



UNIVERSIDADE FEDERAL DO RIO DE JANEIRO  
FACULDADE DE LETRAS

**“AS PESSOA LÁ DO OUTRO LADO NÃO CONCORDA  
COM AS FRASE QUE EU ESCREVO”: A CONCORDÂNCIA  
NOMINAL EM ÁREA RURAL DO RIO DE JANEIRO**

Luiza de Carvalho de Lima

Rio de Janeiro  
2022

LUIZA DE CARVALHO DE LIMA

“AS PESSOA LÁ DO OUTRO LADO NÃO CONCORDA  
COM AS FRASE QUE EU ESCREVO”: A CONCORDÂNCIA  
NOMINAL EM ÁREA RURAL DO RIO DE JANEIRO

Monografia submetida à Faculdade de Letras da  
Universidade Federal do Rio de Janeiro, como  
requisito parcial para obtenção do título de  
Licenciada em Letras na habilitação Português -  
Literaturas.

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Silvia Rodrigues Vieira

Rio de Janeiro  
2022

Lima, Luiza de Carvalho de.

“As pessoa lá do outro lado não concorda com as frase que eu escrevo”: a concordância nominal em área rural do Rio de Janeiro / Luiza de Carvalho de Lima - 2022.

44 f.

Orientadora: Prof<sup>ª</sup>. Dr<sup>ª</sup>. Silvia Rodrigues Vieira.

Monografia (graduação em Letras habilitação Português – Literaturas) – Universidade Federal do Rio de Janeiro, Centro de Letras e Artes, Faculdade de Letras.

Bibliografia: f. 42-44

1. Concordância nominal; 2. Sociolinguística; 3. Ruralidade. I. LIMA/ Luiza de Carvalho de; II. Vieira, Silvia Rodrigues (orient.) III. Universidade Federal do Rio de Janeiro, Faculdade de Letras; IV. “As pessoa lá do outro lado não concorda com as frase que eu escrevo”: a concordância nominal em área rural do Rio de Janeiro.

## AGRADECIMENTOS

Agradeço transcendentemente...

Em primeiro lugar, a Deus e ao mundo espiritual, por todo amor infinito, luz e bondade suprema, por não terem me deixado desistir dessa pesquisa e de mim.

Aos meus pais, Andrea e Fabio, por todo auxílio nos últimos 21 anos e por todos os sacrifícios que fizeram para que eu pudesse estar onde estou, e à minha avó Maria, por ter me dado morada e aconchego no Rio de Janeiro. Obrigada por todo carinho.

À minha orientadora, professora Dr<sup>a</sup> Silvia Rodrigues Vieira, por ter me iluminado tanto nessa pesquisa. Nunca imaginei que uma simples ligação telefônica me faria conhecer um mundo completamente diferente. Obrigada por todo auxílio para que esse trabalho pudesse ganhar vida. Isso seria impossível sem a sua parceria. Você é meu maior exemplo.

Aos amigos da Faculdade de Letras, em especial Isabel Lessa, Mayara Gonçalves, Milena Velloso e Palloma Pimentel, por sempre serem minhas parceiras e compartilharem corredores, bibliotecas e sonhos comigo. A vida universitária foi mais fácil ao lado de vocês.

Aos meus professores do colégio, em especial Luthiene Dalanhese, por não ter saído do meu lado desde meus 15 anos e por sempre me incentivar a seguir meus sonhos, e Carlos Matias, por todo companheirismo durante esse tempo. Obrigada por acreditarem em mim.

A todos os meus amigos, em especial Alan Dutra, Alice Carriço, Carolina Valladares, Felipe Frare, Heloisa Mendes e Isabella Cantarino, pelo apoio incondicional nos últimos anos, mesmo em longa distância às vezes.

À professora Marcia Damaso (*in memoriam*). Você cativou em mim um amor inexplicável pela sintaxe em suas aulas. Onde quer que esteja, que essa pesquisa te dê orgulho.

À Pró-Reitoria de Extensão da UFRJ, por terem me acolhido desde 2019. O apoio que recebi nesses últimos anos foi fundamental para que esse trabalho ganhasse vida e se tornasse, como vocês sempre me ensinaram e incentivaram, algo para além dos muros da universidade.

E a todos que contribuíram para que eu pudesse estar aqui. Essa pesquisa é resultado de muito amor pela Língua Portuguesa e de muita fé na educação brasileira. Que nós nunca nos esqueçamos do porquê de estarmos fazendo isso.

*“O senhor é tão jovem, tem diante de si todo o começo, e eu gostaria de lhe pedir da melhor maneira que posso, meu caro, para ter paciência em relação a tudo que não está resolvido em seu coração. Peço-lhe que tente ter amor pelas próprias perguntas, como quartos fechados e como livros escritos em uma língua estrangeira. Não investigue agora as respostas que não lhe podem ser dadas, porque não poderia vivê-las. E é disto que se trata: de viver tudo. Viva agora as perguntas.”*

Rainer Maria Rilke.

## RESUMO

Este trabalho surgiu do questionamento a respeito do comportamento referente à concordância de número no sintagma nominal em áreas menos urbanas/mais rurais (Guapimirim e Cachoeiras de Macacu) do estado do Rio de Janeiro. Vinculado ao projeto *Contínuos de/em variedades do Português: análises contrastivas* (área: Linguística – Sociolinguística e Dialetoologia), esta pesquisa pretende descrever e analisar possíveis relações entre a concordância dentro do SN a fatores linguísticos e extralinguísticos, a partir dos preceitos da Teoria da Variação e Mudança (WEINREICH; LABOV; HERZOG, 2006 [1968]; LABOV, 2008 [1972]) e de resultados de estudos anteriores acerca da variação na concordância nominal, de forma a contribuir com a análise dos padrões de concordância e com a investigação acerca do *continuum* ruralidade-urbanidade, buscando identificar os traços característicos da comunidade de fala em questão. De modo geral, a análise dos resultados de marcação de pluralidade permitiu fazer considerações sobre a concordância nominal nas áreas rurais, de forma a contribuir, também, com o registro de dados acerca da constituição do Português Brasileiro e suas variedades, uma vez que revelaram que o comportamento nas regiões mais rurais se apresenta, em alguma medida, desfavorável ao uso de marcas, sobretudo quando comparado ao das regiões urbanas contempladas em outros estudos com amostras contemporâneas.

**PALAVRAS-CHAVE:** Concordância nominal; Sociolinguística; ruralidade.

## SUMÁRIO

1. INTRODUÇÃO .....	8
2. PRESSUPOSTOS TEÓRICO-METODOLÓGICOS ELEMENTARES SOBRE A VARIACÃO LINGUÍSTICA .....	11
3. BREVE REVISÃO DA LITERATURA ACERCA DA VARIACÃO DA CONCORDÂNCIA NOMINAL NO RIO DE JANEIRO .....	13
4. METODOLOGIA .....	21
4.1 A comunidade de Guapimirim e Cachoeiras de Macacu .....	21
4.2 Descrição do <i>corpus</i> .....	23
4.3 Tratamento dos dados e a regra variável .....	23
4.4 Descrição das variáveis independentes investigadas .....	25
4.4.1 Variáveis extralinguísticas .....	25
4.4.2 Variáveis linguísticas .....	25
5. ANÁLISE DOS RESULTADOS .....	28
5.1 Distribuição geral dos dados .....	28
5.2 Variáveis estatisticamente relevantes .....	30
5.2.1 Variáveis extralinguísticas .....	30
5.2.2. Variáveis linguísticas .....	33
5.3 <i>Continuum</i> rural-urbano da concordância nominal: breve comparação entre resultados de pesquisas no Rio de Janeiro .....	37
6. CONSIDERAÇÕES FINAIS .....	40
7. REFERÊNCIAS BIBLIOGRÁFICAS .....	42

## 1. INTRODUÇÃO

Em um país como o Brasil, caracterizado pelo multilinguismo e pela diversidade de variedades da língua portuguesa, é inegável a importância de se abordar a variação linguística, sobretudo considerando os diferentes usos, seja nos centros urbanos ou nas áreas mais rurais. Dentre as diversas motivações para a investigação sociolinguística, importa, além de nos perguntarmos quanto a que se deve, em termos sócio-históricos, a variação, questionarmos sobre quais fatores linguísticos e extralinguísticos fomentam o uso de certas variantes por cada comunidade de fala.

Dentre os temas que manifestam a variação no Português Brasileiro, observamos que a concordância de número, verbal ou nominal, é, possivelmente, um dos aspectos mais considerados nas pesquisas variacionistas acerca da fala brasileira. A falta de expressão formal da concordância na fala parece ser um dos traços mais estigmatizados, sendo frequentemente considerada fruto da falta de escolarização e/ou típica de comunidades mais rurais/isoladas, de onde teriam surgido as normas/variedades populares, em oposição às normas/variedades cultas.

Estudos anteriores sobre o assunto realizados principalmente com dados do Rio de Janeiro – como os de Scherre (1988; 1994), Brandão (1994; 2013), Brandão; Vieira (2014) e Rosa (2016; 2019) – também fundamentam a formulação de hipóteses e a interpretação do fenômeno da variação de concordância nominal e suas motivações. Embora muitas pesquisas sociolinguísticas já tenham tratado do tema no âmbito de variedades do Português do Brasil, a partir de *corpora* constituídos em sua maioria na segunda metade do século XX, verifica-se a carência de descrição de dados contemporâneos em comunidades distribuídas no chamado *continuum* rural-urbano, sobretudo no polo mais rural/menos urbano, região normalmente caracterizada por maiores índices de não-marcação de plural.

Os resultados relativos ao fenômeno gramatical estudado revelam, em linhas gerais, que a marcação de plural no Sintagma Nominal (SN) – como em “os pais participam mais” – tende a ocorrer predominantemente nas variedades urbanas e de pessoas mais escolarizadas, enquanto a outra variante – como em “umas paisagem0 bonita0” – tende a ser mais realizada por falantes com baixo grau de escolaridade e das variedades mais rurais. Com isso,

presume-se que, enquanto o falante mais escolarizado apresenta maior emprego da concordância-padrão independentemente do contexto, o falante menos escolarizado apresenta menor uso de marcas de concordância-padrão.

Com base nesse contexto, o presente trabalho tem por objetivo descrever as ocorrências da marcação *versus* não-marcação de pluralidade no âmbito do sintagma nominal em entrevistas sociolinguísticas com falantes de Guapimirim e Cachoeiras de Macacu, cidades mais distantes da urbanidade típica das grandes metrópoles, observando o comportamento linguístico dessas comunidades em comparação com o obtido em estudos anteriores sobretudo em áreas urbanas centrais da região metropolitana do Rio de Janeiro. Supomos que as marcações de número nas comunidades mais rurais se apresentariam em menor frequência, diferentemente das relativas às comunidades mais urbanas (ao menos as controladas em estudos anteriores), onde o acesso à escolarização e o contato entre pessoas com diferentes idades, classes econômicas e escolaridades seriam maiores.

Assim, interessa responder à seguinte pergunta geral: como se apresenta o comportamento da marcação de número nos sintagmas nominais na fala de indivíduos que vivem nos municípios de Guapimirim / Cachoeiras de Macacu, cidades que ainda representam meios menos urbanos do RJ? Essa indagação abarca a problematização central do trabalho. Como hipótese geral, supõe-se que, mesmo com o vasto processo de urbanização e de escolarização no Brasil, além do acesso à tecnologia e aos meios/redes de comunicação social – o que fez mudar as feições das comunidades consideradas rurais –, ainda haveria diferença no comportamento da marcação de plural nos sintagmas nominais entre as áreas rurais e urbanas.

Desse modo, esta pesquisa tem como objetivo analisar o comportamento variável da expressão de plural, por meio do marcador de número -s, em sintagmas nominais, no âmbito da fala de moradores das cidades de Guapimirim e Cachoeiras de Macacu - cidades da região Metropolitana do estado do Rio de Janeiro que podem ser consideradas menos urbanas ou mais rurais -, a partir do *corpus* organizado por Corrêa (2019), em sua tese de doutorado, e ampliado por Souza, em 2020-2021<sup>1</sup>. O presente trabalho pretende compreender e apresentar as possíveis relações entre a concordância dentro do sintagma nominal e fatores linguísticos e

---

<sup>1</sup> Trata-se de atividade da iniciação científica de Jennifer Aguiar de Souza, orientada por Silvia Rodrigues Vieira no âmbito do Projeto Contínuos de/em variedades do Português: análises contrastivas.

extralinguísticos, a partir dos pressupostos da Teoria da Variação e Mudança (WEINREICH; LABOV; HERZOG, 2006 [1968]; LABOV, 2008 [1972], 2003), que consideram a variação como inerente a todo sistema linguístico.

Para responder à questão central, os propósitos formulados para a pesquisa são os seguintes: (i) averiguar se a marcação de número no SN apresentaria baixo índice de concordância nas áreas rurais, sobretudo em consonância com resultados de estudos linguísticos anteriores, como os obtidos por Brandão (cf. 1994; 2013); (ii) descrever os fatores que favorecem ou não a marcação de plural; e (iii) comparar brevemente os resultados obtidos em Guapimirim / Cachoeiras de Macacu com dados de estudos realizados nas zonas urbanas do Rio de Janeiro (SCHERRE, 1988; 1994; BRANDÃO, 2013; BRANDÃO & VIEIRA, 2014; ROSA, 2016; 2019). Com base nesses propósitos, o trabalho visa a colaborar com o registro da constituição do Português Brasileiro (PB) nas áreas rurais (menos urbanas), de modo a prover, em última instância, mais informações sobre a sócio-história da região e suas consequências.

Para desenvolver esse debate, esta pesquisa é estruturada por meio de cinco capítulos, além desta introdução, com a apresentação geral do trabalho. O Capítulo 2, que apresenta aspectos relacionados ao tratamento da variação linguística, aborda os fundamentos teórico-metodológicos elementares adotados com base na Teoria da Variação e Mudança (WEINREICH; LABOV; HERZOG, 1968; LABOV, 1972; 2003); no Capítulo 3, apresenta-se uma breve revisão da literatura acerca da variação na concordância nominal, com destaque aos principais resultados obtidos em pesquisas anteriores; no Capítulo 4, referente à metodologia, encontramos uma breve caracterização das comunidades de fala, a descrição do *corpus*, informações sobre os procedimentos para o tratamento dos dados e a descrição das variáveis linguísticas e extralinguísticas investigadas; o Capítulo 5 constitui-se da análise dos dados coletados no *corpus*, expondo e debatendo os resultados; e as considerações finais, na última seção, destacam as conclusões do trabalho.

Por fim, esperamos que, ao descrever e analisar quantitativamente e qualitativamente a expressão das marcas de número internas ao sintagma nominal na fala das comunidades rurais de Guapimirim / Cachoeiras de Macacu, do estado do Rio de Janeiro, esta pesquisa contribua e auxilie na investigação do fenômeno da concordância nominal no *continuum* rural-urbano do Português do Brasil.

## **2. PRESSUPOSTOS TEÓRICO-METODOLÓGICOS ELEMENTARES SOBRE A VARIAÇÃO LINGUÍSTICA**

A linha teórico-metodológica utilizada neste trabalho considera, como base, os pressupostos gerais da Sociolinguística Variacionista (WEINREICH; LABOV; HERZOG, 1968; LABOV, 1972; 2003). O objeto de estudo da Sociolinguística é, em essência, a fala viva em seu contexto de variação e possível mudança, uma vez que investiga as formas alternantes de expressão vernacular (a língua falada em situações espontâneas), os componentes linguísticos enraizados em um determinado grupo social – uma comunidade de fala –, com suas características próprias, seu comportamento linguístico, além de certa avaliação dos usos linguísticos partilhada.

A Sociolinguística Variacionista, também chamada de Teoria da Variação e Mudança ou Sociolinguística Quantitativa, pressupõe, assim, que a língua sofre influência do meio social e dos seus próprios condicionamentos estruturais. Conseqüentemente, os fatores linguísticos e extralinguísticos podem ou não favorecer o uso de determinadas regras. É a partir dessa perspectiva que a pesquisa se fundamenta, ao analisar e descrever formas variantes (no caso, a marcação ou a não marcação de plural nos vocábulos que constituem o sintagma nominal) em que essa língua se apresenta, que podem ou não evoluir para um quadro de mudança – situação em que uma das formas alternantes suplanta a outra.

Partindo da proposta de Labov (2003) a respeito do funcionamento do sistema linguístico e seus tipos de regras, a presente pesquisa tem por objetivo averiguar o estatuto do fenômeno da variação no sintagma nominal em relação a uma das três regras definidas pelo sociolinguista: categórica, semicategórica e variável. Em termos de produtividade das formas alternantes, enquanto a regra categórica opera com 100% de frequência na fala dos indivíduos de determinado grupo social, a regra semicategórica se apresenta em torno de 95-99%, alternando entre as variantes que ocorrem poucas vezes, mas que podem ser observadas, sendo, assim, raras, porém relatáveis; por fim, a regra variável se apresenta abaixo dos 95%, através da coexistência de variantes na língua.

Apesar de Labov (2003) assumir, nessa proposta, um perfil quantitativo em relação à produtividade das realizações das variantes linguísticas para caracterizar tais fenômenos em cada tipo de regra, é preciso atentar para a natureza dos dados em termos estruturais, observando os contextos em que as variantes se manifestam. Assim, esta pesquisa alia as

análises quantitativa e qualitativa para analisar os resultados de concordância nominal na área rural do Rio de Janeiro. Essa perspectiva foi apresentada por Brandão e Vieira (2014) ao destacarem a necessidade de atentar-nos ao perfil das regras quanto à quantidade e à qualidade das variantes, de modo a perceber que a alternância de formas nem sempre indicaria o *status* de uma regra efetivamente variável, mas, por vezes, um caso de regra semicategórica.

Assume-se, assim, que afirmar que uma língua/variedade admite, por opção gramatical, uma dada estrutura não implica necessariamente o registro categórico dessa estrutura, como se sabe, nem tampouco se pressupõe um comportamento efetivamente variável. Há que se verificar quantitativa – um número restrito de dados – e qualitativamente – contextos específicos em termos estruturais – a especialização dos usos para se determinar o parâmetro gramatical de certa língua/variedade. É essa a hipótese que norteia as reflexões feitas no presente capítulo. (BRANDÃO; VIEIRA, 2014, p. 86)

Assim, utilizando das concepções da Teoria Variacionista, o trabalho parte da análise de dados obtidos a partir do uso real da língua, para investigar sobretudo o chamado problema das restrições, referente aos fatores favorecedores ou desfavorecedores da marcação ou não marcação de plural interna ao SN; com base no comportamento das variáveis extralinguísticas e na comparação entre os resultados obtidos nas comunidades rurais e urbanas, o estudo pode, ainda, permitir reflexões acerca do problema do encaixamento da variação/mudança linguísticas. O problema da avaliação social das variantes, embora não seja objeto de estudo desta pesquisa, também se relaciona às nossas hipóteses interpretativas, tendo em vista ser o fenômeno estudado normalmente valorado como índice de prestígio ou desprestígio social.

A análise das ocorrências foi feita com o suporte do programa GoldVarb-X, após obtenção de dados dos sintagmas nominais extraídos de entrevistas sociolinguísticas realizadas nos municípios de Guapimirim e Cachoeiras de Macacu, codificados em relação a fatores linguísticos, ligados a propriedades estruturais, e extralinguísticos, referentes ao perfil dos falantes quanto a alguns fatores sociais, os quais estão descritos, adiante, no capítulo referente à metodologia.

### 3. BREVE REVISÃO DA LITERATURA ACERCA DA VARIAÇÃO DA CONCORDÂNCIA NOMINAL NO RIO DE JANEIRO

Com o propósito de investigar o fenômeno da variação na concordância nominal, diversos pesquisadores se propuseram a estudar o tema ao longo dos anos. Como referência, apresentaremos alguns resultados de pesquisa sobre a concordância no Brasil, destacando aqueles relativos à região do estado do Rio de Janeiro, como os apresentados em SCHERRE (1988; 1994), BRANDÃO (1994; 2013), BRANDÃO & VIEIRA (2014) e ROSA (2016; 2019).

SCHERRE (1994), ao sintetizar o que foi investigado em sua tese de doutorado (1988), apresenta o fenômeno variável da concordância de número como inerente ao Português do Brasil, uma vez que ocorre produtivamente em contextos sociais e linguísticos semelhantes, manifestando um direcionamento sistemático e previsível. Assim, segundo a pesquisadora, a variação, quanto à concordância na comunidade de fala brasileira em suas diversas variedades, apresenta diferenças maiores em relação à quantidade de marcas de plural, e não, necessariamente, aos contextos linguísticos em que a regra variável se manifesta. Para chegar a tal afirmação, Scherre analisou a fala de 64 informantes da cidade do Rio de Janeiro, com dados extraídos do *Corpus Censo do PEUL*<sup>2</sup>, constituído no período de 1980 a 1982. Esse conjunto de entrevistas sociolinguísticas, que contempla 48 falantes adultos (15-71 anos) e 16 crianças (7-14 anos), permite considerar os indivíduos classificados por sexo (homem e mulher), faixa etária (7-14; 15-25; 26-49 e 50-71) e anos de escolarização (1 a 4; 5 a 8 e 9 a 11).

A análise dos dados foi feita segundo duas perspectivas: atomística, segundo a qual cada elemento flexionável no sintagma nominal constitui uma unidade de análise, e não atomística, em que se considera o sintagma nominal como um todo, de modo a controlar apenas se houve alguma ausência de marcas (ou não). Em ambas as perspectivas, Scherre considerou, como formas alternantes, a presença *versus* ausência da forma explícita do plural. Analisadas as ocorrências, a pesquisadora observou – a partir do cruzamento entre as variáveis *posição linear do elemento no SN* e *classe gramatical do elemento nominal* – que a marca explícita do plural pode ser encontrada: (i) em todos os elementos flexionáveis do SN;

---

<sup>2</sup> PEUL: Programa de Estudos sobre o Uso da Língua, projeto desenvolvido na Faculdade de Letras da Universidade Federal do Rio de Janeiro.

(ii) em alguns elementos flexionáveis do SN; (iii) em apenas um dos elementos flexionáveis do SN; ou (iv) em nenhum dos elementos flexionáveis do SN.

Na análise atomística, os falantes adultos apresentaram índices de 72% de marcação. De acordo com os resultados, recebem mais marcas de plural explícitas todos os elementos determinantes antepostos ao núcleo (98% de marcação) e recebem menos marcas explícitas de plural todos os elementos pospostos ao núcleo (68% de marcação). Com relação à marcação no núcleo, Scherre observou que, na primeira posição do SN, núcleos são mais marcados (95%), enquanto os que ocupam a segunda posição do SN (54%) são menos marcados do que os núcleos encontrados nas demais posições (62%).

Ao analisar o sintagma nominal na perspectiva não atomística, Scherre chegou a um percentual de 51% de concordância. A pesquisadora concluiu que o SN recebe mais marcas explícitas de plural quando localizado à esquerda na oração (63%), enquanto quando localizado à direita na oração e em posições indistintas (sem posição nítida) receberam menos marcas (49% e 40%, respectivamente). Quanto à influência da configuração sintagmática, ou seja, a ordem e a classe dos elementos na oração para a presença de marcas explícitas de plural, a pesquisadora observou que a presença do artigo definido/quantificador (na primeira posição) acompanhado de um substantivo ou categoria substantivada (na última posição) favorecem o emprego de mais marcas explícitas de plural em todos os elementos do SN, ao passo que qualquer outra classe gramatical na primeira posição, acompanhada de adjetivo, possessivo ou quantificador, na última posição, desfavorece a marcação de plural no SN. Além disso, Scherre afirmou que as estruturas com sujeito composto desfavorecem a marcação de plural no SN.

No que concerne aos estudos sobre a concordância nominal no Português do Brasil, também no Rio de Janeiro, Brandão (cf. BRANDÃO, 2013; BRANDÃO; VIEIRA, 2014) desenvolve a investigação do tema, a partir da análise de dados contemporâneos extraídos de entrevistas realizadas entre 2009 e 2011 no âmbito do Projeto *Estudo comparado dos padrões de concordância em variedades africanas, brasileiras e europeias do Português*<sup>3</sup>. Para tal estudo, a pesquisadora analisou a concordância dos sintagmas nominais nas variedades europeia, africana (de São Tomé e Príncipe) e brasileira, a partir do *Corpus Concordância*,

---

<sup>3</sup> Trata-se de projeto de cooperação internacional, que se iniciou em 2008 e, a partir de 2011, passou a integrar os projetos da ALFAL - Associação de Linguística e Filologia da América Latina. No âmbito desse projeto, foi organizado o Corpus Concordância. Para maiores detalhes, consultar <https://corporaport.letas.ufrj.br>.

com dados de 18 informantes separados por sexo/gênero (homem e mulher), faixas etárias (18 a 35, 36 a 55 e 56 a 75) e níveis de escolaridade (5 a 8, 9 a 11 e mais de 11). Os dados do Português Brasileiro relevantes para o presente trabalho foram coletados a partir de gravações de Copacabana e Nova Iguaçu, na região Metropolitana do Rio de Janeiro.

Os resultados da referida pesquisa demonstram que variedades urbanas do Português Brasileiro (PB), segundo a análise atomística, apresentam regra variável com índices de 91,1% para marcação de número nos sintagmas nominais, enquanto o Português Europeu (PE) apresenta regra categórica ou semicategórica, chegando ao índice de 99,9% de marcação.

No que se refere aos condicionamentos da regra, as variáveis estruturais, como *saliência posicional*, *traço de animacidade* e *diferenciação entre as formas singular e plural*, e social, como *nível de escolarização*, *sexo* e *faixa etária*, foram importantes para a ausência/presença das marcas de número no SN no *corpus* brasileiro investigado. Destacam-se, porém, a *posição linear e relativa do constituinte* e o *nível de instrução do informante* como as mais relevantes. Em relação à variável posição linear e relativa do constituinte, constatou-se que os vocábulos em posições pré-nucleares tendem a receber mais marcas de número, ao passo que os vocábulos pós-nucleares registram menos marcas de plural. No que diz respeito à variável relativa ao nível de escolarização do informante, observou-se forte polarização entre os indivíduos com nível fundamental e superior, apresentando os falantes com escolaridade média um pouco menos marcas de número do que os falantes de nível fundamental. Assim, a referida pesquisa, baseada em Labov (2003), confirma que a regra de concordância nominal, na amostra do Português Brasileiro investigada, demonstra ser variável na fala dos indivíduos menos escolarizados, e semicategórica na dos mais escolarizados (de nível superior).

Brandão (1994), em análise anterior sobre a concordância nominal, considerando dados não urbanos – extraídos do *corpus* do Projeto do *Atlas Etnolinguístico dos Pescadores do Estado do Rio de Janeiro* (APERJ), constituído, assim como o do PEUL, na segunda metade do século XX –, observou, em uma análise atomística, que a marcação de número nos constituintes do sintagma nominal apresentou um índice menor, de 44% dos 2.865 SNs coletados, comprovando que o fenômeno se comporta como uma regra variável. Para a realização da análise, a pesquisadora coletou dados de entrevistas com 72 informantes homens, todos pescadores, moradores do Norte e do Noroeste fluminense, área considerada

rural do Rio de Janeiro, distribuídos por três faixas etárias (Faixa A – 18 a 35 anos; faixa B – 36 a 55 anos; faixa C – mais de 56 anos) e com nível de escolaridade de zero a quatro anos.

Ao total, foram obtidos 2.865 dados de SNs, nos quais a pesquisadora constatou que em 93,2% o núcleo se encontra em segunda posição, enquanto a primeira, terceira e quarta posições apresentaram índices menores de realização, com 0,6%, 6,11% e 0,3%, respectivamente, o que permite constatar que a configuração sintagmática do SN influencia o fenômeno da marcação de plural. Os resultados apontaram, também, que os falantes de maior idade apresentaram índices maiores de não marcação, confirmando a hipótese de que, no Português Brasileiro, a tendência seria de aquisição (e não perda) de marcas nos constituintes do SN.

Ainda em relação aos estudos sobre a concordância nominal, Rosa (2016; 2019) analisou a fala de 12 indivíduos moradores da Zona Oeste do Rio de Janeiro, a partir de 26 gravações obtidas, sendo 11 interações espontâneas (gravações secretas, autorizadas posteriormente pelo participante) e 15 gravações controladas (em entrevistas sociolinguísticas). O perfil dos informantes foi categorizado em sexo (11 mulheres e 1 homem), faixa etária (faixa A- 15 a 35 anos; faixa B- 36 a 55 anos; e faixa C- a partir de 56 anos) e escolaridade (Ensino Fundamental, Ensino Médio e Ensino Superior).

A pesquisadora considerou, na análise atomística, a variável dependente *marcação ou não marcação de plural*, e como variáveis independentes a *função do vocábulo no SN* (determinante/quantificador, nuclear ou modificador), *a posição do vocábulo no SN* (1ª posição, 2ª posição, 3ª posição, 4ª posição e 5ª ou 6ª posição), *a relação entre os constituintes do SN* (elemento nuclear, elemento não nuclear anteposto e elemento não nuclear posposto) e *o processo morfofonológico de formação do plural / saliência fônica* (itens terminados com marcas dupla de plural; e itens com as terminações L, ÃO, R, S, vogal oral e vogal nasal, todas no singular), *tonicidade do item no singular* (elemento oxítono, paroxítono, proparoxítono, monossílabo tônico e monossílabo átono), *marcas precedentes, e contexto fonológico subsequente* (consoante, vogal ou pausa). Para a análise não atomística, foi considerada como variável dependente a *concordância padrão* ou a *concordância não padrão*, e como variáveis independentes a *animidade do núcleo, classe do primeiro vocábulo do SN, número de palavras do SN, presença de palavra pós-núcleo e paralelismo formal*.

Na análise atomística, foram coletados e analisados 1956 dados. Desse total, 85,6% apresentaram marcas formais de marcação de plural, contra 14,4% dos elementos com ausência de marcas de plural. No que diz respeito à relevância das variáveis, a *posição linear e relativa do constituinte no SN* e a *saliência fônica* foram as que apresentaram maiores índices de relevância para a marcação de número dos elementos. No que concerne à *posição linear e relativa do constituinte no SN*, Rosa observou que a posição pré-núcleo tende a apresentar mais marcas de marcação de plural, sendo a primeira posição com 99% de marcação e a segunda, com 94,3%, enquanto os elementos pós-núcleo no sintagma tendem a ser menos flexionados, visto que os elementos de segunda posição receberam marcas em 88,6% das ocorrências. Do mesmo modo, os elementos de terceira posição receberam 62,7% de marcas, índice acima dos elementos de quarta e quinta posição, que receberam apenas 38,5% de marcação. Em relação à saliência fônica, a pesquisadora observou que os elementos terminados em *-l* e *-r* e os terminados em *-s* apresentaram índices, respectivamente, de 99,1% e 98,3%, favorecendo a marcação, ao passo que os elementos terminados em vogal nasal e vogal oral desfavorecem a marcação, com 90,2% e 83,7%, respectivamente, confirmando a hipótese de que os elementos com traços [+ saliência] favorecem a marcação, em oposição aos [- saliência], que a desfavorecem.

No que se refere à análise não atomística, dos 1031 dados de SN coletados, 786 (76,2%) apresentaram a concordância padrão, contra 245 dados (23,8%) com a concordância não padrão. Nas rodadas multivariadas, as variáveis consideradas mais relevantes para a marcação foram *registro*, *paralelismo formal* e *classe do primeiro vocábulo do SN*. A pesquisadora observou que o registro semiespontâneo favorece a marcação de plural (86,1%), enquanto o registro espontâneo a desfavorece (51,5%), confirmando a hipótese de que a monitoração estilística influencia o comportamento dos indivíduos quanto às marcas explícitas de concordância nominal no português brasileiro. Na variável paralelismo formal, Rosa constatou que, no plano do paralelismo discursivo, ou seja, de um SN para outro, recebe mais marcas o SN quando precedido por SN com concordância padrão (91,4%), e recebe menos marcas quando precedido por um SN sem concordância padrão (33,3%), o que confirma a tendência de que a presença de marcas favorece marcas nos elementos do sintagma seguinte.

Para sistematizar os resultados dos estudos expostos aqui, a tabela abaixo apresenta uma distribuição dos percentuais totais de marcação de plural (a partir das análises

atomísticas, desenvolvidas em todos os estudos) em função da localidade controlada em cada estudo, o que permite refletir sobre o *continuum* rural-urbano no âmbito da concordância nominal (SCHERRE, 1988; BRANDÃO, 1994; 2013; ROSA, 2019].

**Tabela 1:** Sistematização da marcação explícita de plural no sintagma nominal (análise atomística) a partir da localidade dos falantes nas propostas de Brandão (1994; 2013), Scherre (1994) e Rosa (2019).

AMOSTRA	LOCALIDADE	ESCOLARIDADE	PORCENTAGEM
BRANDÃO (2013) ( <i>Corpus</i> Concordância – 2009-2011)	Copacabana - Rio de Janeiro	Ensino fundamental, médio e superior	92,4%
BRANDÃO (2013) ( <i>Corpus</i> Concordância – 2009-2011)	Nova Iguaçu - Rio de Janeiro	Ensino fundamental, médio e superior	91,1%
ROSA (2019) ( <i>Corpus</i> coletado em 2015)	Zona Oeste do Rio de Janeiro	Ensino fundamental, médio e superior	85,6%
SCHERRE (1988) (Amostra Censo do PEUL – década de 80)	Capital do Rio de Janeiro - várias zonas	Ensino fundamental e médio	72%
BRANDÃO (1994) ( <i>Corpus</i> APERJ – década de 80)	Norte e Noroeste fluminense	Analfabeto ou semi-escolarizados	44%

A comparação dos percentuais nos permite observar que o fenômeno da marcação de pluralidade, o que se associa a maior concordância nominal, se apresenta de forma mais variável nas áreas mais rurais e/ou afastadas dos centros do que nas urbanas e/ou mais próximas dos centros. De acordo sobretudo com o fator escolaridade, também é possível observar que o fenômeno se aproxima mais do perfil de uma regra semicategórica nas áreas mais urbanas, principalmente na fala de indivíduos que têm ensino superior.

Essas diferenças estão na base do debate empreendido em algumas pesquisas, também importantes para o estudo do fenômeno da concordância, que retratam a polarização norma culta *versus* norma popular na formação do PB. Esse tema se associa, em última análise, ao

debate sobre as origens do próprio PB, que apresenta divergências quanto à interpretação das motivações para seu caráter variável.

Naro; Scherre (2007), em estudos acerca das origens do português brasileiro, argumentam que as feições do português no Brasil são “resultado natural da deriva secular inerente na língua trazida de Portugal” (NARO; SCHERRE, 2007, p. 69), isto é, uma herança portuguesa de modo que a língua falada em Portugal e no Brasil partilham de estruturas variáveis semelhantes. Os autores rejeitam a ideia de que um processo de crioulização, ou seja, de fusão de diferentes estruturas a partir do contato direto de pessoas com línguas diferentes para o desenvolvimento do sistema linguístico do português brasileiro, possa ser bem fundamentada. Segundo os autores, as bases da variação foram herdadas do Português trazido de Portugal, tendo encontrado terreno fértil no Brasil, em meio ao intenso contato interlinguístico, para expandir essa variabilidade.

Em contrapartida, apresentando uma segunda concepção, Lucchesi (cf. 2003; 2009, entre outros) alega que o português brasileiro popular é fruto de um processo de aprendizagem linguística irregular e que essa transmissão se daria, principalmente, como efeito colateral da colonização europeia na África, América, Ásia e Oceania. Segundo Baxter; Lucchesi; Ribeiro (2009),

As variedades de segunda língua que se formam nessas condições, mais ou menos defectivas consoante as especificidades de cada contexto histórico, acabam por fornecer os modelos para aquisição da língua materna para as novas gerações de falantes, na medida em que os grupos dominados vão abandonando as suas línguas nativas. (BAXTER; LUCCHESI; RIBEIRO, 2009, p. 101)

De acordo com os autores, seria efeito colateral desse intenso contato interlinguístico, por séculos da história brasileira, a polarização entre as chamadas variedades cultas e populares. Observam essa polarização a partir das extremidades de um *continuum* que envolve o grau de urbanidade/ruralidade e a escolarização, conforme descrevem para o fenômeno da concordância no PB: em uma extremidade, a norma culta, presente nas áreas urbanas, com altos índices de marcação de plural; na outra, localizada nas comunidades mais rurais e, em muitos casos, afrobrasileiras, a norma popular, com baixos índices de marcação de plural.

Não é nosso objetivo, nesta pesquisa, adentrar no debate acerca das origens do PB, uma vez que exigiria elementos de aprofundamento e muita reflexão que não caberiam nos limites e propósitos do presente trabalho. Todavia, não se nega que relacionar ambas as propostas é relevante para o questionamento central da pesquisa, uma vez que ambas admitem o contato linguístico como gatilho para a aceleração do processo de diferenciação das variedades cultas e populares, consideradas em diversos trabalhos anteriores.

Assim, dialogando com pesquisas desenvolvidas a partir de dados produzidos sobretudo no Rio de Janeiro, o presente trabalho busca investigar como seria o comportamento de marcação de número nos sintagmas nominais na fala de moradores de Guapimirim / Cachoeiras de Macacu, questionando o que tal comportamento revelaria sobre o caráter de ruralidade dessas regiões nos dias atuais. Em síntese, a pesquisa busca contribuir para a expansão de resultados acerca de dados contemporâneos da variação no *continuum* rural-urbano considerando comunidades de fala em regiões rurais, normalmente caracterizadas por maiores índices de não-marcação, no que se refere à concordância nominal.

## 4. METODOLOGIA

Conforme já se apresentou, a presente pesquisa investiga a hipótese geral de que, mesmo com o vasto processo de urbanização e de escolarização no Brasil – o que fez mudar as feições das comunidades consideradas rurais –, ainda haveria diferença no comportamento da marcação de plural nos sintagmas nominais entre as áreas rurais e urbanas. Por esse motivo, foram analisados os sintagmas nominais com ao menos uma marca explícita de plural nas falas dos moradores dos municípios de Guapimirim e Cachoeiras de Macacu, caracterizados na subseção abaixo, que poderão ser brevemente comparadas, com base nos resultados das pesquisas anteriores, às de falantes da capital.

Além disso, descrevemos, neste capítulo, o *corpus* utilizado, o tipo de tratamento de dados desenvolvido, as variáveis linguísticas e extralinguísticas consideradas para a análise atomística, de modo a tornar claros os procedimentos realizados no tratamento e na análise dos dados coletados.

### 4.1 A comunidade de Guapimirim e Cachoeiras de Macacu

Situados na região metropolitana do estado do Rio de Janeiro, os municípios de Guapimirim e Cachoeiras de Macacu estão localizados a uma distância de, aproximadamente, 80 km e 104 km da capital, respectivamente. Em 2021<sup>4</sup>, Guapimirim atingiu um número de 62.225 habitantes, enquanto Cachoeiras de Macacu possui um total de 59.652 habitantes. Vizinhas, as referidas cidades são ricas em biodiversidade e de uma grande área natural, com cachoeiras e notáveis fauna e flora, além de grandes áreas de preservação e conservação ambiental.

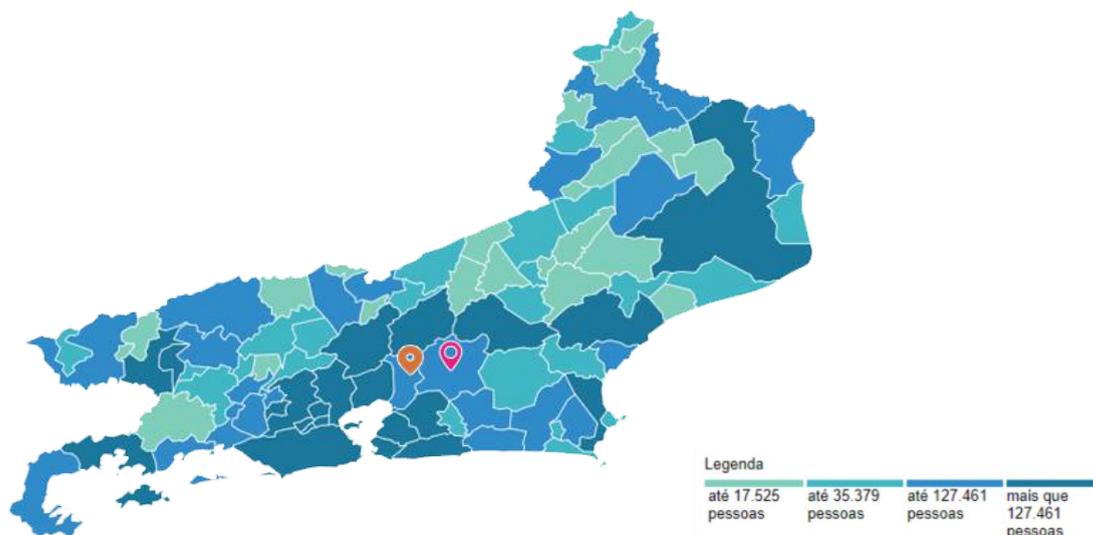
Em índices populacionais, um significativo número de pessoas reside no campo, área mais afastada do centro comercial e de difícil acesso. Enquanto Cachoeiras de Macacu vive economicamente do comércio (vendas), os habitantes relataram nas entrevistas que, em Guapimirim, a principal fonte de renda da população é no ramo de construção civil (obras). Alguns moradores das cidades apontaram, nas gravações realizadas, que a bicicleta é o meio de transporte mais comum na região e que a atividade de lazer se constitui, majoritariamente, dos banhos nas cachoeiras. Seguindo as comparações, enquanto Guapimirim conta com

---

<sup>4</sup> Dados extraídos do site do Instituto Brasileiro de Geografia e Estatística (IBGE) <ibge.gov.br>, consultados em 27/01/2022.

hospital municipal, os moradores de Cachoeiras de Macacu precisam se deslocar para outro município caso necessitem de atendimento; não há universidades em ambas as cidades, e, normalmente, os moradores recorrem aos outros municípios vizinhos em busca de melhores oportunidades de emprego.

**Figura 1:** Localização e população dos municípios de Guapimirim, em laranja, e Cachoeiras de Macacu, em rosa, no mapa do estado do Rio de Janeiro.



Fonte: <https://cidades.ibge.gov.br>

Assim, levando em consideração os dados apresentados, sobretudo consoante os critérios adotados pelo Instituto Brasileiro de Geografia e Estatística (IBGE), e a experiência da pesquisa de campo, Corrêa (2019) – responsável primeira pela constituição do *corpus*, ora em fase de ampliação – considera os municípios de Guapimirim e Cachoeiras de Macacu como áreas rurais, por contemplarem os traços e características demográficas, físicas e econômicas de ruralidade. Embora se suponha haver diferenças entre os municípios em termos de grau de ruralidade (sendo possivelmente Cachoeiras de Macacu ainda menos urbana do que Guapimirim), este trabalho abordou o conjunto de entrevistas, como faz Corrêa (2019), como pertencente a uma comunidade de fala menos urbana / mais rural, como um todo.

## 4.2 Descrição do *corpus*

Conforme já se informou, o *corpus* da pesquisa em questão é composto pela gravação da fala em entrevistas sociolinguísticas de moradores das cidades de Guapimirim e Cachoeiras de Macacu, do estado do Rio de Janeiro, organizado por Corrêa (2019), e complementado com gravações realizadas por Souza, em 2020-2021.

Os informantes foram separados e identificados por sexo (homem ou mulher), escolaridade (ensino fundamental ou ensino médio) e faixa etária (faixa A, faixa B e faixa C), como disposto no quadro abaixo:

**Quadro 1:** Descrição do *corpus*, separado por faixas etárias, níveis de escolarização e sexo dos falantes.

Faixa Etária	Escolaridade	Sexo
Faixa A (entre 15 e 35 anos)	Ensino Fundamental	Mulher
Faixa B (entre 36 e 55 anos)	Ensino Médio	Homem
Faixa C (a partir de 56 anos)		

Para cada cédula de escolaridade e de faixa etária, foram utilizados dois informantes de cada sexo, totalizando, assim: 4 informantes da faixa A com ensino fundamental, 4 informantes da faixa A com ensino médio; 4 informantes da faixa B com ensino fundamental; 4 informantes da faixa B com ensino médio; 4 informantes da faixa C com ensino fundamental, 4 informantes da faixa C com ensino médio, totalizando 24 informantes. Nas entrevistas, com duração aproximada de 20-25 minutos cada, foram feitas perguntas sobre a cidade em que se localizam, o sistema de saúde, trabalho, educação, entre outros temas, de forma que o participante não estivesse ciente de que o objetivo era investigar sua fala e, assim, viesse a se expressar de forma artificial, pouco espontânea.

## 4.3 Tratamento dos dados e a regra variável

Para a análise geral, investigamos a alternância da marcação de pluralidade como variável dependente, ou seja, optamos pelo detalhamento do comportamento dos dados de acordo com a chamada análise atomística. Nesse modelo, cada elemento flexionável do sintagma nominal é um item de análise, de forma a controlar quais fatores agem internamente

no sintagma. Assim, o controle primário dos dados não é exatamente o da concordância, mas quais vocábulos receberam marca de plural e quais as possíveis razões para que a marcação ocorra.

Desse modo, o trabalho considera dois fatores para definir a variável dependente: a *marcação de plural* (Ex. 1) ou *ausência de marcação de plural* (Ex. 2) nos vocábulos que compõem o sintagma nominal.

#### *Marcação de plural*

(1) Muitas pessoas usam o celular. (H, nível 1, faixa A)<sup>5</sup>

#### *Não marcação de plural*

(2) Geralmente essas pessoa<sup>0</sup> que mora em cidade... (H, nível 1, faixa A)

Na análise não atomística, o sintagma nominal inteiro é um item de análise, de forma a apresentar um resultado global da concordância de número nas sentenças. Apesar de não ter sido realizada a análise não atomística, foi feita também a contabilização dos dados gerais consoante a presença total de marcas ou a presença parcial de marcas, para que nos fosse permitido observar, em índices gerais, se a concordância de número nos SNs nas áreas mais rurais se apresentaria como uma regra variável, como indicam estudos anteriores (cf. Brandão 1994). Para isso, foi definida a variável binária: *concordância padrão* (Ex. 3) ou *não concordância padrão* (Ex 4) no SN.

#### *Concordância padrão*

(3) **Grandes empresas** já quiseram se instalar aqui. (H, nível 2, faixa C)

#### *Concordância não padrão*

(4) **as pessoa**<sup>0</sup> são acostumada/ tinha **uns trezentos empregado**<sup>0</sup>. (M, nível 2, faixa A/  
H, nível 1, faixa B)

No que se refere à análise multivariada de natureza atomística, todos os dados de sintagmas nominais – com ou sem marcação de número – foram coletados, com base na

---

<sup>5</sup> Com o objetivo de identificar os falantes dos exemplos expostos, os códigos utilizados foram: H (homem) ou M (mulher), Nível 1 (ensino fundamental) ou Nível 2 (ensino médio) e faixa A (entre 15 e 35), faixa B (entre 36 e 55) ou faixa C (a partir de 56).

audição das gravações e nas transcrições realizadas, e codificados de acordo com fatores das variáveis linguísticas e extralinguísticas descritas na próxima seção. Após a coleta e codificação, os dados foram submetidos à análise com apoio do pacote de programas GoldVarb X, que permitiu a quantificação das ocorrências, oferecendo valores absolutos, percentuais e pesos relativos em relação aos dados de marcação e não marcação de plural consoante os fatores das variáveis controladas.

#### **4.4 Descrição das variáveis independentes investigadas**

A seguir, descrevemos os grupos de fatores extralinguísticos e linguísticos observados na análise variacionista dos dados.

##### **4.4.1 Variáveis extralinguísticas**

Para o controle das variáveis extralinguísticas, foram considerados os fatores *sexo*, *faixa etária* e *escolaridade* dos falantes, descritos anteriormente no item 4.1.

Com base no observado em estudos anteriores, formulamos algumas hipóteses gerais. No que se refere à escolarização, esperamos que essa variável se apresente como mais relevante no que concerne à concordância, promovendo a hipótese de que quanto maior o nível de instrução formal, maior será seu nível de concordância padrão. Concomitantemente, é esperado que a faixa etária também seja uma grande influenciadora na realização de marcação de plural, de modo que os indivíduos com maior faixa etária apresentariam menos marcas de número no SN. Por fim, supomos que o sexo dos falantes, que compõe a terceira variável extralinguística, não apresente relevância na marcação ou não marcação de plural, mas, ainda assim, o controle dessa variável é o que permite observar a possível diferença entre homens e mulheres, tendo em vista seus respectivos papéis na comunidade de fala em questão.

##### **4.4.2 Variáveis linguísticas**

Neste subitem, são apresentadas as variáveis linguísticas controladas na codificação dos dados de SN coletados. A definição das variáveis utilizadas neste trabalho foi feita com base nos resultados de estudos anteriores acerca da concordância nominal (cf. SCHERRE, 1988; BRANDÃO, 2013), privilegiando-se, dado o caráter preliminar desta análise, aquelas relativas à constituição do sintagma nominal.

Com base no caráter da análise adotada, a atomística, foram definidas três variáveis linguísticas específicas para a investigação, as quais permitem o controle de características relacionadas ao sintagma nominal: *a função do vocábulo no SN*, *a posição do vocábulo no SN* e *a relação entre os constituintes do SN*, como descritas abaixo.

a) *Função do vocábulo no SN*

O grupo de fatores dessa categoria nos permite identificar qual a função do elemento dentro do SN. Indiretamente, ele também é analisado na variável *Relação entre os constituintes*, descrita mais à frente. As funções dos vocábulos no SN foram organizados em:

- Determinante / quantificador (**as** flores rosas/ **algumas** pessoa)
- Nuclear (as **montanhas** da Serra)
- Modificador (as geladeiras **grandes**)

Assim, considerando que a estrutura mais típica de um sintagma nominal segue a ordem *determinante/quantificador + núcleo + modificador*, esperamos que os determinantes e quantificadores apresentem mais marcas de número, em comparação com as outras duas categorias, em função do seu caráter especificador. Raposo et al. (2013) classificam os determinantes e quantificadores como especificadores dos nomes. Sobre os determinantes, os autores os definem da seguinte forma:

Palavras que se combinam com nomes, adquirindo o grupo resultante uma dimensão referencial, ou seja, passando a designar entidades particulares do mundo real ou imaginário sobre qual incide o discurso [...]. Os determinantes incluem os artigos definido e indefinido, a forma *algum* (e variantes) e as formas demonstrativas *este*, *esse*, *aquele* e suas variantes. (RAPOSO et al., 2013, p. 329)

Já os quantificadores, também classificados como especificadores, são definidos como “palavras que se combinam com nomes, passando o grupo resultante a designar entidades do mundo em termos de quantidade [...] Têm, por vezes, uma natureza adjetival.” (RAPOSO et al., p. 329). Dessa forma, além de levar em consideração o caráter semântico, esperamos que os vocábulos *determinantes/quantificadores* apresentem mais marcas de número, sobretudo por se apresentarem, normalmente, na posição mais à esquerda do SN.

b) *Posição do vocábulo no SN*

A variável *posição do vocábulo no SN* controla como a marcação ou não marcação é influenciada pela localização em que cada elemento ocupa no sintagma. Assim, levanta-se a hipótese de que os elementos identificados na primeira posição do SN tendem a receber mais marcas, independentemente da função que este exerça na sentença. As posições foram definidas como:

- 1ª posição (**alguns** meninos da escola)
- 2ª posição (alguns **meninos** da escola)
- 3ª/4ª posição (as melhores **aulas**)

c) *Relação entre os constituintes do SN*

A última variável controlada na análise atomística foi a *relação entre os constituintes*, de modo que pudéssemos controlar, de forma integrada, a natureza do elemento, se nuclear ou não, e a posição, se antes ou depois do núcleo. Scherre (1988), ao analisar as *marcas precedentes*, variável que não foi controlada no presente trabalho, observou que a presença de uma marca expressa de número ou a ausência de marcação em um elemento do SN pode influenciar no aparecimento ou não de marcas nos elementos subsequentes. Para este trabalho, levantamos apenas a hipótese de que os elementos não-nucleares antepostos receberiam mais marcas de número, ao contrário dos demais. O elemento pós-nuclear, por suposição, teria menos marcas de número, em comparação com os outros elementos.

Os fatores controlados nestas variáveis foram os seguintes:

- Elemento nuclear (essas **meninas** bonitas)
- Elemento não nuclear anteposto (**essas** meninas bonitas)
- Elemento não nuclear posposto (essas meninas **bonitas**)

## 5. ANÁLISE DOS RESULTADOS

Neste capítulo, analisamos quantitativamente e qualitativamente os resultados da análise atomística – aquela que considera a expressão de plural em cada item constitutivo dos sintagmas nominais coletados –, com o objetivo de averiguar fatores condicionadores do fenômeno estudado, após as rodadas multivariadas realizadas no programa GoldVarb X.

Embora não tenha sido realizado o controle de variáveis independentes de acordo com o modelo não atomístico - aquele que considera a realização da concordância em todos os itens do SN como um todo *versus* a não realização integral da concordância no SN como um todo -, também controlamos, conforme esclarecemos na metodologia, a frequência absoluta e percentual na distribuição geral das ocorrências de acordo com esse modelo.

### 5.1 Distribuição geral dos dados

Na análise atomística, como explicado no capítulo anterior, considera-se cada vocábulo flexionável do sintagma um item de análise. Quantitativamente, chegamos a um total de 379 dados. A tabela abaixo sistematiza o resultado geral acerca da marcação *versus* não marcação de número.

**Tabela 2:** Resultados gerais de marcas de número em cada constituinte do SN (análise atomística).

	Nº absolutos	%
<b>Marcação de número</b>	281	74,1%
<b>Não marcação de número</b>	98	25,9%
<b>Total</b>	379	100%

Como previsto, a variante em que se realiza a marcação de número no elemento do sintagma nominal apresentou mais produtividade, com 281 dados, equivalente a 74,1%, enquanto a não marcação de número apresentou índices de 25,9%, em 98 dados, o que revela que o fenômeno estudado se apresenta como uma regra variável (LABOV, 2003). Porém, por se tratar de uma análise atomística, o índice de 74,1% de marcação não reflete totalmente a realidade da concordância, mas a frequência com que os elementos no SN são marcados.

Para visualizarmos melhor a tendência do *corpus* analisado quanto à realização da concordância no SN visto como um todo, apresentamos, a seguir, a distribuição geral das ocorrências segundo a análise não atomística. Assim, com o objetivo de contribuir para a melhor avaliação das tendências do *corpus* analisado, realizamos a contabilização consoante a presença total de marcas ou a presença parcial de marcas, assumindo o SN inteiro como unidade de análise.

Como aplicação de regra, consideramos a presença de marca de número em todos os elementos do sintagma como *concordância padrão*, e a ausência de pelo menos um marca em um dos termos que compõem o sintagma como *concordância não padrão*. Ao todo, obtivemos 180 ocorrências de sintagmas nominais com ao menos uma marcação de número. A tabela abaixo sintetiza os resultados obtidos.

**Tabela 3:** Resultados gerais de concordância padrão e concordância não padrão nos SNs (análise não atomística).

	Nº absolutos	%
<b>Concordância padrão</b>	96	53,3%
<b>Concordância não padrão</b>	84	46,7%
<b>Total</b>	180	100%

Os resultados da tabela evidenciam que o comportamento da concordância nominal nas áreas mapeadas é altamente variável, com cerca de metade dos dados de cada uma das formas alternantes: 53,3% de ocorrências de *concordância padrão*, como no exemplo (5), e 46,7% de dados com *concordância não padrão*, como no exemplo (6).

(5) na minha rua mesmo, **as pessoas** são unidas (H, nível 1, faixa A)

(6) eu tinha abandonado até **meus amigo0** pra ficar no celular (H, nível 1, faixa A)

Apesar de, nesta pesquisa, não controlarmos os fatores linguísticos que agem internamente na influência do SN global, podemos observar que, dentro da perspectiva não atomística, o fenômeno da concordância padrão de número no sintagma nominal apresenta baixos índices (53,3% de marcação), se comparados, por exemplo, aos verificados em

amostras urbanas contemporâneas, configurando-se, assim, como uma regra variável mais produtiva na área mapeada. Essa análise, ainda que preliminar, nos leva a questionar quais fatores agem a favor da concordância padrão, sobretudo tendo como ponto de partida os resultados aqui obtidos e que desejamos desenvolver, se possível, no âmbito de um futuro mestrado.

Para entendermos melhor o comportamento dos dados quanto à expressão de pluralidade, é necessário analisarmos as variáveis independentes controladas na análise atomística, de modo a identificar os fatores que favorecem ou desfavorecem a marcação de número.

## **5.2. Variáveis estatisticamente relevantes**

Após uma rodada geral (com todas as variáveis controladas) para obtenção dos resultados, realizamos rodadas multivariadas (com e sem *Relação entre os constituintes*, para verificar algum efeito de superposição entre fatores) de modo a identificarmos as variáveis mais relevantes para fenômeno de marcação de número nos elementos dos sintagmas. Considerando a melhor significância do nível de seleção, o programa estatístico apontou, nessa ordem, a *relação entre os constituintes*, *escolaridade*, *sexo do participante*, e *função do vocábulo* como os grupos de fatores mais relevantes para o condicionamento da marcação de plural (valor de aplicação escolhido).

De forma a orientar e facilitar a apresentação dos resultados, estes serão divididos em duas subseções: a primeira, referente às variáveis extralinguísticas; e a segunda, às linguísticas.

### **5.2.1. Variáveis extralinguísticas**

O controle das variáveis extralinguísticas nos permitiu analisar quais grupos sociais realizam mais marcas ou menos marcas de número nos elementos do SN, o que nos possibilita averiguar como é o comportamento da marcação de plural nas comunidades consideradas mais rurais.

Dentre as variáveis extralinguísticas selecionadas, a variável *escolaridade*, de natureza binária, foi indicada como a mais relevante para o comportamento dos dados. Os informantes foram categorizados pelo nível de instrução, sendo ensino fundamental ou ensino médio. Os resultados gerais foram distribuídos na tabela a seguir.

**Tabela 4:** Distribuição dos dados de (não-)marcação de número em cada elemento do SN (análise atomística) segundo a variável extralinguística *Escolaridade*

	Marcação de número	Não marcação de número
<b>Ensino Fundamental</b>	195/281 (69,1%)	87/281 (30,9%)
<b>Ensino Médio</b>	86/97 (88,7%)	11/97 (11,3%)

Assim como na hipótese levantada, os informantes de ensino fundamental implementaram menos marcas do que os falantes de ensino médio, com 69,1% e 88,7% de marcação, respectivamente, o que nos confirma que o nível de escolarização influencia a aplicação das marcas de número nos elementos que compõem o sintagma.

Os resultados do nível de *escolaridade* mostraram que os falantes com ensino médio tendem a apresentar [+ marcas], com peso relativo de .81, enquanto os falantes com apenas o ensino fundamental tendem a apresentar [- marcas], com peso relativo de .37.

**Figura 1:** *Continuum* de marcação de plural no SN de acordo com a escolaridade do participante

[+marcas]	[-marcas]
Ensino médio	Ensino fundamental
.81	.37

A variável *sexo*, a princípio, foi controlada apenas devido à estratificação da amostra. Lucchesi; Baxter; Silva (2009), quando investigam a variável na fala de moradores de comunidades rurais, afirmam que os homens costumam realizar mais marcas por consequência do contexto social em que estão inseridos. Segundo o autor:

Essa maior integração do homem no processo produtivo e na vida social o coloca numa situação de maior exposição aos padrões linguísticos adventícios e numa posição que favorece mais a sua assimilação. Já as mulheres, envolvidas com a criação dos filhos e com as tarefas domésticas, ficam mais restritas à vida isolada da pequena propriedade, não obstante também trabalharem na lavoura e também

levarem os produtos para a feira. Circunscritas ao universo doméstico, acabam por conservar mais os primitivos usos linguísticos, mais diretamente marcados pelo processo de transmissão linguística irregular. (LUCCHESI, 2009, p. 313)

Assim, corroborando a hipótese levantada pelo autor, em nossa análise, os homens realizaram mais a marcação de número do que as mulheres, com índices de 78% e 67,4% de marcação, respectivamente.

**Tabela 5:** Distribuição dos dados de (não-)marcação de número em cada elemento do SN (análise atomística) segundo a variável extralinguística *Sexo do participante*

	Marcação de número	Não marcação de número
Homem	188/241 (78%)	53/241 (22%)
Mulher	93/138 (67,4%)	45/138 (32,6%)

Na variável *sexo*, os homens apresentaram maior tendência de marcação do que as mulheres, conforme indicam os pesos relativos de .58 e .35, respectivamente.

**Figura 2:** *Continuum* de marcação de plural no SN de acordo com o sexo do participante<sup>6</sup>

[+marcas]	[-marcas]
Homem	Mulher
.58	.35

Apesar de não ter sido selecionada como estatisticamente relevante, entendemos que expor os resultados da variável *faixa etária* auxilia no conhecimento do comportamento da comunidade de fala na análise do fenômeno observado. A tabela a seguir apresenta os resultados gerais referentes à variável.

<sup>6</sup> Embora a variável *sexo*, de ordem biológica, não tenha sido configurada efetivamente a partir da concepção de gênero, o que não seria compatível, portanto, com a proposição de um *continuum*, manteve-se essa forma de representação para manter a uniformidade na descrição do conjunto das variáveis e, ainda, para remeter à indagação se a maior ou menor proximidade com a relação de feminilidade e masculinidade afetaria o fenômeno.

**Tabela 6:** Distribuição dos dados de (não-)marcação de número em cada elemento do SN (análise atomística) segundo a variável *Faixa etária do participante*

	Marcação de número	Não marcação de número
<b>Faixa A</b>	128/168 (76,2%)	40/168 (23,8%)
<b>Faixa B</b>	56/79 (70,9%)	23/79 (29,1%)
<b>Faixa C</b>	97/132 (73,5%)	35/132 (26,5%)

De fato, os índices de realização de marcas de plural no SN foram muito próximos em cada faixa etária. Os informantes da Faixa A realizaram um pouco mais as marcas de plural (76,2%), se comparados aos falantes da Faixa B (70,9%), índices próximos dos falantes da Faixa C, que apresentaram 73,5% de marcação.

### 5.2.2 Variáveis linguísticas

Nesta subseção, estão expostos os resultados das variáveis linguísticas que se apresentaram relevantes na análise atomística: a *relação entre os constituintes e a função do vocábulo no SN*.

Em relação à variável *relação entre os constituintes*, Scherre (1996), ao apresentar uma abordagem de análise dos SNs, propõe que os elementos sejam analisados por meio de suas distribuições ao redor do núcleo, sendo anteposto ou posposto a ele. Partindo dessa metodologia, a pesquisadora observou que os elementos antepostos ao núcleo recebem mais marcas do que os elementos pospostos ao núcleo, indicando que não são os modificadores em geral que recebem [-marcas], mas os modificadores pospostos.

Na tabela a seguir, apresenta-se a distribuição dos dados na fala mais rural da amostra investigada.

**Tabela 7:** Distribuição dos dados de (não-)marcação de número em cada elemento do SN (análise atomística) segundo a variável *Relação entre os constituintes do SN*

	Marcação de número	Não marcação de número
<b>Elemento anteposto ao núcleo</b>	182/183 (99,5%)	1/183 (0,5%)
<b>Elemento nuclear</b>	96/182 (52%)	86/182 (48%)
<b>Elemento posposto ao núcleo</b>	6/17 (35,3%)	11/17 (64,7%)

Os resultados confirmam a hipótese apresentada por Scherre: os elementos antepostos ao núcleo receberam mais marcas de número (99,5%), apresentando-se como um contexto de aplicação semicategórica da marcação de plural. Os elementos nucleares e os elementos pospostos ao núcleo receberam 52% e 35,3% de marcação de plural, respectivamente.

O único caso de elemento anteposto ao núcleo não marcado é o apresentado no exemplo (7), onde o vocábulo que não recebeu marca é um modificador, e não uma categoria funcional, como um determinante, por exemplo, confirmando a relevância da variável *Função no vocábulo*. Além disso, observa-se que o vocábulo é um monossílabo tônico terminado em vogal oral, cuja oposição com a forma plural correspondente seria apenas o acréscimo do morfema -s. Ao que parece, o baixo grau da chamada *saliência fônica*, variável não investigada no presente trabalho, pode ter alguma relevância para a marcação de número na fala dos moradores das áreas mapeadas, o que precisa ser aferido na continuidade da pesquisa.

(7) são **má0** companhias (H, nível 1, faixa A)

A respeito da *relação entre os constituintes*, os elementos antepostos ao núcleo apresentaram peso relativo de .94, favorecendo a presença de [+marcas], sendo seguido dos elementos nucleares (peso relativo de .07) e dos elementos pospostos ao núcleo (peso relativo de .02) desfavorecendo a marcação.

**Figura 3:** *Continuum* de marcação de plural no SN de acordo com a relação entre os constituintes

[+marcas]		[-marcas]
Anteposto ao núcleo	Nuclear	Posposto ao núcleo
.94	.07	.02

Para a análise dos dados na variável *função do vocábulo*, definimos a categoria do determinante/quantificador como a que envolve as classes gramaticais que ocorrem em SNs em uma dimensão referencial (cf. Raposo et. al., 2013), combinando com os nomes, classificados como núcleo. A categoria do modificador, por sua vez, envolve os itens considerados adjetivos ou identificadores, acompanhando o núcleo do SN. Essa classificação detalhada tem por objetivo auxiliar na verificação da *função do vocábulo*, diferenciando-a da *relação entre os constituintes* no Sintagma Nominal, discutida anteriormente.

**Tabela 8:** Distribuição dos dados de (não-)marcação de número em cada elemento do SN (análise atomística) segundo a variável *Função do vocábulo*

	Marcação de número	Não marcação de número
<b>Determinante/ quantificador</b>	171/171 (100%)	0/171
<b>Núcleo</b>	96/182 (52,7%)	86/182 (47,3%)
<b>Modificador</b>	14/26 (53,8%)	12/26 (46,2%)

Em consonância com a hipótese levantada para os resultados relacionados à *função dos vocábulos*, os elementos categorizados como determinante/quantificador apresentaram, nos sintagmas coletados, marcas explícitas de número em todas as suas ocorrências, apresentando-se, assim, como uma regra categórica. Esse resultado nos leva a refletir sobre a configuração sintagmática da sentença, uma vez que a marcação se apresenta em menor ocorrência quando na função do núcleo e nos modificadores, com apenas 52,7% e 53,8% de

marcação, respectivamente. Como exposto nos exemplos (8) e (9) a seguir dispostos, independentemente da posição na sentença, os determinantes/quantificadores apresentam [+marcas] de número. Tais resultados nos confirmam a hipótese de que é sobretudo no núcleo, e a partir dele, que ocorre a não marcação.

(8) **Os** pais participam mais, né? (M, nível 1, Faixa C)

(9) **Uns trezentos** empregadoo. (H, nível 1, Faixa B)

Os elementos categorizados como determinantes/quantificadores se apresentaram como uma regra categórica, onde todos apresentaram [+marcas]. O núcleo apresentou peso relativo de .62, favorável à presença de [+marcas], em contraste com o modificador, desfavorecendo a marcação com peso relativo de .03.

**Figura 4:** *Continuum* de marcação de plural no SN no SN de acordo com a função do vocábulo

[+marcas]		[-marcas]
Determinante/quantificador	Núcleo	Modificador
1.0	.62	.03

Os resultados obtidos para as variáveis anteriores contemplam, indiretamente, a relevância da *posição do vocábulo no SN*, que, apesar de não ter sido selecionada como estatisticamente relevante quando a rodada contou com a totalidade das variáveis (devido à superposição dos fatores), confirma a hipótese de que a primeira posição é a que favorece o maior número de marcas.

Considerada uma das variáveis mais importantes em estudos anteriores do fenômeno, a *posição linear* é caracterizada em função do local que o elemento ocupa no SN. Scherre (1988) indicou que a primeira posição do SN é a mais marcada e as demais posições apresentariam índices menores, sendo possível, em alguns casos, estabelecer oposição entre o primeiro e último elemento do sintagma. Essa condição pôde ser confirmada nos dados percentuais (expostos na tabela a seguir), onde os elementos em 1ª posição receberam 98,3%

de marcação, seguido da 2ª posição e 3ª ou 4ª posição, com 53,1% e 45,8% de marcação, respectivamente.

**Tabela 9:** Distribuição dos dados de (não-)marcação de número em cada elemento do SN (análise atomística) segundo a variável *Posição linear no SN*

	Marcação de número	Não marcação de número
<b>1ª posição linear no SN</b>	177/180 (98,3%)	3/180 (1,7%)
<b>2ª posição linear no SN</b>	93/175 (53,1%)	82/175 (47,3%)
<b>3ª ou 4ª posição linear no SN</b>	11/24 (45,8%)	13/24 (54,2%)

Como se pode observar, é a partir da 2ª posição, sobretudo, que ocorrem os cancelamentos de marcas de número nos elementos do SN, apresentando uma perfeita linha decrescente no referente à marcação. Destacamos abaixo uma ocorrência em cada uma das posições controladas.

(10) **uns** lugar bom pra tomar banho. (M, nível 1, faixa B)

(11) tinha que ser mais puxado os **professor**, entendeu? (H, nível 1, faixa A)

(12) uns lugar **bom** pra tomar banho. (M, nível 1, faixa B)

Diante dos resultados, podemos corroborar a hipótese proposta por Scherre (1988), ao afirmar que “a existência de uma marca na primeira oportunidade não é, pois, condição suficiente para que as demais marcas redundantes desapareçam, embora, na maioria dos casos, seja condição necessária” (SCHERRE, 1988, p. 293).

### **5.3 *Continuum* rural-urbano da concordância nominal: breve comparação entre resultados de pesquisas no Rio de Janeiro**

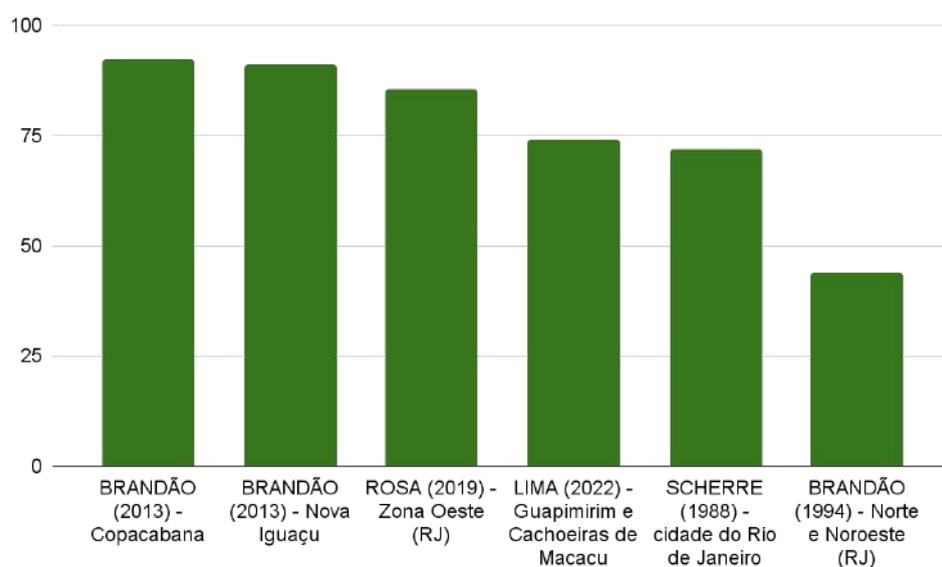
A análise dos resultados aqui apresentados, sobretudo os da análise atomística, foco desta investigação, nos permite situá-los brevemente no *continuum* rural-urbano da fala dos brasileiros, ao menos considerando a realidade fluminense.

O índice geral de 74,1% de marcação de número nos elementos constitutivos do SN permite considerar o fenômeno na fala de indivíduos de Guapimirim / Cachoeiras de Macacu, com ensino fundamental ou médio, como compatível com uma regra variável. Conforme detalhamos na análise, a não marcação de número plural incide de forma mais evidente na fala das mulheres e dos menos escolarizados; em relação ao condicionamento linguístico, observamos que a variação acontece efetivamente nas posições mais costumeiramente ocupadas pelos núcleos e pelos modificadores, e não na da categoria funcional dos determinantes ou quantificadores.

De modo geral, fica claro que os dados constituem expressão de uma norma popular de concordância, que se faz presente na fala dos moradores das áreas, o que ficou evidente no controle não atomístico dos dados, segundo o qual a concordância padrão acontece em cerca de metade das ocorrências.

Em termos comparativos, como demonstrado no gráfico abaixo, buscamos observar os resultados aqui apresentados em relação aos obtidos em pesquisas realizadas anteriormente (cf. Capítulo 3), considerando, aqui, apenas o índice geral de marcação de pluralidade nos elementos do SN em análises atomísticas.

**Gráfico 1:** Continuum rural-urbano na fala fluminense no âmbito da marcação de plural no SN segundo análises atomísticas [Scherre (1988); Brandão (1994; 2013); Rosa (2019)].



No gráfico acima, observamos que os resultados do *corpus* analisado por Brandão (2013) foram os que apresentaram maiores índices de marcação, sendo 92,4% para Copacabana e 91,1% para Nova Iguaçu, seguidos dos apresentados por Rosa (2019), com 85,6% na análise da fala de indivíduos da Zona Oeste do Rio de Janeiro. É importante ressaltar que os três *corpora* ora destacados, assim como os de Scherre (1988), se referem a dados contemporâneos produzidos por indivíduos residentes em áreas progressivamente urbanas. O índice mais baixo de não marcação de plural em Scherre (1988) pode estar relacionado a dois fatores: primeiro, à escolaridade dos indivíduos, que, diferentemente dos demais trabalhos, engloba falantes do nível fundamental; segundo, ao fato de a amostra, nesse caso, ser da década de 80, o que é compatível com o fato de que os índices de pluralidade no Brasil vêm aumentando progressivamente, sobretudo após o amplo processo de urbanização e acesso à escolarização.

De outro lado, os outros dois trabalhos são oriundos de áreas consideradas mais rurais: Guapimirim e Cachoeiras de Macacu, apresentados nesta pesquisa (LIMA, 2022), com 74,1%, próximos de Scherre (1988), na amostra Censo do *Corpus* PEUL, com 72% de marcação, e o de Brandão (1994), na região Norte e Noroeste do Rio de Janeiro, com apenas 44%. A grande diferença entre os índices das comunidades mais rurais também se relaciona aos dois fatores já mencionados: além de haver uma diferença na época da constituição dos *corpora* (enquanto Lima analisa dados contemporâneos, Brandão analisa dados da segunda metade do século XX - em que a natureza da ruralidade pode ser menor e diferenciada nos dias de hoje, com maior influência da cultura urbana), o estudo de Lima considera a fala de indivíduos com ensino fundamental ou médio e o de Brandão considera a fala de indivíduos com nenhuma ou parca escolaridade.

Percebe-se, assim, que apesar da concordância nominal se apresentar como uma regra variável e haver outros fatores influenciando o fenômeno, a tendência é que a aplicação de [+marcas] de número nos elementos dos sintagmas nominais, embora pareça crescer progressivamente entre os falantes rurais e urbanos, é menor nos dados produzidos em localidades mais rurais.

## 6. CONSIDERAÇÕES FINAIS

Ao longo do exposto nesta dissertação, pudemos analisar em que grau se apresentam os índices de marcação de número no sintagma nominal nas amostras das variedades rurais do estado do Rio de Janeiro, a partir das análises atomísticas dos dados coletados. A pesquisa foi feita com base na análise de gravações sociolinguísticas de residentes das cidades de Guapimirim e Cachoeiras de Macacu, cidades pertencentes à região metropolitana do Rio de Janeiro. A partir da coleta dos sintagmas nominais dessas gravações, acredita-se que as amostras dos 24 informantes, estratificados quanto à escolaridade, faixa etária e sexo, tenham nos apresentado um panorama do comportamento linguístico dos falantes da comunidade escolhida.

Focalizamos, majoritariamente, os resultados da análise atomística, isto é, em como cada um dos elementos que compõem o sintagma nominal se comporta e em como esse comportamento influencia a marcação de número na fala desses moradores. Verificamos, assim, que a marcação expressa de número nos SNs apresenta alta produtividade e se apresenta como uma regra variável (74,1% de marcação) nos municípios de Guapimirim e Cachoeiras de Macacu. No controle geral da análise não atomística, segundo a qual analisamos o sintagma nominal globalmente, a fim de termos uma análise quantitativa da concordância nos dados coletados, conferimos que, dos 180 SNs coletados, 53,3% apresentaram concordância padrão.

Em termos estruturais, o estudo demonstrou o forte condicionamento da relação entre os constituintes do SN e da função de cada constituinte, confirmando resultados de pesquisas anteriores, que demonstram que as posições mais à esquerda do sintagma, ocupada sobretudo por determinantes e quantificadores, se mostraram mais favoráveis à incorporação de marcas de plural no SN. Para o controle mais detalhado dos fatores linguísticos, propõem-se, em etapas futuras da pesquisa, a reconfiguração e a ampliação do controle das variáveis independentes, a começar pela *saliência fônica* (Cf. SCHERRE; NARO, 1998) e pela variável compósita *Posição e função dos constituintes no SN* (BRANDÃO, 2013).

No que concerne à caracterização sociolinguística do *continuum* rural-urbano, as cidades de Guapimirim e Cachoeiras de Macacu ainda demonstram tendências à não-marcação de número, fator característico das áreas mais rurais. Além disso, as variáveis extralinguísticas se mostraram relevantes, sobretudo a relacionada ao *sexo do participante*,

segundo a qual os homens aplicaram [+ marcas] do que as mulheres, e *escolaridade*, com os indivíduos com nível fundamental apresentando menores índices de marcação do que os de nível médio.

Esses resultados preliminares, por ora, podem ser sistematizados e problematizados da seguinte forma:

(i) Apesar da tendência à maior aplicação de [+ marcas] nas últimas décadas, as áreas rurais, ainda atualmente, confirmam a hipótese da presença de [- marcas] de número nos elementos do SN em comparação com as áreas urbanas do estado do Rio de Janeiro controladas, o que pode estar altamente relacionado ao perfil da escolarização dos indivíduos em cada localidade e ao acesso a meios/redes de comunicação social, o que precisa ser aprofundado; e

(ii) Considerando a estratificação dos falantes, os homens apresentaram índices maiores de realização de marcas do que as mulheres. Embora as amostras tenham sido coletadas consoante o mesmo parâmetro numérico, é necessário aprofundar a análise levando em conta o perfil social desses falantes em termos qualitativos, observando quais espaços as mulheres ocupam nessas regiões.

De um modo geral, a análise preliminar do método não atomístico nos permitiu fazer considerações sobre a concordância nominal nas áreas rurais, de forma a contribuir, também, para o registro de dados acerca da constituição do Português Brasileiro e suas variedades, uma vez que pudemos observar que o comportamento nas regiões de Guapimirim e Cachoeiras de Macacu se apresenta, em alguma medida, desfavorável ao uso de marcas, sobretudo quando comparado ao das regiões urbanas contempladas em outros estudos com amostras contemporâneas.

Assim, espera-se que a continuidade da investigação, possivelmente no âmbito do Mestrado, permita dar continuidade à pesquisa de modo a providenciar, além da ampliação de variáveis estruturais, a análise da concordância nominal por meio da abordagem não atomística dos dados. Ademais, espera-se contar com a ampliação das gravações coletadas com moradores das zonas ainda mais afastadas do centro das cidades para que, assim, possamos aprofundar a interpretação sobre as comunidades de fala do Rio de Janeiro no *continuum* rural-urbano, delimitando o que é característico de cada área e suas possíveis motivações.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

BAXTER, A. A concordância de número. In: LUCCHESI, D; BAXTER, A; RIBEIRO, I (Orgs.) *O Português afro-brasileiro*. Salvador: EDUFBA, 2009.

BRANDÃO, S. F. Em torno de um velho tema: o cancelamento da marca de número na fala das comunidades rurais brasileiras. *Revista Internacional de Língua Portuguesa* 12, p. 50-57, dez. 1994.

BRANDÃO, S. F.; VIEIRA, S. R. Concordância nominal e verbal: contribuições para o debate sobre o estatuto da variação em três variedades urbanas do português. *Alfa*, v. 56, n. 3, p. 1035-1064, 2012.

BRANDÃO, S. F. Concordância nominal. In: VIEIRA, S. R.; BRANDÃO, S. F. (Org). *Ensino de gramática: descrição e uso*. 2. ed. São Paulo: Contexto, 2013. [2007]. p. 55-83.

BRANDÃO, S. F. Concordância nominal em três variedades do Português: resultados gerais, novas indagações. *Cuadernos de la ALFAL*, v. 7, p. 36-52, 2015.

BRANDÃO, S. F.. Variação e estatuto de variedades do Português. *Diadorim* (Rio de Janeiro), v. Especial, p. 83-104, 2016.

CORRÊA, C. M. M. L. *Concordância verbal de terceira pessoa do plural em comunidades rurais e urbanas do estado do Rio de Janeiro: avaliação e produção*. 2019. Tese (Doutorado em Letras Vernáculas) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2019. 183 f.

LABOV, W. Some sociolinguistic principles. In: PAULSTON, C.; TUCKER, R. *Sociolinguistic. The essential readings*. New York: Blackwell, 2003. p. 234-250.

LABOV, W. *Padrões sociolinguísticos*. Tradução de Marcus Bagno, Maria Marta Pereira Scherre e Caroline R. Cardoso. São Paulo: Parábola, 2008 [1972].

LUCCHESI, D. O conceito de ‘transmissão linguística irregular’ e o processo de formação do português do Brasil. In: RONCARATI, C.; ABRAÇADO, J. (Orgs.). *Português brasileiro: contato lingüístico, heterogeneidade e história*. Rio de Janeiro: 7Letras, 2003. p. 272-284.

LUCCHESI, D., and BAXTER, A. A transmissão linguística irregular. In: LUCCHESI, D., BAXTER, A.; RIBEIRO, I. (Orgs.). *O português afro-brasileiro*. Salvador: EDUFBA, 2009. p. 101-024.

LUCCHESI, D. *Língua e sociedade partidas: a polarização sociolinguística do Brasil*. São Paulo: Contexto, 2015.

NARO, A. J.; SCHERRE, M. M. P. *Origens do português brasileiro*. São Paulo: Parábola, 2007.

RAPOSO, E. B. P.; BACELAR DO NASCIMENTO, M. F.; MOTA, M. A. C. da; SEGURA, L.; MENDES, A. (Coord.). *Gramática do Português*. Lisboa: Fundação Calouste Gulbenkian, 2013, Vol. I.

ROSA, F. F. S. *Concordância nominal interna ao SN e monitoração estilística*. Monografia (Conclusão de Curso) - Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2016.

ROSA, F. F. S. *A influência da monitoração estilística na concordância nominal de número interna ao SN*. 2019. Dissertação (Mestrado em Letras Vernáculas) - Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2019. 137 f.

SCHERRE, M. M. P.. *Reanálise da Concordância Nominal em Português*. 1988. Tese (Doutorado em Linguística) - Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1988.

SCHERRE, M. M. P. 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. 12, p. 37-49. dez. de 1994.

VIEIRA, S. R.; BRANDÃO, S. F. Tipologia de regras linguísticas e estatuto das variedades línguas: a concordância em português. *Linguística*, Montevideo, v. 30, n. 2, p. 81-112, dez. 2014.

VIEIRA, S. R.; BRANDÃO, S. F. CORPORAPORT: Variedades do Português em análise. Rio de Janeiro, Faculdade de Letras/UFRJ. Disponível em: <[www.corporaport.lettras.ufrj.br](http://www.corporaport.lettras.ufrj.br)>

VIEIRA, S. R. (Org.). *A concordância verbal em variedades do português: a interface fonética-morfossinaxe*. Rio de Janeiro: Vermelho Marinho, 2015.

WEINREICH, U.; LABOV, W. & HERZOG, M. *Fundamentos empíricos para uma teoria da mudança linguística*. (Tradução de Marcos Bagno). São Paulo: Parábola, 2006 [1968].