

# A concordância nominal de número em Bacabal-MA: um estudo sociolinguístico

**João Vitor Cunha Lopes**

Graduado em Letras

Universidade Federal do Maranhão

E-mail: [joaovitorcunhalopes@outlook.com](mailto:joaovitorcunhalopes@outlook.com)

 ORCID: <https://orcid.org/0000-0001-6696-5364>

**Revista Falange Miúda**

ISSN 2525-5169

**Periodicidade:**

Chamada temática

**Volume 5**

**Número 2**

**Recebido em:** 04/03/2020

**Aprovado em:** 19/06/2020

## **Resumo**

Com base nos pressupostos da Sociolinguística Variacionista (WEINREICH; LABOV; HERZOG, 2006 [1968]; LABOV, 2008[1972]), esta pesquisa se propôs a investigar os fatores linguísticos e sociais que se correlacionam à realização de concordância nominal de número em sintagmas nominais simples no português falado em Bacabal, Maranhão. Para tanto, foram analisadas 12 entrevistas sociolinguísticas realizadas com falantes dessa cidade. Os resultados da análise de regressão logística, realizada no programa R (R CORE TEAM, 2019), apontam para o fato de que a escolaridade e a faixa etária dos falantes são as variáveis sociais que se correlacionam ao fenômeno. Em relação às variáveis linguísticas, a análise evidenciou que todas as variáveis se correlacionam à realização da concordância.

## **Palavras-chave:**

Concordância nominal de número, Português bacabalense, Variação linguística.

## 1 Introdução

Este estudo é uma contribuição para a descrição do português falado em Bacabal. Nele<sup>1</sup>, analisa-se a realização da concordância nominal de número em sintagmas nominais simples (constituídos de dois elementos), como em “os meninos” e “os meninoØ”, na fala de 12 informantes da zona urbana dessa cidade.

No Maranhão, há muito se vem descrevendo o português maranhense (cf. RAMOS, BEZERRA, ROCHA, 2005; 2006; 2010; TEIXEIRA, 2017), no entanto, há poucos estudos que se ocupam em analisar a fala de Bacabal e da região Mearim. Apesar de o fenômeno da concordância nominal já ter sido amplamente descrito no português brasileiro (cf. BRAGA, 1977; SCHERRE, 1988, 1996; CARVALHO, 1997; SCHERRE E NARO, 1998; LOPES, 2001; SALOMÃO, 2010; RAMOS E BEZERRA, 2010; FIAMENGUI, 2011; OUSHIRO, 2015; SILVA, 2014; TEIXEIRA, 2017; SCARDUA, 2018), ele ainda não foi amplamente descrito com base na variedade maranhense do português, de modo a se estabelecer os padrões de realização dessa variável no português falado no estado.

Por mais que o Atlas Linguístico do Maranhão – ALiMA (cf. RAMOS, 1999), projeto de pesquisa vinculado ao Departamento de Letras da Universidade Federal do Maranhão, tenha realizado e apresentado resultados de pesquisa acerca da variedade do português falado em Bacabal (RAMOS E BEZERRA, 2010; ALVES, 2010), esta cidade ainda não teve a sua realidade linguística descrita com base nos procedimentos teórico-metodológicos estritos da Sociolinguística Variacionista (LABOV, 2008[1972]).

Sabe-se que o português brasileiro é uma língua heterogênea, característica comum a todas as línguas naturais, que oferece várias possibilidades de uso para produzir uma mesma expressão dentro de um determinado contexto, constituindo, assim, o que Labov (2008 [1972]) denomina de variantes linguísticas. No caso do fenômeno de concordância nominal de número, observa-se que, além da marcação prescrita pela gramática normativa do português, que “exige” a marcação de plural em todos os elementos do sintagma nominal (cf. BECHARA, 2009; ROCHA LIMA, 2011), a não marcação redundante também pode ser realizada pelo falante, atestando a variação entre essas formas.

Com a proposta de descrever um fenômeno variável morfossintático que se percebe muito frequente no português brasileiro e, conseqüentemente, no português falado em Bacabal – a variação na concordância nominal de número –, buscou-se investigar a variação nessa variável, a fim de responder a duas perguntas: como se dá a realização da concordância nominal de número? Quais os padrões de usos desse fenômeno?

---

<sup>1</sup> Este artigo faz um recorte de um estudo mais amplo realizado na graduação. Aqui, são apresentados os resultados da análise **não atomística** de variáveis linguísticas e sociais.

Com base nisso, são objetivos deste estudo: i) investigar a realização de concordância nominal de número em sintagmas nominais simples na fala de bacabalenses; ii) verificar se o fenômeno apresentado se trata de um caso de variação estável ou se há um indício de mudança em curso; iii) analisar a relevância de fatores linguísticos já testados em outras pesquisas (formação do plural no elemento à direita, tonicidade do elemento nuclear, número de sílabas do elemento nuclear, classe morfológica do primeiro elemento pré-nuclear, animacidade do SN) e; iv) analisar a relevância de fatores sociais nesse caso de variação/mudança, a saber sexo/gênero, faixa etária e escolaridade dos falantes. A análise sociolinguística desse fenômeno contribuirá para que se faça uma descrição mais acurada da fala da cidade em foco e, conseqüentemente, da fala maranhense, ainda pouco descrita.

A seção, a seguir, trata, ainda que brevemente, sobre a sociolinguística variacionista. Logo em seguida, trata-se do tratamento dado ao estudo da concordância nominal de número em sintagmas nominais pela gramática normativa e por três estudos variacionistas que se organizam numa agenda de descrição desse fenômeno linguístico com base no português brasileiro.

Em seguida, apresenta-se a metodologia utilizada para realização da pesquisa. São apresentadas algumas informações relevantes sobre a cidade de Bacabal. Também são descritos os instrumentos utilizados, como o roteiro de entrevista. Apresentam-se qualitativamente as variáveis linguísticas e sociais consideradas para este estudo. Na seção de análise, são apresentados a distribuição geral dos dados e os resultados estatísticos obtidos por meio das análises das variáveis sociais no programa R (cf. R CORE TEAM, 2019). Este artigo se encerra com as considerações finais e com a apresentação das referências bibliográficas que subsidiaram teoricamente a realização desta pesquisa.

## 2 A sociolinguística variacionista

Em contrapartida a um pensamento homogêneo de língua e falante ideais (SAUSSURE, 1969 [1916]; CHOMSKY, 1978 [1965]; 2018 [1957]), a Sociolinguística tem interesse em uma análise que contemple o falante em situações reais de usos da língua, a fim de verificar quais aspectos sociais permeiam a fala do indivíduo (cf. LABOV 2008 [1972]). Labov (2008 [1972]: 21), em seu estudo sobre a realização dos ditongos /aj/ e /aw/, em palavras como *life* ‘vida’ e *house* ‘casa’, afirma que,

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

A Sociolinguística é uma área de estudo que, de uma forma mais sistemática, busca evidenciar quais aspectos linguísticos e sociais são relevantes para a explicação da variação e das mudanças que ocorreram, bem como as possíveis mudanças que possam ocorrer na língua (cf. MARTELOTTA, 2011). As pesquisas em Sociolinguística pensam a língua como “uma instituição social e, portanto, não pode ser estudada como uma estrutura autônoma, independente do contexto situacional, da cultura e da história das pessoas que a utilizam como meio de comunicação” (MARTELOTTA, 2011: 141). Assim, a inclusão da variação na língua, como foco dos estudos linguísticos, foi um dos diferenciais em relação ao Estruturalismo de Saussure e ao Gerativismo de Chomsky.

Ainda na década de 60, quatro anos depois da primeira conferência sobre sociolinguística<sup>2</sup>, os pesquisadores Weinreich, Labov e Herzog (2006 [1968]) estabeleceram os pressupostos teórico-metodológicos da área dos estudos linguísticos que se ocupa da variação e da mudança linguísticas. Os autores afirmam que a variação linguística, além de estar sujeita a pressões estruturais, também sofre pressões sociais. Eles sugerem um novo modelo de língua, no qual fossem acomodados “os fatos do uso variável e seus determinantes sociais e estilísticos” (ibidem: 34). Essa nova linha de estudo baseia-se no fato de que a variação é inerente a qualquer língua e pode ser analisada e sistematizada como uma estrutura heterogênea, a partir da fala em seu contexto social.

Esses autores foram os responsáveis pela implementação dos pressupostos teóricos da Sociolinguística Variacionista. William Labov, por meio de suas pesquisas, contribuiu de forma significativa para o desenvolvimento da metodologia da área, e desde os seus primeiros estudos de ordem fonética sobre o inglês falado na ilha de Martha’s Vineyard, em 1963, no estado de Massachusetts (Estados Unidos) e o inglês falado em Nova Iorque, em 1966, o seu intuito foi o de contrapor-se ao modelo gerativo, devido à ausência do componente social (cf. TARALLO, 2007) na análise linguística. Em 1972, a partir da publicação do livro Padrões Sociolinguísticos<sup>3</sup>, Labov trouxe algumas discussões acerca das duas de suas principais pesquisas, além de outros estudos, fazendo com que a Sociolinguística Laboviana/Variacionista recebesse uma maior robustez no modelo que subsidiasse a busca por padrões gerais de usos.

Labov defendeu que a língua deixasse de ser tratada como algo homogêneo<sup>4</sup> para, assim, ser tratada como uma realidade heterogênea, considerando a sistematização da variação, característica intrínseca a toda língua natural. Com

---

<sup>2</sup> Em maio de 1964, um grupo composto por 25 pesquisadores reuniram-se para discutir pesquisas cujo foco fosse mais social, em uma conferência realizada na Universidade da Califórnia, em Los Angeles (UCLA), promovida pelo pesquisador William Bright (cf. CAMACHO, 2013; CALVET, 2002).

<sup>3</sup> LABOV, William (1972). *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

<sup>4</sup> A heterogeneidade é tratada como sendo algo que pode ser sistematizado. Weinreich, Labov e Herzog (2006 [1968]: 125) afirmam que “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 dessas estruturas heterogêneas”.

relação ao termo Sociolinguística, o próprio Labov (2008 [1972]: 13) afirma ter resistido por vários anos a essa nomenclatura, uma vez que, segundo o autor, “implica que pode haver uma teoria ou prática linguística bem-sucedida que não é social”. Justamente por esse caráter, não seria “necessário” nomear uma vertente para definir os estudos linguísticos pelo olhar social. No entanto, o problema relacionado a isso repousa no fato de que tal afirmação não era evidenciada nos estudos da época. Inegavelmente, a partir dessas pesquisas, a observação da língua e da sociedade se tornaram algo considerado indissociável<sup>5</sup>.

A Sociolinguística Variacionista considera que toda língua é heterogênea, devido à presença da variação (formas variantes concorrentes), e estabelece que não há somente uma forma linguística para referir-se à um mesmo significado ou expressão, em um determinado contexto, estabelecendo, assim, o que Labov (2008 [1972]) denomina de variantes linguísticas. Ao conjunto de variantes linguísticas, denomina-se variável dependente, cujo emprego “não é aleatório, mas influenciado por grupos de fatores (ou variáveis independentes) de natureza social ou estrutural” (MOLLICA, 2015: 11). Ou seja, a existência das variantes está intimamente ligada a fatores ou variáveis independentes, que podem influenciar os usos.

É importante salientar que o termo “variável” pode se referir tanto ao fenômeno que se encontra em variação quanto aos grupos de fatores linguísticos e sociais. Para Labov (2008 [1972]: 93), uma variável “indica um foco de distribuições significativas dentro da unidade, condicionando aquilo que, de outro modo, seria considerado variação livre ou não condicionada”.

Em se tratando da variável linguística morfossintática analisada neste estudo, propõe-se que, tanto a variante ‘marcação redundante de plural’ (os meninos chegaram cedo), quanto a variante ‘não marcação redundante de plural’ (os meninoØ chegaram cedo) mantêm o mesmo valor de verdade da noção de número (vários meninos), pois o determinante “os” carrega a ideia de plural e, de certa forma, torna-se “desnecessária” a marcação no elemento seguinte, considerando-se os resultados de pesquisas que indicam que o elemento de primeira posição dentro do SN quase sempre terá a marcação de plural (SCHERE, 1988; OUSHIRO, 2015).

Este trabalho se embasa nos pressupostos teórico-metodológicos da Sociolinguística Variacionista (WEINREICH; LABOV; HERZOG, 2006 [1968]; LABOV, 2008 [1972]). Quanto ao estudo do processo de realização da concordância nominal de número no sintagma nominal, especificamente, foram consultadas algumas pesquisas como referencial metodológico (SCHERE, 1988; OUSHIRO, 2015; RAMOS E BEZERRA, 2010; TEIXEIRA, 2017). A abordagem dessas pesquisas, bem como a visão prescritivista da realização de concordância nominal, serão apresentadas na seção três, a seguir.

---

<sup>5</sup> Labov (2008 [1972]: 306) afirma que, para alguns linguistas que priorizavam apenas os fatores internos, “a influência da sociedade” interferia no “desenvolvimento” da língua, e que, por isso, poderia ser considerada “assistemática” e até ‘catastrófica’, como previsto pela visão de Martinet.

### 3 A concordância nominal de número

#### 3.1 Visão gramatical

De acordo com os conceitos previstos nas gramáticas normativas de Bechara (2009) e de Rocha Lima (2011), depreende-se que a concordância nominal de número é considerada redundante e obrigatória, pois as palavras dentro do sintagma nominal (doravante SN) devem se harmonizar, flexionando-se com as palavras de que são dependentes. A concordância nominal, segundo Bechara (2009: 543), “pode ser estabelecida de palavra para palavra ou de palavra para sentido”.

A concordância de palavra para palavra ocorre da seguinte forma:

- 1 – Quando a estrutura apresentar apenas uma palavra determinada, a palavra determinante deve concordar, em gênero e número com a palavra determinada (como em ele cantou aquela música linda).
- 2 – Quando a estrutura apresentar mais de uma palavra determinada, a palavra determinante deve concordar, em número e gênero comum da palavra determinada ou concordar com a palavra anteposta mais próxima (como em a língua e a literatura portuguesas/ a língua e a literatura portuguesa).
- 3 – Quando a estrutura apresentar apenas uma palavra determinada e mais de um determinante, a palavra determinada irá para o plural ou ficará no singular, sendo, neste último caso é facultativo o uso do artigo (a exemplo de as literaturas brasileira e portuguesa; a literatura brasileira e portuguesa; ou a literatura brasileira e a portuguesa).

Rocha Lima (2011: 376) propõe que a concordância do adjetivo com o substantivo aconteça conforme os seguintes preceitos gerais:

- a) caso o adjetivo modifique um só substantivo, este tomará o gênero e o número deste (como em homem alto, mulher alta, homens altos, mulheres altas);
- b) caso haja vários substantivos, de gêneros diferentes e do singular, o adjetivo pode ir para o masculino plural ou concordar apenas com o substantivo mais próximo.

O autor afirma ainda que, no caso acima, essa escolha está sujeita às exigências da eufonia e da clareza, e se subordina, principalmente, à intenção do autor (como em o pai e a mãe extremosos, ou extremosa), extremosos se o adjetivo se referir tanto ao pai quanto à mãe, ou extremosa se o adjetivo se referir apenas à mãe. Se o adjetivo preceder os substantivos, será feita a concordância com o primeiro destes (como em boa hora e local escolheste!).

As mesmas condições são seguidas quando os substantivos são de gêneros e números variados (a exemplo de agastamentos e ameaças fingidos ou fingidas prantos, lamentações e mágoas dolorosos ou dolorosas propósitos e tentativas malogradas).

Por conseguinte, vê-se que as gramáticas normativas, como as de Bechara (2009) e Rocha Lima (2011), consideram a escrita literária clássica e propõem a prescrição da realização de concordância nominal a partir da lista de uma gama de regras fixas que não refletem a língua em seu uso real.

### 3.2 Estudos variacionistas sobre concordância nominal de número

O fato de haver regras que regem a concordância nominal de número não é garantia de que a marcação redundante de plural aconteça em situações reais de fala. Estudos de cunho variacionista (cf. BRAGA, 1977; SCHERRE, 1988, 1996; SCHERRE E NARO, 1998; CARVALHO, 1997; LOPES, 2001; SALOMÃO, 2010; RAMOS E BEZERRA, 2010; FIAMENGUI, 2011; SILVA, 2014; OUSHIRO, 2015; TEIXEIRA, 2017; SCARDUA, 2018) têm demonstrado que a realização da concordância nominal em SNs é influenciada por variáveis sociais e linguísticas.

A ausência de concordância nominal geralmente está relacionada a níveis de escolaridade mais baixos e à informalidade da situação de fala, dado a se tratar de um fenômeno linguístico socialmente estigmatizado (cf. SCHERRE E NARO, 1998). Quanto às variáveis linguísticas analisadas por outros pesquisadores (cf. SCHERRE, 1988; SILVA, 2014; OUSHIRO, 2015; TEIXEIRA, 2017), as que têm apresentado maior correlação com esse fenômeno são a posição linear que o elemento ocupa no SN (posição de uma palavra dentro do SN: 1º posição – **as** casas amarelas; 2º posição – as **casas** amarelas; 3º posição – as casas **amarelas**, e assim por diante) e a classe gramatical de seus elementos (classes gramaticais das palavras do SN: artigo – **as** casas ; substantivo – as **crianças**; adjetivo – as casas **amarelas**).

A fim de delinear um brevíssimo panorama dos estudos sobre concordância nominal no português falado no Brasil, revisitam-se, aqui, alguns dos estudos variacionistas já realizados sobre o tema (SCHERRE, 1988; OUSHIRO, 2015; TEIXEIRA, 2017).

Em sua tese de doutorado, Scherre (1988) analisou elementos passíveis de flexão dos SNs (simples e complexos), a fim de descrever e explicar a correlação de variáveis linguísticas e sociais, quanto à realização da marcação de plural, e evidenciar a existência de um sistema que gerencia essa variação. Além disso, ela discutiu a hipótese funcionalista de Kiparsky (1972), a fim de evidenciar que uma teoria linguística não deve considerar apenas aspectos gramaticais. A autora também pretendeu verificar se havia variação estável ou se tratava-se de um processo de mudança em curso, e observar se o fenômeno em tela revelava um caso de variação inerente, ou seja, se a variação grupal reflete a variação individual. Para tanto, ela analisou uma amostra composta por 64 informantes da cidade do Rio de Janeiro, estratificada de acordo com a faixa etária (7-14 anos, 15-25 anos, 26-49 anos, 50-71 anos) dos falantes, sua escolarização (primário, ginásial e colegial) equivalentes ao ensino fundamental e ensino médio.

As variáveis linguísticas consideradas relevantes para a **análise atomística** (análise de cada elemento de forma isolada) foram marcas precedentes e posição, saliência fônica (dimensões processos e tonicidade), relação dos elementos não nucleares em função do núcleo e posição dos elementos nucleares no SN, formalidade dos substantivos e adjetivos, grau dos substantivos e adjetivos, animacidade dos substantivos, contexto fonético/fonológico seguinte e função resumitiva. Para a análise **não atomística** (análise de todo o SN) foram considerados a pluralidade do contexto, configuração sintagmática do SN, saliência fônica (dimensão e processos), grau/formalidade do SN, pluralidade do SN, função textual e localização do SN. As variáveis sociais consideradas foram escolaridade, idade e sexo. A autora postulou nesse estudo um princípio geral chamado “paralelismo formal”, considerando que, no SN, marca de plural leva à marca de plural e marca zero leva à marca zero.

Controlando esses fatores, Scherre observou que o fenômeno linguístico concordância nominal é variável, inerente e está internalizado nas mentes dos falantes, apresentando duas tendências claras em dois subagrupamentos distintos dos falantes: 1) variação estável com gradação etária nos falantes de ambiente não humilde e concordância alta; 2) processo de mudança linguística em direção a um sistema sem marcas nos falantes de ambiente humilde e concordância baixa.

Mais recentemente, Oushiro (2015) realizou uma pesquisa com dados da fala paulistana, composta por uma amostra de 118 entrevistas, com perfis estratificados de acordo com as faixas etárias de 20-34 anos, 35-59 anos e 60 anos ou mais, com escolarização de ensino médio e ensino superior e região de residência (bairros centrais ou mais periféricos, em São Paulo) e o sexo dos informantes. A fim de discutir identidades sociais e o possível impacto de significados das variantes nos processos de variação e mudança, a autora analisou quatro variáveis sociolinguísticas: a realização de /e/, a realização de /r/ em coda silábica, a concordância nominal de número e a concordância verbal de primeira e terceira pessoa do plural.

Para concordância nominal de número, a autora controlou sintagmas simples e complexos de forma **atomística** e **não atomística**, com os seguintes fatores: sexo/gênero, faixa etária, escolaridade, região de residência, classe social, origem dos pais, mobilidade geográfica, processos morfofonológicos de formação do plural, tonicidade do item singular, posição relativa, classe morfológica, contexto fônico seguinte e paralelismo intrassintagmático (somente para 3<sup>a</sup> posição – os meus dois, 4<sup>a</sup> posição – os meus dois amigos e a 5<sup>a</sup> posição – os meus dois amigos pobres). Oushiro observou que, no que concerne às variáveis sociais, é a classe social do falante, bem como o seu sexo/gênero e sua escolaridade, os fatores que mais se correlacionam ao fenômeno estudado. Quanto às variáveis linguísticas, destacam-se a posição linear e classe de palavras.

A autora observou que há um quadro de mudança na direção da variante marcação redundante de plural em regiões mais centrais e explica que esta mudança parece estar relacionada ao enfraquecimento da relação de marca zero com

identidades locais e a prevalência do discurso do “erro”. Para uma proposta de análise perceptual de variantes linguísticas (ECKERT, 2008; 2012), ela sugere que a variante não padrão esteja relacionada ao índice de masculinidade.

No contexto maranhense, Teixeira (2017) realizou uma pesquisa sobre a concordância nominal, a partir de uma amostra de fala composta por 20 entrevistas realizadas em uma comunidade do bairro Campo de Belém, situada no município de Caxias – MA. Esses informantes foram distribuídos de acordo com seu sexo (9 homens e 11 mulheres), dois níveis de escolarização (ensino médio e superior) e duas faixas etárias (dos 18 aos 34 e dos 35 aos 50 anos). Foram analisadas as variáveis linguísticas saliência fônica (tonicidade do núcleo), classe morfológica do item pré-nuclear, posição linear e posição relativa dos constituintes e as variáveis sociais sexo, faixa etária e escolaridade. As variáveis linguísticas posição linear e posição relativa dos constituintes foram as que mais se correlacionaram com a variante considerada não prestigiada.

É interessante evidenciar que, em relação às variáveis sociais sexo e faixa etária, a autora constatou que não há muita diferença entre homens e mulheres, e que os adultos são responsáveis pela maior parte das ocorrências não prestigiadas. A variável social mais expressiva é a escolaridade, pois a formação básica dos informantes aparece como o fator social que mais se correlaciona à realização da variante considerada não-padrão.

Essas pesquisas que se ocuparam em analisar a realização da concordância nominal de número no português brasileiro ajudaram, de um modo geral, na definição das variáveis independentes investigadas neste estudo. Assim, será possível verificar se as hipóteses selecionadas se confirmam também na fala dos bacabalenses, e se os padrões dessa variante se aproximam aos padrões observados aos já alcançados por outros pesquisadores.

#### **4 *Corpus* e métodos da pesquisa**

Para a presente pesquisa, foi feita uma análise qualitativa de dados, a fim de que fossem observados, a partir da amostra da fala bacabalense, em quais contextos ocorreram os casos de concordância nominal, bem como a sua variante não padrão. Essa etapa é importante porque, a partir daí, foi possível realizar a análise quantitativa desses mesmos dados, de modo que fossem obtidos os valores estatísticos das variáveis que se correlacionam aos fatores sociais analisados. É necessário que se reitere o fato de que toda investigação sociolinguística de cunho variacionista é realizada com base em uma fase inicial de “coleta de dados”, seguida de uma “redução e apresentação de dados”, culminando, mas não se encerrando, na “interpretação e explicação de dados” (GUY; ZILLES, 2007: 20).

De certo modo, as etapas que subsidiam a realização de uma pesquisa sociolinguística são apreciadas neste estudo, do seguinte modo: (i) construção da amostra de fala – gravação das entrevistas; (ii) transcrição das entrevistas no Elan (cf. HELLWIG; GEERTS, 2019); (iii) extração dos dados (SNs) em planilha do Excel; e (iv) análise dos dados no R (cf. R CORE TEAM, 2019).

O locus da pesquisa foi a cidade de Bacabal, que está localizada ao centro norte do estado, a 240 quilômetros da capital, São Luís. Bacabal é considerada uma das cidades mais importantes do estado, devido à sua localização estratégica que conduz à capital, São Luís, pessoas de vários outros estados. Conforme informações contidas no site da Prefeitura, o estabelecimento do território bacabalense deu-se no ano de 1876, com a chegada do Coronel Lourenço Vieira à região, a partir da fundação de uma fazenda com sede no local que hoje se localiza a Praça Nossa Senhora da Conceição. Bacabal é a cidade polo da Região Mearim que é composta por mais oito municípios, a saber Altamira do Maranhão, Bom Lugar, Brejo de Areia, Conceição do Lago Açu, Lago Verde, Olho d'Água das Cunhãs, São Luís Gonzaga do Maranhão e Vitorino Freire.

#### 4.1 A amostra de fala bacabalense

A composição da amostra que será apresentada a seguir foi influenciada diretamente pela limitação de tempo para a realização desta pesquisa. Ainda que o número de participantes não esteja de acordo com o que é adotado em vários estudos da área, buscou-se estratificar a amostra analisada aqui (cf. Quadro 1) conforme o modelo utilizado por outras pesquisas (cf. TEIXEIRA, 2017; SCARDUA, 2018), que se ocupam do mesmo tema. Desse modo, foram gravadas e transcritas 12 entrevistas com informantes que nasceram em Bacabal, ou que se mudaram para a cidade com até 3 anos de idade. Todas as entrevistas foram gravadas com o auxílio do aplicativo RecForge II (Versão 1.2.7.5.g), disponibilizado para Android. Os informantes foram estratificados de acordo com o seu sexo/gênero, três faixas etárias (18 a 30 anos; 31 a 49 anos e 50 anos ou +) e sua escolaridade (ensino médio e ensino superior), conforme o quadro abaixo.

**Quadro 1:** Perfis sociolinguísticos da amostra da fala bacabalense

Sexo/gênero	Faixa etária	Escolaridade	Perfil Sociolinguístico
Masculino	1 (18-30)	Ens. Médio	1. M1M
		Ens. Superior	2. M1S
	2 (31-49)	Ens. Médio	3. M2M
		Ens. Superior	4. M2S
	3 (50 +)	Ens. Médio	5. M3M
		Ens. Superior	6. M3S
Feminino	1 (18-30)	Ens. Médio	7. F1M
		Ens. Superior	8. F1S
	2 (31-49)	Ens. Médio	9. F2M

		Ens. Superior	10. F2S
	3 (50 +)	Ens. Médio	11. F3M
		Ens. Superior	12. F3S

**Fonte:** Elaborado pelo autor

## 4.2 Roteiro de entrevista

O roteiro para as entrevistas é composto de duas partes e baseou-se no roteiro utilizado pelo Projeto SP2010<sup>6</sup>. Na primeira parte, conversa-se sobre assuntos gerais, como o bairro em que mora o falante, sua infância, sua família, educação, sua ocupação, suas redes sociais e atividades de lazer. A inclusão desses assuntos de cunho mais pessoal contribui para o desenvolvimento de uma conversa mais fluida e espontânea. A segunda parte do roteiro compreende assuntos relacionados à cidade de Bacabal, bem como avaliações sobre determinadas variantes linguísticas (como concordância nominal de número padrão e não padrão, pronúncia do /r/ pós-vocálico como tepe e retroflexo, e subjuntivo/indicativo). Ao final da entrevista, pediu-se ao informante que lesse uma lista de palavras, um texto jornalístico e um depoimento com marcas de oralidade.

Para a transcrição das entrevistas, utilizou-se o Elan (cf. HELLWIG; GEERTS, 2019), que, de acordo com Oushiro (2014), diz respeito a um programa computacional para anotação de arquivos de áudio e vídeo, que vem sendo muito utilizado por estudiosos da língua e apresenta vantagens como a sincronização entre arquivo de mídia e a transcrição/anotação, ferramentas mais sofisticadas de buscas dentro de um corpus, flexibilidade de formatos de exportação da transcrição, além do fato de ser gratuito.

## 4.3 Variável dependente

A variável dependente realização de concordância nominal de número tem as seguintes variantes: marcação redundante de plural (CN) e não marcação redundante de plural (CNØ) (as casaS; os faróIS; os anÕES; os lugarES /as casaØ). No exemplo, abaixo, o documentador é representado pela sigla D1 e o informante é representado pela sigla S1.

(1)D1: você gosta de morar aqui?

S1: sim é é bom

S1: as pessoas algumas pessoas fala que é perigoso mas ah eu pessoalmente não acho não

D1: uhum

D1: eu sei por que você escolheu morar nesse bairro?

S1: eu acho assim na época não foi nem tanto assim... por escolha não foi mesmo

S1: por necessidade

<sup>6</sup> MENDES, R.B.; OUSHIRO, L. (2013) Documentação do Projeto SP2010 – Construção de uma amostra da fala paulistana. Disponível em <<http://projetosp2010.fflch.usp.br/producao-bibliografica>>. Acesso em: 03 de março de 2020.

D1: uhum

S1: eh que a gente veio embora do interior e eh porque meus pai já tinha alguns conhecido aqui...

(RicardoL.M2M)

Como evidencia o excerto, ocorre que o mesmo falante utiliza, em sua fala, tanto a marcação redundante de plural, como em “as pessoas” e “algumas pessoas”, quanto a não marcação redundante de plural, como em “meus pai” e “alguns conhecido”. Nota-se que difere daquilo que se prescreve sobre a marcação redundante (cf. BECHARA, 2009: 543-547).

#### 4.4 Variáveis linguísticas

A partir da análise qualitativa dos dados, foram definidas as variáveis/hipóteses linguísticas. A primeira variável é a formação do plural do elemento à direita, que já teve sua relevância estatística testada e confirmada no que concerne à realização de concordância nominal zero (cf. SCHERRE, 1988). Nessa variável, o conceito de saliência fônica proposto por Lemle e Naro (cf. SCHERRE, 1988) é central por considerar que, em alguns casos, as formas singular e plural são mais salientes em sintagmas como feijão/feijões, do que naqueles como casa/casas. Com base nisso, a hipótese é a de que as formas mais salientes sejam mais propensas a serem marcadas com plural do que as menos salientes. Foram analisadas as variantes plural regular “S”, plural em “IS”, plural em “ÕES” e plural em “ES”.

Foi analisada também a tonicidade do elemento nuclear (SCHERRE, 1988), para a qual se levanta a hipótese de que sintagmas com oxítonos e os monossílabos, como em pé, fê, país e leitão, por terem acento na sílaba que vai receber o morfema de plural, favoreçam mais a aplicação da regra de concordância do que os paroxítonos e proparoxítonos. Assim, foram analisados sintagmas com monossílabos tônicos, oxítonos, paroxítonos, proparoxítonos. Também foi analisado o número de sílabas do elemento nuclear (SCHERRE, 1988), observando-se a hipótese de que quanto maior o número de sílabas do elemento nuclear do SN, maior a probabilidade de haver a marcação de plural. Foram controladas as variantes com uma sílaba, duas sílabas, três sílabas e quatro ou mais sílabas.

As variáveis classe morfológica do elemento pré-nuclear e animacidade do SN também foram analisadas. Com relação à variável classe morfológica do elemento pré-nuclear, estudos evidenciam que os artigos e pronomes são mais propensos a receberem a marcação do plural do que os substantivos e adjetivos (SCHERRE, 1988; CARVALHO, 1997; OUSHIRO, 2015). Foram controladas as variantes adjetivo, artigo definido, artigo indefinido, preposição mais artigo, preposição mais demonstrativo, pronome demonstrativo, pronome indefinido, pronome possessivo e substantivo.

Por fim, foi analisada a animacidade do SN, com a hipótese de que sintagmas com traços mais humanos (pessoas) sejam mais propensos a receber a marcação de

concordância do que os sintagmas com traços menos humanos “coisas” (SCHERRE, 1988; BRANDÃO E VIEIRA, 2012; OUSHIRO, 2015). Foram controladas as variantes mais humano e menos humano.

#### 4.5 Variáveis sociais

Considerando uma pesquisa sociolinguística, que prioriza aspectos sociais, além de coletar a fala de indivíduos em situações reais de uso, outro princípio fundamental, é necessário analisar algumas características sociais desses falantes, a fim de descrever e buscar uma explicação para o comportamento de determinados fenômenos da língua. Por isso, para entender como as características sociais podem influenciar na realização da concordância nominal de número, foram analisadas as variáveis sociais sexo/gênero, faixa etária e escolaridade.

As variáveis sexo/gênero, faixa etária, escolaridade, raça, profissão e outras de cunho social são variáveis essenciais para o desenvolvimento de uma investigação sociolinguística. Na verdade, compreender como aspectos sociais contribuem para o processo de variação e mudança de uma língua é uma das principais contribuições da Sociolinguística (LABOV, 2008 [1972]).

A faixa etária é essencial para o interesse do estudo por verificar através do construto do tempo aparente (LABOV, 2008 [1972]) se o fenômeno se encontra em variação estável ou se há indício de mudança em curso. Alicerçado na hipótese da aquisição da linguagem cuja fase final se daria por volta dos quinze anos, o construto do tempo aparente propõe que “o estado atual da língua de um falante adulto reflete o estado da língua adquirida quando o falante tinha aproximadamente quinze anos de idade” (NARO, 1992: 82 apud SANTOS; PIMPÃO, 2018: 261). Por meio da análise da faixa etária, com aparentes recortes no tempo, podem-se observar as diferenças ou semelhanças entre os falantes, já que o comumente esperado é que os mais velhos tendam a realizar mais a forma padrão relativamente aos mais jovens (cf. TEIXEIRA, 2017)

O sexo/gênero, por sua vez, permite observar se há diferenças na realização da concordância nominal entre homens e mulheres, tendo em vista que alguns estudos têm afirmado que as mulheres tendem a favorecer a utilização das formas padrão ou mais bem avaliadas socialmente (cf. SILVA, 2014; OUSHIRO, 2015). Oushiro (2015) evidencia que os homens produziram mais a concordância sem a marca de plural CNO do que as mulheres tanto na análise atomística quanto na análise não atomística, com os respectivos pesos relativos (0.64 e 0.63) e (0.38 e 0.36).

Com a análise da escolaridade, pode-se depreender, a depender do grau de escolarização do falante, se ele tende a realizar mais a forma considerada padrão. Estudos (cf. OUSHIRO, 2015; TEIXEIRA, 2017) têm mostrado que os anos de formação superior proporcionam ao falante o uso mais frequente da marcação do

plural. Teixeira (2017) mostra que os informantes com ensino médio realizaram a variante não prestigiada em 76,9% das ocorrências relativamente aos informantes com nível superior, que fizeram uso da variante não prestigiada em 54,9% das ocorrências, com os respectivos pesos relativos 0.661 e 0.418. Desse modo, buscou-se verificar qual a estratificação social da concordância nominal de número na variedade bacabalense do português. As variáveis sociais apresentadas foram analisadas a fim de se compreender como aspectos sociais contribuem no processo de variação do fenômeno estudado.

#### 4.6 O programa R

As análises estatísticas necessárias para se verificar quais fatores sociais e linguísticos se correlacionam à realização e à não realização da concordância de SNs no português bacabalense foram feitas na versão mais recente do R (cf. R CORE TEAM, 2019). Oushiro (2014) explica que esse programa é voltado para a análise de dados linguísticos, e pode ser utilizado para realizar computações estatísticas e gráficas, como também compilar e anotar corpora, produzir listas de frequências e outras tarefas.

Além disso, o R permite que o pesquisador otimize o seu tempo durante a realização de diversas tarefas. Gries (2009) afirma que, por ser uma linguagem de programação, o R permite ao pesquisador um controle direto das funções, diferentemente de outros programas, em que o pesquisador depende do desenvolvedor. Isso o torna muito mais versátil do que qualquer outro programa pronto, como o GoldVarb X. Em outras palavras, o R permite ao usuário customizar tarefas, o que, nas palavras de Oushiro (2014: 134) “significa [...] que ao invés de clicar em botões com funções limitadas e pré-definidas, o usuário normalmente define as funções que deseja executar através de linhas de comando”.

Levshina (2015) explica que o R permite realizar a análise de regressão múltipla, que possibilita mensurar o efeito particular de cada variável independente enquanto se trabalha com as outras variáveis. Essa observação de Levshina traz uma das mais importantes contribuições dessa linguagem de computação. Com o R, o pesquisador tem a possibilidade de analisar duas ou mais variantes (análise multinomial); já, no GoldVarb X, isso não é possível. Scherre (2012: 4-5) afirma que, o GoldVarb X “só efetua a análise de pesos relativos de duas variantes (binomial), em um nível (*one level*) ou em múltiplos níveis (*up and down*)”.

Na seção, a seguir, são apresentados a distribuição geral dos dados obtidos a partir das entrevistas realizadas com bacabalenses e os resultados da análise estatística das variáveis sociais e linguísticas.

## 5 Análise dos dados

Nesta seção, apresentam-se a distribuição geral dos dados, a partir das entrevistas com bacabalenses, e os resultados das análises quantitativas dos SNs, realizadas no R (R CORE TEAM 2019), obtidos com base em modelos de regressão logística. Os SNs foram analisados sob duas perspectivas: a **não atomística**, em que se analisa o SN como um todo (as casas) e a **atomística**, que considera como dado de análise cada elemento do SN de forma isolada (as/ casas). Enfatiza-se que as duas análises se complementam e ajudam a revelar diferentes processos, seja no nível da palavra, seja no nível do sintagma (OUSHIRO, 2015). No presente artigo, apresentam-se apenas os resultados da análise **não atomística**. Das 12 entrevistas, foram extraídos sintagmas nominais plurais com duas palavras em que pelo menos uma tenha sido marcada com plural. Foram excluídos dados invariáveis como férias, numerais cardinais e dados em que a noção de plural é semântica como “muita coisa”. Nesta análise, foram considerados 988 dados. A tabela 1, a seguir, que traz a distribuição geral das ocorrências, já indicia variação na realização da concordância nominal de número, principalmente na fala dos mais velhos e dos menos escolarizados.

### 5.1 Distribuição geral dos dados

**Tabela 1:** Distribuição geral dos SNs

		Dados não atomísticos		
		CNØ	%	Total
Sexo/Gênero	Masculino	132	25	516
	Feminino	139	29	472
Faixa Etária	1 (18 a 30 anos)	45	12	365
	2 (31 a 49 anos)	107	31	344
	3 (50 anos ou mais)	119	42	279
Escolaridade	Ensino Médio	193	39	486
	Ensino Superior	78	15	502

**Fonte:** Elaborado pelo autor.

Para a variável sexo/gênero, a tabela 1 mostra que as proporções de realização e não realização parecem não diferir muito entre homens e mulheres. No entanto, evidencia-se que a taxa de não realização é um pouco maior entre as falantes. Os homens produziram a não realização em 25% das ocorrências, enquanto as mulheres não utilizaram a concordância padrão em 29% do total de dados.

Considerando-se a faixa etária dos informantes, observa-se que a taxa de realização redundante de concordância nominal de número é muito maior entre os mais jovens (88%). Por outro lado, os mais velhos foram os que menos produziram a

realização de concordância padrão (58%) e, conseqüentemente, foram os que mais realizaram a concordância não padrão (42%). Os informantes da 2ª faixa etária, por sua vez, realizaram a concordância não padrão em 31% das ocorrências.

Em relação à escolaridade, as proporções mostram que os informantes com mais escolaridade tenderam a realizar mais a concordância, em relação aos informantes com menos escolaridade. Os informantes com ensino superior produziram a não realização em 15% das ocorrências, enquanto os informantes com ensino médio não utilizaram a concordância padrão em 39%.

## 5.2 Análise não atomística

As análises de regressão logística aqui apresentadas visam a examinar quais variáveis sociais e linguísticas se correlacionam ao fenômeno estudado. Oushiro (2017:182) afirma que “o interesse nas análises de regressão logística é verificar o efeito simultâneo de múltiplas variáveis previsoras, a fim de chegar a um modelo para descrever, explicar e prever o comportamento da variável resposta”. Desse modo, o principal intuito é o de mensurar o efeito das variáveis independentes sobre a variável dependente, a fim de chegar a um modelo que explique de maneira satisfatória o fenômeno em tela.

Em uma primeira análise de regressão logística, consideraram-se todas as variáveis e as suas respectivas variantes (tabela omitida aqui). No entanto, percebeu-se que em algumas variantes de algumas variáveis havia uma quantidade pequena de dados. Diante disso, foi necessário amalgamá-las, a fim de que fossem obtidos resultados que possibilitassem explicar e descrever o fenômeno de forma mais adequada. Para sustentar a ideia da amalgamação de algumas variantes, recorreu-se ao princípio heurístico da Navalha de Ockham proposto por Guilherme Ockham. Naro e Scherre (2007: 182) afirmam que, “esta doutrina ensina a cortar qualquer excrescência teórica que não contribua para explicar os dados relevantes, exatamente como se usava na época uma navalha para eliminar excessos de tinta por acaso pingados nos documentos manuscritos”.

Apoiando-se na Navalha de Ockham, partindo da premissa de que menos é mais, foram amalgamadas variantes de fatores linguísticos que apresentaram poucos ou nenhum dado em algumas células, ao invés de simplesmente excluí-las, a fim de garantir a obtenção de mais valores significativos. Conforme o exposto, as variantes das variáveis número de sílabas do elemento nuclear, tonicidade do elemento nuclear e classe morfológica do elemento pré-nuclear foram amalgamadas. Após todas as amalgamações, realizou-se uma nova análise com as variáveis sociais e linguísticas. Com os resultados dessa última análise, dispostos na tabela 2, a seguir, é possível afirmar que todas as variáveis se correlacionam, em alguma medida, ao fenômeno de concordância nominal de número, com exceção da variável social sexo/gênero, que não apresentou diferenças significativas entre os dados de homens e mulheres.

Na tabela 2, abaixo, os valores da coluna estimativa devem ser lidos em relação ao **intercept** (ou valor de referência – aqui, a não realização de concordância nominal), que nessa tabela foram definidos, pelo R, por meio de seu valor **default** (ordem alfabética e/ou ordem numeral crescente, exceto para faixa etária, cujo valor de referência são os falantes da 3ª faixa etária). Abaixo do intercept são apresentadas as variáveis previsoras, ou independentes. A coluna seguinte traz informações sobre as medidas estimadas para cada variante, que aqui são cada uma das variáveis previsoras. A **estimativa** é o resultado da soma entre o nível de referência (intercept) e o valor da variante que esteja sendo analisada. A coluna seguinte apresenta o **erro padrão**. De acordo com Oushiro (2017), trata-se de uma medida de dispersão que mostra quão preciso é o valor da estimativa: quanto menor, maior é o grau de precisão do modelo de análise. Em seguida, são apresentados os **valores de z e p**. O primeiro calcula o valor de significância (divisão entre os valores da estimativa pelo erro padrão), e que mostra a probabilidade de se observar uma distribuição, para o caso de a hipótese nula (H0) ser verdadeira – aqui, por exemplo, uma das H0 é a de que não há correlação entre deixar de realizar concordância padrão (variável resposta ou dependente) e o sexo dos falantes<sup>7</sup> (variável preditora ou independente), ou seja, não há diferença entre as estimativas de homens e mulheres; o segundo, por sua vez, apresenta os níveis de significância. O valor convencional para que se calcule a significância é o limite de 5%, e a representação se dá por meio do símbolo p ( $p < 0,05$ ). Os níveis de significância são apresentados por meio do uso de asteriscos (\*). Assim, se  $p < 0,05$ , usa-se um \*, dois \*\* para  $p < 0,01$  e três \*\*\* para  $p < 0,001$ .

**Tabela 2:** Estimativas, erros padrão e valores de significância para todas as variáveis, com todas as amalgamações necessárias

	Estimativa	Erro padrão	Valor-z	Significância (p)
<i>Intercept</i>	-3,228	0,722	-4,468	<0,001 ***
Masculino	0,130	0,169	0,768	0,442
Escolaridade Superior	-1,309	0,170	-7,668	<0,001 ***
2ª Faixa Etária	-0,329	0,188	-1,745	0,080.
1ª Faixa Etária	-1,613	0,220	-7,304	<0,001 ***
Plural com -is	-1,003	1,092	-0,919	0,358
Plural com -ões	0,521	0,702	0,741	0,458

<sup>7</sup> Desse modo, a estimativa para que os homens deixem de realizar a concordância nominal é de 0,130 em relação às mulheres. Da mesma maneira, a estimativa para que os falantes com escolaridade universitária deixem de realizar a concordância nominal é de -1,309, em comparação aos falantes com ensino médio. Esses valores devem ser somados ao do intercept para que se tenha o valor real da estimativa dos homens ( $-3,228 + 0,130 = -3,098$ ) e os ouvintes com ensino superior ( $-3,228 + (-1,309) = -4,537$ ) de deixarem de realizar a concordância padrão. Nesse sentido, o valor de  $p < 0,001$  para o Ensino Superior, por exemplo, aponta para o fato de que é muito pequena a chance de alcançar esses resultados em uma situação em que a hipótese nula viesse a ser verdadeira (caso não houvesse diferença entre Ensino Médio e Ensino Superior). Isso indica que a diferença entre essas variantes é significativa.

Plural com -s	1,899	0,347	5,468	<0,001 ***
Paroxítono	0,812	0,375	2,164	<0,05 *
Proparoxítono	1,370	0,546	2,509	<0,05 *
Três sílabas	-0,510	0,254	-2,003	<0,05 *
Uma ou Duas sílabas	-0,259	0,265	-0,975	0,329
Artigo	1,100	0,466	2,357	<0,05 *
Preposição + outra classe	1,143	0,482	2,373	<0,05 *
Pronome	0,981	0,472	2,078	<0,05 *
SN - Menos Humano	0,540	0,191	2,829	<0,01**

Modelo:glm(formula=Variavel~Sexo.Genero+Escolaridade+Faixas+Formação.do.plural.do.elemento.a.direita+Tonicidade.do.elemento.nuclear+Numero.de.silabas.do.elemento.nuclear+Classe.morfologica.do.elemento.pre.nuclear +Animacidade.do.SN, Family = binomial, data = cn  
Intercept: Feminino, Escolaridade Médio, 3ª Faixa etária, Plural com es, Monossílabo tônico ou Oxítono, Quatro sílabas ou mais, Adjetivo ou Substantivo, SN – Mais Humano.

**Fonte:** Elaborado pelo autor

Os resultados da análise estatística, resumidos na tabela 2 acima, mostram que não há diferença significativa entre informantes do sexo feminino (valor de referência) e falantes do sexo masculino (estimativa 0,130,  $p = 0,442$ ), que deixaram de realizar concordância padrão, respectivamente, 29% e 25% das vezes, no total dos dados analisados. Foram verificadas diferenças significativas, no entanto, entre falantes com ensino médio (valor de referência) e falantes com ensino superior, que desfavorecem a não realização de concordância nominal de número (estimativa -1,309,  $p < 0,001$ ). Os informantes com mais escolarização deixaram de realizar a concordância em 15% do total de ocorrências na perspectiva não atomística. São observadas diferenças significativas entre os falantes da 3ª faixa etária (valor de referência), com 50 anos ou mais, e o falantes da faixa etária mais jovem, de 18 a 30 anos, que tendem a realizar menos concordância não padrão (estimativa -1,613,  $p < 0,001$ ) do que os falantes com mais idade. Não se diferenciaram significativamente os falantes com mais idade e os falantes da faixa etária entre 31 e 49 anos (estimativa -0,329,  $p = 0,080$ ).

Em relação às variáveis linguísticas, os resultados da tabela 2 apontam para o fato de que os elementos nominais com formação do plural em -es (valor de referência) se diferenciam significativamente daqueles em que a formação do plural é constituída com -s, que desfavorecem a realização da concordância padrão (estimativa 1,899,  $p < 0,001$ ), mas não dos sintagmas com formação de plural com -ões (estimativa 0,521,  $p = 0,458$ ).

Quanto à tonicidade do elemento nuclear, são observadas diferenças significativas entre os elementos monossílabos tônicos e oxítonos (amalgamados, valor de referência) e os paroxítonos (estimativa 0,812,  $p < 0,05$ ) e proparoxítonos (estimativa 1,370,  $p < 0,05$ ). Os elementos monossílabos tônicos e oxítonos

apresentaram mais a concordância padrão do que os sintagmas paroxítonos e proparoxítonos.

No tocante ao número de sílabas do elemento nuclear do SN, são percebidas diferenças significativas entre os elementos com quatro ou mais sílabas (valor de referência) e os com três sílabas (estimativa -0,510,  $p < 0,05$ ), que favoreceram mais marcas de concordância padrão, mas não entre sintagmas com quatro ou mais sílabas e os com uma ou duas sílabas (estimativa -0,259,  $p = 0,329$ ).

No que se refere à classe morfológica do elemento pré-nuclear do SN, há diferenças significativas entre a variante adjetivo/substantivo (valor de referência) e a variante artigo (estimativa 1,100,  $p < 0,05$ ), entre adjetivo/substantivo e preposição + outra classe (estimativa 1,143,  $p < 0,05$ ), e entre adjetivo/substantivo e pronome (estimativa 0,981,  $p < 0,05$ ). Para todos esses casos, a formação de plural em sintagmas adjetivos e substantivos são os que mais desfavorecem a realização de concordância padrão.

Por fim, em relação à animacidade do SN, observa-se que há diferença significativa entre os elementos com traços mais humanos (valor de referência) e aqueles com características menos humanas (estimativa 0,540,  $p < 0,01$ ), cuja estimativa positiva favorece mais a marcação redundante de plural do que os elementos com traços mais humanos.

### 5.3 Encaixamento social da concordância nominal de número no português falado em bacabal

Para a variável sexo/gênero, o resultado da análise estatística mostrou que não há diferença significativa entre os dados de homens e mulheres. Considerando-se os dados não atomísticos, os homens realizaram a concordância não padrão 25% das ocorrências, enquanto as mulheres não utilizaram a concordância padrão em 29% do total de dados. Desse modo, a hipótese de que as mulheres tendem mais à concordância padrão não se confirma.

Em comparação a outros estudos variacionistas, o resultado dessa variável se mostrou divergente, tendo em vista que algumas pesquisas evidenciaram que há correlação e mostram que as mulheres são mais sensíveis às consideradas normas de prestígio. Scherre (1988) evidenciou que essa variável foi verificada como relevante, e que as mulheres tendem à a marcação da concordância padrão. Oushiro (2015) constatou em seu estudo que os homens são os que mais favorecem a concordância não padrão, enquanto que as mulheres a desfavorecem.

Para o contexto do português maranhense, Teixeira (2017), que estudou a concordância nominal de número na cidade de Caxias, também verificou que a variável sexo/gênero não foi considerada relevante na análise. A autora mostra que o peso relativo da variante sexo feminino (0.50) não é significativo e, portanto, não se pode afirmar que mulheres favorecem o uso da concordância padrão. Em se tratando de frequências, as mulheres produziram a concordância não padrão em 62,7% dos

dados analisados, já os homens realizaram a concordância não padrão em 61,9% do total de dados, o que revela uma diferença muito pequena entre esses grupos de falantes.

Quando se trata da escolaridade dos informantes, o resultado da perspectiva não atomística evidencia que os falantes com ensino médio se diferenciam significativamente daqueles com ensino superior, que favorecem a realização de concordância nominal de número. Os falantes com mais escolaridade produziram a concordância não padrão em 15% das ocorrências, enquanto os informantes com ensino médio não utilizaram a concordância padrão em 39% dos dados. Confirmasse, então, a hipótese de que pessoas com mais escolaridade tendem à preservação da marcação redundante de plural.

Oushiro (2015) também evidenciou que a escolaridade se correlaciona ao fenômeno de concordância. A partir da estratificação dos informantes em três níveis de escolaridade (ensino fundamental, ensino médio e ensino superior), a autora percebeu, segundo suas próprias palavras, uma correlação gradual. Nesse sentido, tanto na perspectiva atomística quanto na não atomística, os informantes com ensino fundamental favoreceram fortemente a concordância não padrão (P.R. 0.85, para a análise atomística, e P.R. 0.81, para a análise não atomística); os que possuem o ensino médio favoreceram em grau menor (P.R. 0.67, na análise atomística, e P.R. 0.64, na não atomística), enquanto aqueles que concluíram o ensino superior desfavorecem bastante a variante não padrão (P.R. 0.32 e P.R. 0.35, respectivamente).

Teixeira (2017) observou que houve o desfavorecimento da concordância padrão por parte dos falantes com ensino médio, com diferença significativa corroborada pelo peso relativo 0.66. A autora enfatizou que o ingresso no ensino superior estimulou os falantes a produzirem mais a concordância padrão. No que se refere às taxas, os falantes com ensino superior deixaram de realizar a concordância padrão em 54% do total das ocorrências. Já os informantes com ensino médio produziram a concordância não padrão em 76% das ocorrências.

Quanto à faixa etária, observou-se diferença significativa entre os falantes da 3ª faixa etária e os falantes da faixa etária mais jovem, de 18 a 30 anos, que tendem a realizar menos concordância não padrão do que os falantes com mais idade, tanto na análise não atomística quanto na atomística. Verifica-se, portanto, que não há um indício de mudança em curso e sim uma variação estável, com a tendência de preservação das marcas de plural em todos os elementos do SN. Em Oushiro (2015), constatou-se que essa variável não foi selecionada como estatisticamente significativa, tanto para análise atomística quanto para não atomística. A autora verifica que a concordância nominal de número parece ser um caso de variação estável na cidade de São Paulo.

Teixeira (2017:88) evidenciou que essa variável não tem tanta relevância estatística, pois “os pesos calculados também estão próximos, 0.55 para informantes

adultos, e 0.45 para informantes jovens; contudo, a diferença é um pouco mais relevante no caso de informantes adultos”. Em relação às frequências, não houve a realização de concordância padrão entre os mais jovens, em 57,7% dos dados, e na faixa dos adultos, 67% dos dados são ocorrências da variante não padrão. Diante do exposto, considerando a amostra analisada, o encaixamento social indica que a concordância nominal de número, em Bacabal, é fenômeno variável estável que apresenta forte correlação à escolaridade e à faixa etária dos informantes. Esses resultados se aproximam, de certo modo, a outras pesquisas aqui resenhadas.

## 6 Considerações finais

A análise evidenciou que a maioria das variáveis se mostrou significativa. Quase todas as variáveis mostraram valores significativos, com exceção da variável social sexo/gênero, que não apresentou diferença significativa entre os dados de homens e mulheres. Revelou-se, por meio da análise da amostra da fala bacabalense que a não marcação redundante de plural, ou forma não padrão da concordância, foi produzida por todos os informantes. Isso confirma a ideia de que a variação da concordância nominal de número é inerente ao português brasileiro (cf. SCHERRE, 1988; OUSHIRO, 2015).

Em síntese, este estudo verificou que, em geral, conforme mencionado acima, há variação na realização da concordância nominal de número em sintagmas nominais. Nesse sentido, por outro lado, também se evidenciou que os falantes bacabalenses tendem mais à preservação das marcas de plural em todos os elementos do que a não marcação redundante de plural. Nota-se, porém, que os informantes mais velhos e os menos escolarizados, tendem a um desfavorecimento maior da norma padrão.

Com a análise de dados coletados em uma comunidade urbana não tão grande, mas mista, ainda que a amostra de fala analisada tenha uma quantidade de informantes idiossincrática, o que faz com que ela não ofereça generalizações mais profundas da realidade sociolinguística da cidade, espera-se ter contribuído para o início do estabelecimento de padrões sociolinguísticos em Bacabal. Dito de outro modo, acredita-se que esta pesquisa foi capaz de apresentar os padrões gerais de uso de uma das variáveis sociolinguísticas mais produtivas do português brasileiro. Espera-se que, a partir daqui, outras pesquisas possam ser desenvolvidas.

## Referências

- ALVES, Cibelle Corrêa Béliche. *O uso do tu e do você no português falado no Maranhão*. 2010. 143f. Dissertação de Mestrado. Universidade Federal do Ceará, Fortaleza, 2010.
- BECHARA, Evanildo. *Moderna gramática portuguesa*. – 37. ed. rev., ampl. e atual. conforme o novo acordo ortográfico. Rio de Janeiro: Nova Fronteira, 2009.
- BRAGA, M. L. *A concordância de número no sintagma nominal no triângulo mineiro*. 1977. Dissertação (Mestrado em Linguística) - Faculdade de Letras, Pontifícia Universidade Católica do Rio de Janeiro - PUC/RJ, Rio de Janeiro, 1977.
- CALVET, Louis-Jean. *Sociolinguística: uma introdução crítica*. Trad. Marcos Marcionilo. São Paulo: Parábola Editorial, 2002.
- CAMACHO, Roberto Gomes. *Da linguística formal à linguística social*. São Paulo: Parábola Editorial, 2013.
- CARVALHO, R. C. *A concordância de número no sintagma nominal na fala urbana de Rio Branco*. 1997. 182f. Dissertação (Mestrado em Linguística) - Universidade Estadual de Campinas, Instituto de Estudos da Linguagem - UNICAMP, Campinas, 1997.
- CASTILHO, Ataliba T. de. *Nova gramática do português brasileiro*. 1. ed, 3ª reimpressão. São Paulo: Contexto, 2014.
- CHOMSKY, Noam. *Aspectos da teoria da sintaxe*. Trad. de J.A. Meireles, E. Raposo. Coimbra: Armênio-Amado Editor, 1978.
- CHOMSKY, Noam. *Estruturas sintáticas*. Tradução e comentários de Gabriel de Ávila Othero e Sérgio de Moura Menuzzi. – Petrópolis, RJ: Vozes, 2018. – (Coleção de Linguística) Título original: Syntactic structures, 1957.
- FIAMENGUI, Ana Helena Rufo. *A marcação de pluralidade no SN na fala e na escrita de adolescentes da região de São José do Rio Preto*. 2011. 143f. Dissertação (Mestrado em Estudos Linguísticos) – Instituto de Biociências, Letras e Ciências Exatas. Universidade Estadual Paulista, São José do Rio Preto.
- GUY, Gregory; ZILLES, Ana. *Sociolinguística quantitativa – um instrumental de análise*. São Paulo: Parábola Editorial, 2007.
- GRIES, Stefan Th. *Quantitative Corpus Linguistics with R: a practical introduction*. New York and London: Routledge, 2009.
- HELLWIG, B.; GEERTS, J. ELAN – *Linguistic Annotator*. Versão 5.6. 2019. Disponível em: <<https://tla.mpi.nl/tools/tla-tools/elan/>>. Acesso em: 14 de junho de 2019.
- LABOV, William. *Padrões sociolinguísticos*. Tradução de Marcos Bagno, Maria Marta Pereira Scherre, Caroline Rodrigues Cardoso. São Paulo: Parábola Editorial, 2008 [1972].
- LEVSHINA, Natalia. *How to do Linguistics with R: data exploration and statistical analysis*. Amsterdam / Philadelphia: John Benjamins Publishing Company, 2015.
- LOPES, Norma da Silva. *Concordância nominal, contexto linguístico e sociedade*. 408 f. Tese de Doutorado. Universidade Federal da Bahia, Salvador, 2001.
- MARTELOTTA, Mário Eduardo, *Manual de linguística (org.)*. 2. ed. São Paulo: Contexto, 2011.

MOLLICA, Maria Cecília. Fundamentação teórica: conceituação e delimitação, p. 9-14. In Maria Cecília Mollica, Maria Luiza Braga (Organizadoras) *Introdução à sociolinguística: o tratamento da variação*. 4. Ed. 3ª reimpressão. São Paulo: Contexto, 2015.

OUSHIRO, Livia. Transcrição de Entrevistas sociolinguísticas com o Elan, p.117-132. In Raquel Meister Ko. Freitag (Organizadora). *Metodologia de Coleta e Manipulação de Dados em Sociolinguística*. São Paulo: Editora Edgard Blücher, 2014.

\_\_\_\_\_. Tratamento de dados com o R para Análises Sociolinguísticas, p.134-177. In Raquel Meister Ko. Freitag (Organizadora). *Metodologia de Coleta e Manipulação de Dados em Sociolinguística*. São Paulo: Editora Edgard Blücher, 2014.

\_\_\_\_\_. *Identidade na pluralidade: avaliação, produção e percepção linguística na cidade de São Paulo*. 372 f. Tese de Doutorado em Linguística. Universidade de São Paulo, São Paulo, 2015.

\_\_\_\_\_. *Introdução à Estatística para Linguistas*, v.1.0.1 (dez/2017). Disponível em: <<https://doi.org/10.5281/zenodo.822069>>. Licença Creative Commons 4.0 Atribuição – Não comercial. Acesso em: 21 de novembro de 2019.

RAMOS, Conceição de Maria Araujo; BEZERRA, José de Ribamar Mendes. O processo de formação de plural em itens nominais na fala Maranhense, p. 13-27. In: *O português falado no Maranhão: múltiplos olhares*. São Luís: EDUFMA, 2010

\_\_\_\_\_. (Coord.). *Projeto Atlas Linguístico do Maranhão*. São Luís: Universidade Federal do Maranhão. (Departamento de Letras). 88p. Projeto em andamento, 1999.

\_\_\_\_\_. (Orgs.). *O português falado no Maranhão: estudos preliminares*. São Luís: EDUFMA, 2005.

\_\_\_\_\_. (Orgs.). *A diversidade do português falado no Maranhão: o atlas linguístico do Maranhão em foco*. São Luís: EDUFMA, 2006.

\_\_\_\_\_. (Orgs.). *O português falado no Maranhão: múltiplos olhares*. São Luís: EDUFMA, 2010.  
R CORE TEAM (2019). R: *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Disponível em: <<http://www.R-project.org/>>. Acesso em: 14 de junho de 2019.

ROCHA LIMA, Carlos Henrique da. *Gramática normativa da língua portuguesa*. 49. ed. Rio de Janeiro: José Olympio, 2011.

SANKOFF, D.; TAGLIAMONTE, S.; SMITH, E. GoldVarb X: *A variable rule application for Macintosh and Windows*. Department of Linguistics, University of Toronto, 2005.

SANTOS, Wendel Silva dos; PIMPÃO, Tatiana Schwochow. *Variação estável ou mudança? A expressão do modo subjuntivo em três variedades do português brasileiro*. Caderno Seminal Digital Especial, n° 30 v. 30 (JAN-DEZ/2018) – e-ISSN 1806-9142. Disponível em: <<http://dx.doi.org/10.12957/cadsem.2018.33066>>. Acesso em: 07 de junho de 2019.

SANTANA, Amanda de Lima. *As vogais médias pretônicas na fala de sergipanos em São Paulo*. Dissertação de Mestrado. 147f. Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2018.

SAUSSURE, Ferdinand de. *Curso de linguística geral*. São Paulo: Cultrix, 1969.

SCARDUA, Juliana Rangel. *Análise da concordância nominal na fala de Vitória/ES: o linguístico, o social e o estilístico*. 213f. Dissertação (Mestrado em Linguística) – Programa de Pós-Graduação em Linguístico. Universidade Federal do Espírito Santo, 2018.

SALOMÃO, Mircia Hermenegildo. *A variação de pluralidade nas estruturas predicativas da variedade falada na região de São José do Rio Preto*. Dissertação de mestrado. UNESP, 2010.

SCHERRE, M. M. P. & NARO, A. J. *Sobre a concordância de número no português falado do Brasil*. In Ruffino, Giovanni (org.) *Dialettologia, geolinguística, sociolinguística*. (Atti del XXI Congresso Internazionale di Linguistica e Filologia Romanza) Centro di Studi Filologici e Linguistici Siciliani, Università di Palermo. Tübingen: Max Niemeyer Verlag, 5:509-523, 1998.

SCHERRE, Maria Marta Pereira. *Reanálise da concordância nominal em português*. 555f. Tese de Doutorado em Linguística. Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1988.

\_\_\_\_\_. *Padrões sociolinguísticos do português brasileiro: a importância da pesquisa variacionista*. Revista do Programa de Pós-Graduação em Estudo de Linguagens. Universidade do Estado da Bahia – UNEB. Departamento de Ciências Humanas – DCH I. NÚMERO 04 – junho de 2012.

SILVA, Fernando Gomes da. *Alagoanos em São Paulo e a concordância nominal de número*. 103 f. Dissertação de Mestrado. Universidade de São Paulo, São Paulo, 2014.

TARALLO, Fernando. *Pesquisa sociolinguística*. 8. ed. São Paulo: Ática, 2007.

TEIXEIRA, V. P. *Variação Linguística e Fluxos Migratórios: a concordância nominal de número na fala dos moradores do bairro Campo de Belém do município de Caxias – MA*. 2017. 124f. Dissertação – Programa de Pós-Graduação em Letras, Universidade Federal do Piauí, Teresina, 2017.

WEINREICH, Uriel; LABOV, William; HERZOG, Marvin. I. *Fundamentos empíricos para uma teoria da mudança linguística*. São Paulo: Parábola Editorial, 2006.

# The nominal number agreement in Bacabal-MA: a sociolinguistic study

**Revista Falange Miúda**  
ISSN 2525-5169

**Periodicity:**  
Thematic issue

**Volume 5**  
**Number 2**

**Received in:** 04/03/2020  
**Approved in:** 19/06/2020

**Abstract:**

Based on the assumptions of Sociolinguistics Variationist (WEINREICH; LABOV; HERZOG, 2006 [1968]; LABOV, 2008[1972]), this research aimed to investigate the restriction of linguistic and social factors that correlate with the achievement of nominal number agreement in simple nominal phrases in the spoken Portuguese Bacabal, Maranhão. To this end, 12 sociolinguistic interviews with speakers of this city were analyzed. The results of the logistic regression analysis, made in the R program (R CORE TEAM, 2019), point to the fact that level of education and the age group of the speakers are the social variables that correlate with the phenomenon. In relation to linguistic variables, the analysis showed that all variables correlate to the achievement of agreement.

**Keyword:**

Nominal number agreement. Bacabal Portuguese. Linguistic Variation