR. Com relação à variável *sexo*, as mulheres, mesmo apresentando índices próximos aos dos homens, apresentam tendência favorável à retenção de *marcas de plural*.

No capítulo a seguir, apresentamos as considerações finais acerca da variável *presença ou ausência de marcas explícitas de plural no falar do Cariri cearense*.

5 CONSIDERAÇÕES FINAIS

Nesta pesquisa, buscamos analisar a variável *concordância nominal de número entre os constituintes flexionáveis do sintagma nominal* (SN), sob o arcabouço da Teoria da Variação e Mudança desenvolvida por Labov e seus colaboradores. Os dados foram coletados a partir de 24 entrevistas do *corpus O Português falado no Ceará* inserido no PROFALA. Os informantes pertencem aos municípios do Crato, Juazeiro e Barbalha, região CRAJUBAR,

A estratificação dos falantes apresenta-se da seguinte forma: 12 informantes do sexo masculino e 12 do sexo feminino, distribuídos em três faixas etárias – seguindo as presentes no PROFALA (15-25 anos; 26-49 anos e 50 anos ou mais).

Quanto ao nível de escolarização, analisamos dois níveis: de 1 a 8 anos de escolarização e de 9 a 11 anos de escolarização. Assim, para descrevermos o fenômeno, coletamos todos os sintagmas nominais com marca formal ou explícita de plural e também aqueles que apresentam marca semântica de plural – “[...] *dez pessoas aqui* [...]” (AMA – 63). A análise considerou todos os elementos do sintagma nominal como unidade de análise, perspectiva chamada, por Scherre (1988), de atomística, totalizando 3304 dados após codificados e submetidos ao programa Goldvarb X.

Considerando que a variação linguística não acontece de forma aleatória, mas de forma ordenada – motivada por fatores linguísticos e extralinguísticos –, controlamos os seguintes fatores linguísticos: *posição nuclear; posição e classe gramatical em relação ao núcleo; classe gramatical do sintagma nominal; processos morfofonológicos de formação de plural; tonicidade das sílabas dos itens lexicais singulares e marcas precedentes de plural no âmbito do sintagma nominal* e fatores extralinguísticos; *sexo; escolaridade e faixa etária*. Desses fatores, foram excluídos pelo Goldvarb, após rodada inicial: *classe gramatical, tonicidade das sílabas dos itens lexicais e escolarização*.

Os resultados para a variável *posição linear* do constituinte no sintagma foram uniformes com os resultados de outras pesquisas. É a primeira posição bastante favorável à presença de plural.

No tocante à *posição e classe gramatical em relação ao núcleo*, os resultados demonstram que os constituintes antepostos ao núcleo se apresentam favoráveis à presença de plural no SN na região CRAJUBAR, ficando os constituintes pospostos ao núcleo, principalmente a primeira posição, com índices que não favorecem a presença de marca de plural no SN. Nesse caso, o que está determinando não é a classe, mas a posição em relação ao núcleo no SN.

Quanto à variável *marcas precedentes*, em análise a partir da terceira posição no SN, identificamos a influência favorável do fator *presença de duas ou mais marcas formais de plural precedendo o elemento sob análise* com peso relativo significativo para *marcação de plural*, seguido da presença de *marcas formais na primeira posição*. Os dados nos levaram a observar que a carga semântica posta aos numerais também acarreta resultados bastante favoráveis ao uso de marcas de plural. Já em *misturas de marcas com zero*, precedendo o elemento sob análise, o peso relativo apresenta-se muito baixo, visto que só encontramos uma ocorrência, constatando que a *presença de marca formal de plural antecedendo o constituinte sob análise* leva à presença de marca, enquanto a *ausência de marca antecedendo o constituinte* leva à ausência de marca.

Em relação à variável *processos morfofonológicos de formação de plural*, os constituintes presentes no fator *plural duplo com alternância vocálica, podendo haver ou não inserção de -s*, todos eles foram marcados ratificando o princípio da *saliência fônica*. O fator *plural com alteração silábica nos itens terminados em l, podendo haver ou não inserção de -s*, e *plural nos itens terminados em -r, com inserção de -e ou -es*, ou seja, plural com formação irregular, destacaram-se e se mostraram favoráveis à marcação de plural, opondo-se visivelmente aos constituintes com plural regulares e aos itens *terminados em -s, com inserção de -es*.

Enfim, com relação às variáveis linguísticas selecionadas pelo Goldvarb, nossa primeira hipótese quanto aos fatores linguísticos no âmbito do sintagma nominal, supomos, atuam, no âmbito do sintagma nominal, como possíveis condicionadores da variação da concordância nominal de número, no falar da região CRAJUBAR. Diante dos resultados, comprovamos essa influência dos fatores linguísticos na variação da concordância nominal de número no falar da região do CRAJUBAR.

Quanto aos fatores extralinguísticos, causou- nos surpresa a variável *escolaridade* não ser selecionada, visto que, em várias pesquisas, ela se mostra muito relevante ao uso das regras do fenômeno *concordância nominal de número*. Acreditamos que essa não seleção se deve ao fato de termos controlado apenas dois níveis de escolaridade. Então, foram selecionadas pelo programa estatístico as variáveis sociais *sexo* e *faixa etária*.

No tocante à variável *sexo* na região CRAJUBAR, nossa hipótese se confirma, tendo em vista que são as mulheres quem mais utilizam as *marcas explícitas de plural* na *concordância nominal de número* do CRAJUBAR. Contudo, os homens não destoam tanto

dos índices apresentados pelas mulheres. Conforme a literatura da teoria sociolinguística, são as mulheres quem mais utilizam as variáveis de prestígio não estigmatizadas.

No que se refere à variável *faixa etária*, nossa hipótese não se confirmou, visto que são os jovens (15 a 25 anos) que favorecem a *presença de marcas explícitas de plural*, com uma diminuição nos índices dos adultos (26 a 49 anos); e informantes com 50 anos em diante.

Sobre os dois cruzamentos *sexo vs. faixa etária* e *sexo vs. escolarização*, os resultados só confirmam que mulheres favorecem, com pouca diferença dos homens, a *presença de marcas explícitas de plural* e que são os jovens (15 a 25 anos) quem mais retêm as marcas de plural na concordância nominal de número.

Quanto à terceira rodada sem as variáveis *classe gramatical* e *posição linear*, o Goldvarb X selecionou em primeiro lugar, por ordem de significância, a variável *posição e classe gramatical em relação ao núcleo*. Os índices obtidos ratificaram a primeira posição anteposta ao núcleo como a que mais retém marcas explícitas de plural, assim como os constituintes que ocupam a função de núcleo na primeira posição do SN.

Neste estudo, podemos dizer que os resultados encontrados na região do CRAJUBAR, municípios Crato, Juazeiro e Barbalha, do Cariri cearense, não destoam muito das outras cidades do Brasil em que a concordância de número foi investigada.

Ressaltamos, também, que o fenômeno em estudo se mostra em variação, motivada não só por fatores morfossintáticos, mas, principalmente, por fatores sociais.

Esperamos que o nosso trabalho *Concordância nominal de número na região do CRAJUBAR* à luz da Teoria da Variação e Mudança venha a contribuir, de alguma maneira, para compreendermos quais contextos linguísticos e extralinguísticos são favoráveis ao uso da variante *concordância nominal de número na região do Cariri cearense*. Por conseguinte, esperamos contribuir com os estudos sociolinguísticos do português brasileiro.

REFERÊNCIAS

ABAURRE, Maria Luiza M.; PONTARA, Marcela. **Gramática - Texto: análise e construção de sentido**. São Paulo: Moderna, 2006.

AMARAL, Emília. *et al.* **Novas Palavras**. Nova edição, v. 3. São Paulo: FTD, 2010.

ANDRADE, Leila Minatti. **Rupturas e contínuos de concordância nominal de número em textos orais de informantes de Tubarão (SC) e São Borja (RS)**. Dissertação (Mestrado em Ciências da Linguagem). Tubarão: Universidade do Sul de Santa Catarina, 2003.

ANDRADE, Patrícia Ribeiro de. **Um fragmento da constituição sócio-histórica do português do Brasil: variação na concordância nominal de número em um dialeto afro-brasileiro**. Dissertação de mestrado, UFBA, Instituto de Letras, 2003.

BAGNO, Marcos. **Não é errado falar assim!** Em defesa do Português Brasileiro. São Paulo: Parábola Editorial, 2009, p. 166.

BECHARA, Evanildo. **Moderna Gramática Portuguesa**. 37. ed. Rio de Janeiro: Lucerna, 2001, p. 543-547.

BORTONI-RICARDO, Stella Maris. **The urbanization of ruras dialect speakers: A sociolinguistic study in Brazil**. New York, Cambridge University Press, 1985.

_____. **Educação em língua materna: a sociolinguística na sala de aula**. São Paulo: Parábola Editorial, 2004.

BRAGA, Maria Luiza; SCHERRE, Maria Marta Pereira. A concordância de número no SN na área urbana do Rio de Janeiro. In: **Encontro Nacional de Linguística**, Anais... Rio de Janeiro: PUC, 1976.

BRAGA, Maria Luiza. **A concordância de número no sintagma nominal no triângulo mineiro**. Dissertação de Mestrado. Rio de Janeiro: PUC, 1977.

BRANDÃO, Sílvia Figueiredo. Concordância Nominal. In: VIEIRA, S. R.; BRANDÃO, F. S. **Ensino de Gramática: descrição e uso (orgs.)**. São Paulo: Contexto, 2014.

CAMACHO, R. G. Sociolinguística. In: MUSSALIM, F; BENTES, A.C. (orgs.). **Introdução à Linguística: domínios e fronteiras**. Vol. 1, 6. ed. São Paulo: Cortez, 2006, p. 49.

CARVALHO Hebe Macedo de. **A Alternância Indicativo/Subjuntivo nas orações substantivas em função dos tempos verbais presente e imperfeito na língua falada do Cariri**. Fortaleza: UFC, doutorado em Linguística, 2007, p.74

_____. **Concordância Nominal: uma análise variacionista**. Dissertação (Mestrado em Letras). João Pessoa: UFPB, 1997.

CARVALHO, Raimunda Coelho de. **A concordância de número no sintagma nominal na fala urbana de Rio Branco**. Campinas: UNICAMP, mestrado em Linguística, 1997.

CIPRO NETO, Pasquale; INFANTE, Ulisses. **Gramática da Língua Portuguesa**. São Paulo: Scipione, 2008.

FERNANDES, Marisa. **Concordância nominal na Região Sul**. 1996. Dissertação (Mestrado em Linguística). Universidade Federal de Santa Catarina. Florianópolis, 1996.

FIORIN, José Luiz. **Introdução à Linguística**. São Paulo:Contexto, 2002, p. 12.

FISCHER, J. **Social influence of a linguistic variant**. *Word*, v. 14, 1958, p.47-56.

FREIRE, Maria Valeria da Silva; FERREIRA, Francisco Diego Guedes; LIMA, Maria Messias Ferreira. **Anais do II Colóquio de Geografia Agrária do Cariri Cearense – Ano 2014** ISBN: 978.8565425-17-9 79 O paradigma rural x urbano: uma tipologia para o Ceará e a microrregião do Cariri, 2014, p.83

GIACOMOZZI, G; VALÉRIO, G; REDA, C. M. **Descobrimos a Gramática: língua portuguesa, 8º ano**. Nova edição. São Paulo: FTD, 2010.

GRACIOSA, D. M. D. **Concordância verbal na fala culta carioca**. Dissertação de Mestrado. UFRJ, Rio de Janeiro: UFRJ, 1991.

GRYNER, H. **A variação de concordância com verbos impessoais na cidade de Petrópolis**. UFRJ, Rio de Janeiro. 139p. Dissertação de Mestrado, 1977.

GUIMARÃES, Maria Aparecida Souza. **Concordância nominal de número no português popular do Brasil: estudo de variação e mudança no vernáculo conquistense**. Dissertação de mestrado. Programa de Pós-Graduação em Linguística, UESB, 2014.

GUY, G. R. **Linguistic variation in Brazilian Portuguese: aspects of the phonology, syntax, and language history**. Philadelphia, University of Pennsylvania. 391p. Ph.D. Dissertation, mimeo, 1981.

GUY, Gregory. **A Identidade lingüística da comunidade de fala: paralelismo interdialeto nos padrões de variação lingüística**. *Organon, Revista do Instituto de Letras da UFRGS*, Porto Alegre, v. 28 e 29, 2000, p. 18-21

GUY, R. Gregory; ZILLES, Ana. **Sociolinguística Quantitativa: instrumental de análise**. São Paulo: 2007, p. 22.

HUDSON, R. A. **Sociolinguistics**. Cambridge: Cambridge University Press. 1980, p. 1.

KENEDY, Eduardo. **Curso básico de Linguística Gerativa**. São Paulo, Contexto, 2013, p. 18.

LABOV, W. Where does the linguistic stop? A response to Beatriz Lavandera. In: **Working Papers Sociolinguistics**, nº. 44, 1978.

- _____. **Principles of linguistic change: internal factors.** Oxford: Blackwell, 1994, p. 83.
- _____. Some Sociolinguistic Principles. *In:* PAULSTON, C. B. e TUCKER, G. R. (ors) *Sociolinguistics. The essential Readings.* Blackwell Publishing, 2003, p. 245.
- _____. **Principles of linguistic change: social factors.** Oxford: Blackwell, 2001.
- _____. **Padrões Sociolinguísticos.** Trad. de Marcos Bagno, Maria Marta Pereira Scherre, Carolina Rodrigues Cardoso. São Paulo: Parábola, 2008, p. 13-244.
- LAVANDERA, Beatriz. Where does Sociolinguistic variable stop? *In: Language Society.* n. 7, 1978, p. 171-182.
- LEMLE, Miriam; NARO; Anthony Julius. **Competências básicas do português.** 1977. Relatório final de pesquisa. Apresentado às instituições patrocinadoras. Fundação Movimento Brasileiro de Alfabetização (MOBRAL) e Fundação Ford, Rio, 1977.
- LOPES, Norma da Silva. **Concordância Nominal, Contexto Linguístico e Sociedade.** 2001. Tese (Doutorado em Letras). Universidade Federal da Bahia, 2001.
- MARTINS, Flávia Santos. **Variação na concordância nominal de número na fala dos habitantes do Alto Solimões (Amazonas).** 2013. Tese (Doutorado em Linguística). Universidade Federal de Santa Catarina, 2013.
- MEIRA, Gilberto Almeida. **Estudo comparativo entre as normas popular e culta do português de Vitória da Conquista: concordância nominal de número.** 2015. Dissertação (Mestrado em Linguística). UESB, Bahia, 2015.
- MEYERHOFF, Miriam. **Introducing Sociolinguistics.** London and New York: Routledge, 2006, p. 1-24.
- MOLLICA, Cecília. Fundamentação teórica: conceituação e delimitação. *In:* MOLLICA, Cecília; BRAGA, Maria Luiza (orgs.). **Introdução à Sociolinguística: o tratamento da variação.** São Paulo: Contexto, 2003, p. 9-14.
- MONTEIRO, José Lemos. **Para compreender Labov.** Petrópolis, RJ: Vozes, 2000.
- NARO, A. J. **The social and structural dimensions of syntactic change.** *Language LSA,* 1981.
- NARO A. J; SCHERRE M. M. P. **Sociolinguistic correlates of negative evaluation: variable concord in Rio de Janeiro.** *Language Variation and Change,* v.26, 2014, p.338.
- NICOLAU, E. M. dos D. **A ausência de concordância verbal em português: uma abordagem sociolinguística.** Dissertação (Mestrado em Ciências da Linguagem). UFMG, Belo Horizonte, 1984.
- NINA, T. de J. C. **Concordância nominal/verbal do analfabeto na microrregião de Bragantina.** Dissertação (Mestrado em Linguística). PUC, Rio Grande do Sul, 1980.

PAIVA, Maria da Conceição. A variável gênero/sexo. In MOLLICA, Cecília; BRAGA, Maria Luiza (orgs.). **Introdução à Sociolinguística: o tratamento da variação**. São Paulo: Contexto, 2003, p. 33-34.

PAIVA, M. da C. A. de; DUARTE, M. E. L. Quarenta anos depois: a herança de um programa na Sociolinguística brasileira. In: WEINREICH, U; LABOV, W; HERZOG, G. **Fundamentos empíricos para uma teoria da Mudança Linguística**. Trad. Marcos Bagno. São Paulo: Parábola Editorial, 2003, p. 179.

PLANO. **Territorial do Desenvolvimento Rural e Sustentável: Cidadania do Cariri-
MDA/SDT/Agropolos**. Fortaleza: Instituto Agropolos do Ceará, v. 1, 2010.

PONTE, V. M. L. A. **A concordância nominal de uma comunidade de Porto Alegre**. 1979. Dissertação (Mestrado em Linguística). Rio Grande do Sul: PUC, 1979.

POPLACK, Shana. **The notion of the plural in Puerto Rico Spanish: competing constraints on /s/ deletion**. In: LABOV, William. (eds.). *Locating language in time and space*. Philadelphia, University of Pennsylvania, 1980.

QUEIROZ, Ivan da Silva. **Região metropolitana do Cariri cearense, a metrópole fora do eixo**. Revista Mercator, Fortaleza, V.13, n.3, p. 93 – 104, set/dez. 2014.

RODRIGUES, A. C. de S. **A concordância verbal no português popular em São Paulo**. Dissertação (Mestrado em Linguística). USP, São Paulo, 1987.

SANTOS, Lilia Soares Miranda. **Sobre a ausência de concordância nominal no português falado em Pedro Leopoldo-MG: uma abordagem variacionista**. Dissertação de mestrado. Belo Horizonte: UFMG, 2010.

SARMENTO, Leila Lauer. **Gramática em textos**. São Paulo: Moderna, 2012, p. 488.

SAUSSURE, Ferdinand de. **Curso de Linguística Geral**. Trad. Antonio Chelini, José Paulo Paes, Izidoro Blikstein. São Paulo: Cultrix, 2001.

SCHNEIDER, Simone Daise. **Concordância nominal na fala de crianças de 3 a 6 anos de idade do município de Novo Hamburgo: variação linguística na infância**. 2012. Tese (Doutorado em Linguística Aplicada). Universidade Federal do Rio Grande do Sul. Porto Alegre, 2012.

SCHERRE, Maria Marta Pereira. **A regra de concordância de número no sintagma nominal de Porto Alegre**. Dissertação (Mestrado em Linguística). Rio de Janeiro: PUC, 1978, p. 66.

_____. **Reanálise da concordância nominal em português**. 1988. Tese de Doutorado. Rio de Janeiro, UFRJ, 1988.

_____. Aspectos da concordância de número no português do Brasil. **Revista Internacional de Língua Portuguesa (RILP) - Norma e Variação do Português**. Associação das Universidades de Língua Portuguesa, 1994.

_____. **Doa-se lindos filhotes de poodles**: variação linguística, mídia e preconceito. São Paulo: Parábola, 2005, p. 133.

_____. (1992). The serial effect on internal and external variables. **Language Variation and Change**. Cambridge University Press. 4(1):1-13.

_____. Duas dimensões do paralelismo verbal no português popular do Brasil. **Delta**. São Paulo, 1993, p. 1-14.

TARALLO, Fernando. **A pesquisa sociolinguística**. 8. ed. São Paulo: Ática, 2007, p. 6 – 65.

VOTRE, Sebastião Josué. Relevância da variável escolaridade. MOLLICA, Cecília; BRAGA, Maria Luiza (orgs.). **Introdução à Sociolinguística**: o tratamento da variação. São Paulo: Contexto, 2003, p. 54.

WEINER, J. & LABOV, W. Constraints on the agentless passive. In.: **Journal of Linguistics** 19, [1977], 1983.

WEINREICH, U; LABOV, W; HERZOG, M. **Fundamentos empíricos para uma teoria da mudança linguística**. São Paulo: Parábola Editorial, 2006, p. 36-126.