



UNIVERSIDADE FEDERAL DE PERNAMBUCO  
CENTRO DE ARTES E COMUNICAÇÃO  
DEPARTAMENTO DE LETRAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS

**MARÍLIA ADRIELLE SIQUEIRA DE OLIVEIRA**

**USO VARIÁVEL DO ARTIGO DEFINIDO DIANTE DE ANTROPÔNIMOS: UM  
ESTUDO SOCIOLINGUÍSTICO SOBRE O PORTUGUÊS FALADO EM RECIFE**

Recife

2021

**MARÍLIA ADRIELLE SIQUEIRA DE OLIVEIRA**

**USO VARIÁVEL DO ARTIGO DEFINIDO DIANTE DE ANTROPÔNIMOS: UM  
ESTUDO SOCIOLINGUÍSTICO SOBRE O PORTUGUÊS FALADO EM RECIFE**

Dissertação apresentada ao Programa de Pós-Graduação em Letras da Universidade Federal de Pernambuco, como requisito parcial à obtenção do Título de Mestra em Letras.

**Área de concentração:** Linguística

**Orientador:** Prof. Dr. Marcelo Amorim Sibaldo.

Recife

2021

O48u Oliveira, Marília Adrielle Siqueira de  
 Uso variável do artigo definido diante de antropônimos: um estudo sociolinguístico sobre o português falado em Recife / Marília Adrielle Siqueira de Oliveira. – Recife, 2021.  
 144p.: il.

Orientador: Marcelo Amorim Sibaldo.  
 Dissertação (Mestrado) – Universidade Federal de Pernambuco. Centro de Artes e Comunicação. Programa de Pós-Graduação em Letras, 2021.

Inclui referências e apêndices.

1. Artigo definido. 2. Antropônimos. 3. Sociolinguística Variacionista. 4. Recife. I. Sibaldo, Marcelo Amorim (Orientador). II. Título.

410 CDD (22. ed.) UFPE (CAC 2021-147)

**MARÍLIA ADRIELLE SIQUEIRA DE OLIVEIRA**

**USO VARIÁVEL DO ARTIGO DEFINIDO DIANTE DE ANTROPÔNIMOS: UM  
ESTUDO SOCIOLINGUÍSTICO SOBRE O PORTUGUÊS FALADO EM RECIFE**

Dissertação apresentada ao Programa de Pós-Graduação em Letras da Universidade Federal de Pernambuco como requisito parcial para a obtenção do título de Mestre em Letras.

Aprovada em: 26/05/2021.

**BANCA EXAMINADORA**

---

Prof. Dr. Marcelo Amorim Sibaldo (Orientador)  
Universidade Federal de Pernambuco

---

Prof<sup>a</sup>. Dr<sup>a</sup>. Simone Azevedo Floripi (Examinadora Externa)  
Universidade Tecnológica Federal do Paraná

---

Prof. Dr. Adeilson Pinheiro Sedrins (Examinador Externo)  
Universidade Federal do Agreste Pernambucano

Este trabalho é dedicado a um amor ancestral, à minha alma gêmea, Maria Adriana.  
Obrigada por evoluir comigo, nessa e em muitas outras vidas.

## **AGRADECIMENTOS**

Agradeço primeiramente a minha mãe, Maria do Carmo, por todo amor e dedicação; obrigada pela vida e por ser meu exemplo maior desde sempre, enquanto mulher, mãe e professora. Agradeço à minha alma gêmea, Adriana, somos filhas gêmeas da mesma matéria, mas nosso amor é ancestral, te amo de outras vidas. Agradeço a Mariana, minha irmã, pelas lições de vida que a sua vida me ensina. Agradeço às minhas tias, Francisca e Messias, por cuidarem do mundo como mães. Vocês são minha força. A minha força é mulher!

Agradeço, também, a Luis Patrício, meu companheiro, por ser você, pelos encontros, pela paciência e pelo amor; tem sido lindo dançar a vida ao teu lado.

A todos os meus professores e professoras da Pós-Graduação, obrigada! Agradeço ao professor Adeilson, meu orientador de sempre, por toda dedicação e carinho. Agradeço, em especial, ao professor Marcelo Sibaldo por ter aceitado a tarefa de me orientar durante o mestrado.

Agradeço aos amigos que Recife me deu: Raul, os dias difíceis foram mais leves ao teu lado; Jamillys, você é um exemplo de força e calma. Agradeço, também, a Ítalo e João pelo carinho. Grata, também, a todos(as) os(as) colegas do PPGL – Programa de Pós-Graduação em Letras da UFPE – pelas trocas constantes de conhecimento.

Agradeço a todos os informantes e colaboradores desta pesquisa, esta dissertação é sobre vocês e para vocês.

Agradeço ao presidente Luís Inácio Lula da Silva, por ter me permitido sonhar. Graças ao investimento em educação pública e de qualidade, eu – sertaneja – serei mestra em Linguística pela Universidade Federal de Pernambuco.

Por fim, agradeço a mim, pela força e dedicação de sempre, por cada lágrima derramada no começo, por não ter desistido de mim e por ter ousado viver meu sonho!

“(...) Meu fado é o de não saber quase tudo. Sobre o nada tenho profundidades.”  
(BARROS, 2015, p.125)

## RESUMO

A pesquisa em tela filia-se aos estudos da Sociolinguística Variacionista e elege a variação na realização do artigo definido diante de antropônimos como objeto de investigação. Nessa esteira de pensamento, o nosso objetivo geral é apresentar um estudo descritivo e quantitativo sobre o licenciamento do artigo definido em sintagmas antroponímicos no português falado na cidade do Recife. De maneira adjacente, verificamos se tal fenômeno sofre influência de Variáveis Linguísticas e Extralinguísticas na comunidade de fala. De modo amplo, nossa hipótese geral para esta pesquisa é de que, assim como propõem alguns estudos desenvolvidos no estado (PEREIRA (2017) e OLIVEIRA (2018)), Recife tem preferência pela variante sem artigo. Assim sendo, em termos metodológicos, foram analisados dados oriundos de 12 informantes, todos nativos da capital pernambucana. Além disso, para ser informante desta pesquisa, o indivíduo deveria ter o português brasileiro como língua materna. A estratificação dos informantes levou em consideração as variáveis sociais *Sexo* e *Faixa etária*. Todo o respaldo necessário para a coleta e a publicação dos dados foi assegurado mediante submissão e aprovação do Comitê de Ética em Pesquisa (CEP) da Universidade Federal de Pernambuco (UFPE). No que concerne às demandas linguísticas, consideramos as seguintes variáveis: Tipo de informação; Contexto preposicionado; Referência do antropônimo; Função Sintática; *Status* informacional e Tipologia do antropônimo. Os dados observados nesta pesquisa dão conta de que a capital pernambucana, de fato, tem preferência pela ausência do artigo definido em sintagmas antroponímicos. Isto é, em termos percentuais, 63.6% (467/734) das ocorrências registradas foram licenciadas sem determinantes; ao passo que em 36,4% (267/734) das ocorrências, os entrevistados optaram pela presença do artigo definido. De acordo com os resultados fornecidos pelo programa *GoldVarbX*, ferramenta analítica utilizada na quantificação dos dados percentuais, percebemos que as variáveis extralinguísticas *Sexo*, *Faixa etária* e as variáveis linguísticas *Tipo de informação*, *Referência do antropônimo* e *tipologia do antropônimo* são as principais condicionadoras da variação na comunidade de fala.

**Palavras-chave:** Artigo definido; Antropônimos; Sociolinguística Variacionista; Recife.



## ABSTRACT

This research is affiliated with studies on Variationist Sociolinguistics and elects the variation of definite article before anthroponyms as the object of study. In this line of thought, our goal is to present a descriptive and quantitative study about the licensing of definite article in anthroponymous phrases in Brazilian Portuguese spoken in the city of Recife. In an adjoining manner, we want to verify if said phenomenon is influenced by Extralinguistic and Linguistic Variables in the speech community. In a broader sense, our general hypothesis for this research is that, as it is proposed by some studies developed in the state (PEREIRA (2017) and OLIVEIRA (2018)), Recife has a preference for the variable without the article. Therefore, in methodological terms, we analyzed data from 12 informants, all of which are native from the capital of Pernambuco. Besides that, to be an informant of this research, the individuals needed to have Brazilian Portuguese as their first language. The separation of the informants took into account the social variables of *Sex* and *Age levels*. The data collection and publication were altogether backed through submission to and approval of the Research of Ethics Committee of the Federal University of Pernambuco. Regarding the linguistic demands, the following variables: Information type; Prepositional context; Anthroponymous reference; Syntactic function; Informational status and Anthroponymous typology. The data observed in this research confirm that the capital of Pernambuco has, in fact, a preference for the absence of the definite article in anthroponymous phrases. That is, on percentage, 63.6% (467/734) of registered occurrences were licensed without determinants; while in 36.4% (267/734) of the occurrences, the interviewed opted for the presence of the definite article. According to the official results provided by the GoldVarbX program, an analytical tool used on the quantification of percentage data, it was seen that the Extralinguistic Variables *Sex*, *Age levels* and the Linguistic Variables: *Information type*, *Anthroponymous reference* and *Anthroponymous typology* are the main conditioners of variation in the speech community.

**Keywords:** Defined article; Anthroponyms; Variationist Sociolinguistics; Recife.

## LISTA DE ILUSTRAÇÕES

Esquema 1 - Modelo para identificar construções denominativas.....	27
Figura 1 - Localização da Ilha de Martha's Vineyard no mapa.....	47
Esquema 2 - Variável linguística dependente e variantes possíveis para o fenômeno.....	57
Esquema 3 - Critérios de inclusão e exclusão de informantes.....	71
Gráfico 1 - Panorama geral de realização e não realização do artigo definido diante de antropônimos.....	88
Esquema 4 - Variáveis e fatores favoráveis à marcação do determinante em Recife.....	108
Esquema 5 - Construções denominativas retiradas do nosso banco de dados.....	124

## LISTA DE QUADROS

Quadro 1 - Definição dos tipos de antropônimos conforme Amaral e Seide (2020).....	19
Quadro 2 - Realização do artigo definido diante de antropônimos e diante de pronomes possessivos, apresentada no trabalho de Callou & Silva (1997).....	35
Quadro 3 - Realização do artigo definido diante de antropônimos de acordo com a origem geográfica.....	36
Quadro 4 - Centralização dos ditongos [ay] e [aw] por grupos faixa etária.....	48
Quadro 5 - Estratificação dos informantes.....	70
Quadro 6 - Sistema de codificação para a qualificação.....	77
Quadro 7 - Variáveis e fatores extralinguísticos.....	80
Quadro 8 - Variáveis e fatores linguísticos.....	81
Quadro 9 - Resultado da variável Sexo do informante.....	90
Quadro 10 - Resultado da variável Sexo do informante.....	95
Quadro 11 - Resultado da variável Tipo de informação.....	98
Quadro 12 - Resultado da variável Referência do antropônimo.....	103
Quadro 13 - Resultado da variável Tipologia dos antropônimos.....	106
Quadro 14 - Resultado da variável não significativa Contexto preposicionado.....	110
Quadro 15 - Resultado da variável não significativa Função sintática.....	112
Quadro 16 - Resultado da variável não significativa <i>Status</i> informacional.....	116
Quadro 17 - Cruzamento de dados entre as variáveis Tipo de informação e Referência dos antropônimos.....	118
Quadro 18 - Cruzamento de dados entre as variáveis Tipo de informação e Tipologia dos antropônimos.....	119
Quadro 19 - Cruzamento de dados entre as variáveis extralinguísticas Sexo e Faixa etária.....	121

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>14</b>
<b>2</b>	<b>O OBJETO DE ESTUDO: ARTIGO DEFINIDO + ANTROPÔNIMOS NO PB.....</b>	<b>17</b>
2.1	A ANTROPONÍMIA BRASILEIRA.....	17
2.1.1	Aspectos estruturais da antroponímia em PB.....	22
2.2	O USO VARIÁVEL DO ARTIGO DEFINIDO: O QUE DIZEM ALGUMAS PESQUISAS SOCIOLINGÜÍSTICAS.....	33
2.2.1	Callou e Silva (1997).....	34
2.2.2	Campos Júnior (2013).....	35
2.2.3	Sedrins, Pereira e Siqueira (2015).....	37
2.2.4	Pereira (2017).....	39
2.2.5	Oliveira (2018).....	41
<b>3</b>	<b>BASE EPISTÊMICA: A SOCIOLINGÜÍSTICA VARIACIONISTA.....</b>	<b>45</b>
3.1	WILLIAM LABOV E A SOCIOLINGÜÍSTICA VARIACIONISTA.....	46
3.2	A SOCIOLINGÜÍSTICA NO BRASIL: BREVES INCURSÕES.....	52
3.3	SOCIOLINGÜÍSTICA VARIACIONISTA: CONCEITOS BÁSICOS.....	55
<b>4</b>	<b>OS PROCESSOS METODOLÓGICOS.....</b>	<b>64</b>
4.1	A COMUNIDADE DE FALA.....	65
4.2	OS PROCEDIMENTOS ÉTICOS: SUBMISSÃO DO PROJETO DE PESQUISA AO COMITÊ DE ÉTICA EM PESQUISA.....	67
4.3	A SELEÇÃO DOS INFORMANTES.....	68
4.3.1	Riscos.....	72
4.4	A COLETA DOS DADOS.....	73
4.4.1	Armazenamento dos dados coletados.....	75
4.5	TRANSCRIÇÃO E CODIFICAÇÃO DOS CONTEXTOS SELECIONADOS.....	75
4.6	VARIÁVEIS SELECIONADAS.....	79
4.6.1	Variáveis extralingüísticas.....	79
4.6.2	Variáveis lingüísticas.....	81
4.7	GOLDVARBX: PROCESSAMENTO E RODAGEM DOS DADOS.....	83
<b>5</b>	<b>DESCRIÇÃO E ANÁLISE QUANTITATIVA DOS DADOS.....</b>	<b>85</b>

5.1	VARIÁVEL LINGUÍSTICA DEPENDENTE: RESULTADOS GLOBAIS DE APLICAÇÃO E NÃO APLICAÇÃO DO ARTIGO DEFINIDO EM SINTAGMAS ANTROPONÍMICOS NA CIDADE DO RECIFE.....	86
5.2	VARIÁVEIS APONTADAS COMO SIGNIFICATIVAS PELO PROGRAMA <i>GOLDVARBX</i> .....	89
5.2.1	<b>Sexo.....</b>	<b>90</b>
5.2.2	<b>Faixa etária.....</b>	<b>94</b>
5.2.3	<b>Tipo de informação.....</b>	<b>97</b>
5.2.4	<b>Referência dos antropônimos.....</b>	<b>101</b>
5.2.5	<b>Tipologia dos antropônimos.....</b>	<b>104</b>
5.3	VARIÁVEIS APONTADAS COMO NÃO SIGNIFICATIVAS PELO PROGRAMA <i>GOLDVARBX</i> .....	110
5.3.1	<b>Contexto preposicionado.....</b>	<b>110</b>
5.3.2	<b>Função sintática.....</b>	<b>112</b>
5.3.3	<b>Status informacional.....</b>	<b>115</b>
5.4	CRUZAMENTO DE ALGUMAS VARIÁVEIS SIGNIFICATIVAS.....	117
5.4.1	<b>Tipo de informação <i>versus</i> Referência dos antropônimos.....</b>	<b>117</b>
5.4.2	<b>Tipo de informação <i>versus</i> Tipologia dos antropônimos.....</b>	<b>119</b>
5.4.3	<b>Sexo <i>versus</i> Faixa etária.....</b>	<b>120</b>
5.5	ALGUNS ASPECTOS ACERCA DA ANTROPONÍMIA NA CAPITAL PERNAMBUCANA.....	122
5.5.1	<b>Antropônimos sem artigos: caso de não-opcionalidade em que o antropônimo é utilizado como etiqueta linguística.....</b>	<b>123</b>
5.5.2	<b>Antropônimos com artigos: nomes próprios de pessoas individualizados por meio de um modificador têm uso obrigatório de artigos definidos?.....</b>	<b>125</b>
5.5.3	<b>Antropônimos com artigos: nomes próprios de pessoas deslocados à esquerda das sentenças favorecem a aplicação do artigo definido?.....</b>	<b>126</b>
6	<b>CONCLUSÕES.....</b>	<b>129</b>
	<b>REFERÊNCIAS.....</b>	<b>132</b>
	<b>APÊNDICE A – TERMO DE ASSENTIMENTO LIVRE E ESCLARECIDO.....</b>	<b>134</b>

<b>APÊNDICE B – TERMOS DE CONSENTIMENTO LIVRE E ESCLARECIDO.....</b>	<b>136</b>
<b>APÊNDICE C – TERMO DE COMPROMISSO E CONFIDENCIALIDADE.....</b>	<b>140</b>
<b>APÊNDICE D – FICHA SOCIAL.....</b>	<b>141</b>
<b>APÊNDICE E – ROTEIRO DE PERGUNTAS.....</b>	<b>143</b>
<b>APÊNDICE F – CARTA DE ANUÊNCIA.....</b>	<b>144</b>

## 1 INTRODUÇÃO

A pesquisa em tela, filiada aos estudos sociolinguísticos, tem como principal objetivo apresentar um trabalho quantitativo e descritivo acerca da variação do artigo definido anteposto antropônimos em dados de fala na capital pernambucana: Recife. De maneira adjacente, esta pesquisa almeja examinar, ainda, a frequência de uso do artigo definido diante de nomes próprios de pessoas na comunidade de fala; observando, sobretudo, se há condicionamento por variáveis linguísticas e/ou extralinguísticas.

O arcabouço teórico-metodológico adotado no curso desta investigação apoia-se, basilarmente, nos principais estudos publicados acerca da Sociolinguística Variacionista. De tal modo, nos respaldaremos nos postulados de Labov (2008) [1972], Tarallo (2007), Freitag (2014, 2016), Mollica (2015) e Coelho *et al* (2016), a fim de circunstanciar nossa base epistêmica.

Como elegemos a variação do artigo definido diante de antropônimos como o nosso objeto de análise nesta investigação, nos apoiaremos, também, em alguns trabalhos sociolinguísticos que versam acerca dos sintagmas antroponímicos<sup>1</sup> no Português brasileiro (PB), a saber: Callou & Silva (1997), Campos Jr. (2011), Sedrins, Pereira e Siqueira (2015), Pereira (2017) e Oliveira (2018).

Nessa esteira de pensamento, observaremos também o que pontuam as áreas do conhecimento que se propõem a estudar os nomes próprios, de modo amplo: Onomástica; e, de modo mais específico, os nomes próprios de pessoas: Antroponomástica.

É correto ressaltarmos que muitas pesquisas científicas adotam os antropônimos como objeto de análise, focalizando, conforme já mencionamos, a natureza variável de alguns fenômenos ligados à antroponímia, mas, também, elegendo aspectos mais formais como tema de suas investigações. Assim sendo, o fenômeno da variação do artigo definido em sintagmas de natureza antroponímica no Brasil é mote para verificações linguísticas vinculadas não apenas à Sociolinguística, mas também à Dialetologia e, de modo correlato, aos estudos de cunho Gerativista<sup>2</sup>.

---

<sup>1</sup> Nomenclatura disponível em Amaral e Seide (2020), utilizada para designar os sintagmas que têm antropônimos na posição de núcleo, acompanhados, ou não, por artigos definidos.

<sup>2</sup> Sedrins (2017).

Assim, podemos inferir que o fenômeno da variação do artigo definido frente a antropônimos sofre forte influência de diferentes fatores, sejam eles linguísticos e/ou extralinguísticos, conforme demonstram diversos estudos<sup>3</sup>. Não obstante, Amaral e Seide (2020) destacam que, “embora o fenômeno tenha sido foco de vários estudos, desconhece-se como se dá em muitas partes do país” (AMARAL e SEIDE, 2020, p.133), fator que confirma a necessidade de uma descrição mais substancial acerca do falar da capital pernambucana.

Com isso, destacamos a importância de mapeamentos linguísticos que (assim como esta pesquisa) se propõem a descrever como intercorre o fenômeno da antroponímia em regiões específicas do país. Por assim dizer, essa pesquisa se justifica, sobretudo, pela sua contribuição no que diz respeito à criação de um mapa linguístico abrangente acerca do fenômeno da variação no território estadual. Isso porque levaremos em consideração os trabalhos já produzidos no interior do estado<sup>4</sup> e a nossa investigação na cidade do Recife para substanciar a descrição desse fenômeno em Pernambuco.

Dando visibilidade a nossa empreitada metodológica, para execução desta pesquisa, analisaremos a fala de doze informantes, estratificados por meio das variáveis *faixa etária* e *sexo*. Todos os informantes que se voluntariaram a participar desta pesquisa carregavam a particularidade de serem habitantes da comunidade de fala por, pelo menos, 5 anos e terem o português brasileiro (PB) como língua materna. Além disso, por meio do programa computacional GoldVarbX, quantificaremos nossos dados e analisaremos se haverá influências das seguintes variáveis linguísticas observadas: *Contexto preposicionado*, *Função sintática*, *Status informacional*, *Tipo de informação*, *Referência do antropônimo* e *Tipologia do antropônimo*.

A hipótese geral desta pesquisa pressupõe que – com base nos estudos de Callou & Silva (1997), Sedrins, Pereira e Siqueira (2015), Pereira (2017) e Oliveira (2018) – os falantes pernambucanos têm preferência pela não marcação dos determinantes em sintagmas antroponímicos.

Em termos ainda metodológicos, é válido abalizarmos que a pesquisa em tela respeita a todos os pré-requisitos éticos do Comitê de Ética em Pesquisa (CEP) da Universidade Federal de Pernambuco (UFPE). Ou seja, haja vista a necessidade de garantir respaldo legal para o desenvolvimento e a divulgação dos resultados, a realização da presente pesquisa obedecerá às

---

<sup>3</sup> Silva (1998); Callou & Silva (1997); Campos Jr. (2011).

<sup>4</sup> Pereira (2017) e Oliveira (2018).



cláusulas éticas da Resolução 510/16 do Conselho Nacional de Saúde, que regulamenta as pesquisas realizadas com seres humanos.

Assim sendo, para além desta introdução, este trabalho está organizado em mais quatro capítulos. No segundo capítulo, iremos expor algumas considerações acerca do nosso fenômeno analítico; para tanto, nos respaldaremos nas seguintes explanações: considerações sobre a antroponímia brasileira e o que versam alguns estudos de caráter sociolinguístico. No terceiro capítulo, debateremos sobre o alicerce epistêmico que baseia esta pesquisa: a Sociolinguística Variacionista e suas principais publicações no Brasil. No quarto capítulo, apresentaremos detalhadamente os procedimentos metodológicos seguidos no curso desta investigação. No quinto capítulo, por fim, exibiremos os resultados obtidos por meio da utilização do programa computacional GoldVarbX e, dando seguimento às análises, faremos algumas ponderações acerca dos dados percentuais e das variáveis apontadas como significativas e não significativas para a variação; por fim, exibiremos alguns cruzamentos de dados entre as variáveis significativas e algumas considerações sobre a antroponímia na capital pernambucana. Em seguida, apresentaremos nossas considerações finais acerca desta pesquisa.

## 2 O OBJETO DE ESTUDO: ARTIGO DEFINIDO + ANTROPÔNIMOS NO PB

Conforme já mencionamos, o objeto de estudo desta dissertação é o uso variável do artigo definido anteposto a nomes próprios de pessoas. Por isso, é relevante investigarmos quais são as propriedades sintático-semânticas que norteiam o nosso fenômeno analítico. Para tal, nos debruçaremos sobre alguns estudos acerca do comportamento estrutural da antroponímia em língua portuguesa. Em um segundo momento, revisitaremos estudos de caráter sociolinguístico que já investigaram o fenômeno e colocaram em relevo sua natureza variacional no português brasileiro (PB).

### 2.1 A ANTROPONÍMIA BRASILEIRA

É válido apontarmos que as discussões propostas nesta primeira parte da seção serão respaldadas na recente publicação “*Nomes próprios de pessoas: introdução à antroponímia brasileira*”, de Amaral e Seide (2020). A natureza descritiva e analítica dessa obra fomenta inúmeras discussões que são importantes para o curso deste trabalho, haja vista a similaridade entre os nossos objetos de estudo e o aprofundamento teórico apresentado pelos autores.

A área do conhecimento que se ocupa de estudar os nomes próprios é chamada de *Onomástica*. De modo mais amplo, conforme apontam Amaral e Seide (2020), a onomástica se incumbe de investigar os vários aspectos que envolvem os nomes próprios: “gramaticais, etimológicos, sócio-históricos, geográficos” (AMARAL e SEIDE, 2020, p. 32). Nessa linha argumentativa, tal área é um campo de conhecimento que, mesmo autônomo, faz interface com outras disciplinas. Assim, como o objeto investigativo desta pesquisa é analisar o comportamento variável do artigo definido frente aos antropônimos, é válido ressaltarmos que à subárea que contempla esses nomes – no vasto campo na onomástica – dá-se o nome de *Antroponomástica*. Desse modo, nesta subseção iremos discorrer acerca das principais características dos nomes próprios (sobretudo de pessoas) na língua portuguesa brasileira.

Antes disso, é importante destacarmos que, no âmbito internacional, os estudos antroponomásticos reúnem uma quantidade considerável de subáreas que são substanciais para o conhecimento da disciplina e das discussões que apresentaremos nos limites desta seção. De acordo com Amaral e Seide (2020), entre os muitos campos investigativos, destacamos aqueles que conglobam números expressivos de pesquisas, a saber: Antroponomástica ficcional,

Antroponomástica comparada, Antroponomástica aplicada aos estudos de tradução, Política linguística antroponímica, Jurisdição antroponímica e Antroponomástica histórica<sup>5</sup>.

Na concepção de Amaral e Seide (2020), o interesse pela Antroponomástica no Brasil seguiu o curso dos estudos linguísticos, já que, quando comparado a outros países, as pesquisas dessa área são mais recentes no país. Isso se deve ao fato de a pesquisa brasileira, sobretudo em linguística, estar – assim como também pontua Freitag (2016) – fortemente vinculada à criação de programas de pós-graduação em Letras. Para os autores, os primeiros registros de pesquisas onomásticas surgiram em meados século XX e se debruçaram, principalmente, sobre as propriedades dos nomes próprios.

Após um detalhado apanhado histórico, Amaral e Seide (2020) fazem uma descrição quantificada de grande parte dos estudos antroponímicos já realizados no Brasil (incluindo os de caráter sociolinguístico que citaremos na seção 1.3). Mesmo considerando a importância desse trabalho de fôlego efetivado pelos autores, assumimos que, a partir daqui, nosso principal interesse na obra se concentrará nas descrições tipológicas e sintáticas articuladas pelos autores. De tal modo, ampliaremos nossas discussões acerca dos antropônimos, com vistas a circunstanciar as análises que apresentaremos no capítulo 5.

Sobre as principais características dos nomes próprios, categoria mais ampla que contempla os antropônimos, Amaral e Seide (2020) assumem quatro propriedades inerentes a esse conjunto:

- a) Possibilitam a identificação direta de um referente único em um universo de conhecimento compartilhado por emissor e receptor; b) Possuem capacidade de referir, independentemente da presença de determinante; c) Não apresentam traços semânticos identificadores de classe e d) São grafados com maiúscula inicial. (AMARAL E SEIDE, 2020, p.59)

Entre as questões propostas pelos autores, uma delas parece importante para a discussão do nosso objeto investigativo. Ora, se bem é verdade que os nomes próprios têm a capacidade

---

<sup>5</sup> De acordo com Amaral e Seide (2020) cada subárea destacada compreende um campo investigativo específico que vale ser observado. *Antroponomástica ficcional* é uma área que compreende pesquisa cujo objetivo é estudar os nomes de personagens fictícios sejam eles do cinema, televisão, teatro ou séries. A *antroponomástica comparada* se incumbe de comparar um ou mais sistemas antroponímicos. Antroponomástica aplicada aos estudos de tradução se encarrega de observar como se dão os processos de tradução relacionados aos estudos de antropônimos. A Política linguística antroponímica e jurisdição antroponímica diz respeito às regras e leis que regem os nomes (grafia, pronúncia etc.) e seus usos. Já a Antroponomástica histórica congloba, entre outras questões, pesquisas que abrangem o estudo histórico da evolução dos nomes. Para saber mais acerca de cada uma das áreas elencadas, sugerimos a leitura do capítulo (1) da obra “*Nomes próprios de pessoas: introdução à antroponímia brasileira*”, de Amaral e Seide (2020).

de referir independentemente da presença de um determinante, podemos supor que o uso facultativo do artigo definido frente a antropônimos (um tipo de nome que compõe a categoria dos nomes próprios) confirma tal assertiva. No entanto, parece pertinente destacarmos que a natureza desse nome também deve ser destacada, dado que – conforme verificaremos na subseção 2.1.1 – alguns contextos parecem exigir a presença marcada do artigo definido no português falado no Brasil. Por esse motivo, consideraremos tanto a natureza sintática dos antropônimos (com ou sem determinantes) quanto sua tipologia.

À vista disso, interessa-nos caracterizar, primeiramente, os tipos de antropônimos que constituem tal categoria, para observarmos posteriormente seus comportamentos sintáticos. Conforme apontam Amaral e Seide (2020, p. 72), “os antropônimos não constituem uma subclasse homogênea, fazem parte desse conjunto de antropônimos: prenomes, sobrenomes, apelidos e vários outros elementos”. Ou seja, dada à amplitude dos tipos de antropônimos que podemos encontrar no português brasileiro, iremos utilizar a tipologia proposta pelos autores como variável linguística independente para a quantificação e posterior descrição dos nossos dados.

Nessa altura, cabe ressaltarmos que o nosso objetivo com essa demarcação tipológica é verificar se a natureza dos antropônimos selecionados em nosso *corpus* irá desempenhar (ou não) um papel relevante para uso variável do artigo definido em nossas análises. Nessa configuração, o quadro (1), a seguir, proposto por Amaral e Seide (2020), exemplifica os principais tipos de antropônimos recorrentes no PB e suas características básicas:

**Quadro 1** - Definição dos tipos de antropônimos conforme Amaral e Seide (2020)

	<b>Tipo</b>	<b>Definição e principais características</b>
<b>1</b>	<b>Prenome</b>	Antropônimo que distingue o indivíduo dentro dos grupos sociais de maior intimidade. Antecede o sobrenome e pode ser simples, composto ou justaposto.
<b>2</b>	<b>Sobrenome</b>	Antropônimo que identifica o pertencimento do indivíduo a uma família. Geralmente provêm dos genitores e sucede o prenome.
<b>3</b>	<b>Agnome</b>	Antropônimo que indica uma relação de parentesco com outro indivíduo, via de regra, por via patrilinear.
<b>4</b>	<b>Apelido (ou alcunha, ou cognome)</b>	Antropônimo que se atribui a um indivíduo geralmente por outra pessoa e que costuma aludir a uma característica física ou intelectual ou ainda a um fato ou comportamento social.

5	<b>Hipocorístico</b>	Antropônimo formado a partir de uma alteração morfológica (abreviação, diminutivo, aumentativo etc.) de outro antropônimo. Geralmente criado em ambientes de maior intimidade.
6	<b>Pseudônimo</b>	Antropônimo empregado por um indivíduo em lugar do seu nome civil e escolhido pelo próprio portador do nome próprio.
7	<b>Codinome</b>	Antropônimo empregado para ocultar a identidade de um indivíduo. Pode ser escolhido pelo próprio portador ou por outrem e frequentemente possui traço negativo.
8	<b>Heterônimo</b>	Antropônimo atribuído a um indivíduo fictício, criado pelo portador de outro antropônimo.
9	<b>Nome artístico (e nome de palco)</b>	Antropônimo empregado por um indivíduo em lugar do seu nome civil e pelo qual se faz conhecido em sua atividade profissional, especialmente em áreas como música, cinema, teatro, televisão e afins.
10	<b>Nome de guerra</b>	Antropônimo empregado como substituto do nome civil em ambientes restritos, especialmente no meio militar, na maçonaria, na prostituição e no crime organizado.
11	<b>Nome religioso</b>	Antropônimo empregado por membros de comunidades religiosas em lugar do nome civil.
12	<b>Nome social</b>	Antropônimo pelo qual a pessoa, especialmente transexual e travesti, se identifica e é reconhecida socialmente.
13	<b>Nome de urna</b>	Antropônimo escolhido pelo candidato às eleições proporcionais para registrar-se na Justiça Eleitoral.
14	<b>Nome parlamentar</b>	Antropônimo escolhido pelo indivíduo eleito a cargo legislativo para ser usado em documentos oficiais da casa legislativa.

**Fonte:** Amaral e Seide (2020, p.100)

Nessa linha argumentativa, a tipologia apresentada pelos autores nos fornece subsídios para delimitarmos em nossa análise – inclusive na codificação dos nossos dados – se (e quais) os tipos específicos de antropônimos são mais sensíveis ao uso do artigo e/ou os que parecem anular sua ocorrência. Desse modo, com base nas descrições propostas pelos autores, as sentenças retiradas do nosso *corpus* de análise<sup>6</sup>, a seguir, exemplificam cada um dos tipos de antropônimos definidos acima:

- Prenome
  - (1) **Mariane** sempre foi uma amiga e tanto.
- Sobrenome
  - (2) **O Queiroz** ninguém viu, né?!

---

<sup>6</sup> Exemplos extraídos do nosso *corpus* analítico, com dados do português falado na cidade do Recife.

- Agnome
  - (3) **O Tony Júnior** é uma alegoria do Bolsonaro.
- Apelido (ou alcunha, ou cognome)<sup>7</sup>
  - (4) **Lilia** fazia isso por mim.
- Hipocorístico
  - (5) **Brunão** é meu amigo.
- Pseudônimo
  - (6) Eu gosto dos livros do **Alan Kardec**.
- Codinome
  - (7) **O miliciano**, já que está gravando e não vou me comprometer depois (...)
- Heterônimo
  - (8) Como diz naquela poesia bem triste de **Ricardo Reis**.
- Nome artístico (e nome de palco)
  - (9) **A Gloria Groove**, também conhecida como a dona do mundo.
- Nome religioso
  - (10) Eu gosto **do Julio Lancellotti**.
- Nome social
  - (11) **A Pablo Vittar** é resistência para mim.
- Nome de urna
  - (12) **Geraldo Júlio** não faz nada aqui (...)
- Nome parlamentar
  - (13) **O Jean Wyllys** devia ter cuspidado mais (...)

É válido justificarmos que o único tipo de antropônimo que não foi localizado nos nossos dados, diz respeito à categoria utilizada para substituir o nome civil em ambientes específicos como, por exemplo, na prostituição e/ou no meio militar: o nome de guerra. Desse modo, podemos verificar que as nuances que abrangem a categoria dos antropônimos são bastante amplas e correspondem a uma gama representativa dos nomes próprios de pessoas encontrados no português falado no Brasil; prova disso são os exemplos coletados em nossa pesquisa que contemplam grande parte das categorias sugeridas por Amaral e Seide (2020).

---

<sup>7</sup> É válido ressaltarmos que, no ato da coleta de dados, a inquiridora responsável pelo andamento da pesquisa questionou a natureza dos nomes que poderiam ser interpretados enquanto apelidos. Após os participantes confirmarem (ou não) essas informações, os antropônimos tidos como apelidos foram identificados, tipologicamente, já durante as entrevistas.

É importante ressaltarmos que, ao longo de nossas análises, percebemos que as definições tipológicas de “Nome artístico” e “Pseudônimo” propostas pelos autores dão margem para uma possível relação entre as principais características de ambas. De tal maneira, ressaltamos que a problematização e/ou atualização conceitual de tais categorias não foram feitas no transcorrer desta pesquisa. Todavia, é importante explicar que consideramos como “Nome artístico/Nome de palco” todos aqueles antropônimos que referenciavam uma pessoa, em geral, famosa; conhecida, sobretudo, por sua carreira profissional. Na categoria que indica os “Pseudônimos”, por sua vez, tivemos uma ocorrência que diz respeito ao exemplo (6) supramencionado.

Já no que diz respeito à qualificação tipológica que elege os Nomes religiosos, destacamos a presença da entidade (o papa, por exemplo) antecedendo o antropônimo em si. Essa tipologia também não foi problematizada no curso desta dissertação e, por isso, foi desconsiderada na nossa codificação e análise dos dados. Todavia, como outros exemplos dessa categoria se mostraram recorrentes em nossos dados (como é o caso da construção “O presidente Lula”), achamos prudente destacarmos a importância desse tipo de construção para trabalhos futuros.

Voltaremos a falar da tipologia dos antropônimos no capítulo 4, quando discutiremos, mais especificamente, sobre as variáveis independentes utilizadas nas análises dos nossos dados. Na subseção seguinte, iremos observar alguns dos aspectos sintáticos mais relevantes da antroponímia em língua portuguesa, haja vista que esse é o nosso ambiente de variação do artigo.

### **2.1.1 Aspectos estruturais da antroponímia em PB**

Os estudos de Müller e Negrão (1988) são substanciais para a observação do comportamento dos nomes próprios de pessoas no PB. De acordo com as autoras, o nome próprio carrega a particularidade de ser semanticamente completo, além de denotar “um indivíduo específico no mundo real ou de algum mundo possível” (MÜLLER E NEGRÃO, 1988, p. 530). As autoras expõem, também, a primazia de que a estrutura [artigo definido + nome próprio] é redundante na língua, uma vez que o nome próprio já expressa um indivíduo específico no mundo, ou seja, com ampla margem de definição.

Sintaticamente, as autoras destacam que a combinação de um nome próprio com um sintagma verbal já é material suficiente para a formação de uma sentença completamente

gramatical<sup>8</sup> no português brasileiro. Sem se distanciarem do critério sintático, as autoras seguem argumentando que o uso do artigo diante de nomes próprios parece independer da posição sintática ocupada pelo sintagma nominal<sup>9</sup>. Assim, elas argumentam que, nesses moldes, a presença/ausência do artigo diante de nomes próprios seria equivalente; “dizendo-nos apenas que seus falantes procedem de regiões distintas” (MULLER E NEGRÃO, 1988, p. 531).

Apesar disso, Sedrins (2017) parece expandir essa afirmação – de uso facultativo no PB – ao acrescentar que “apenas quando o nome próprio se refere a um indivíduo único é que podemos ter a opcionalidade na realização do artigo definido” (SEDRINS, 2017, p. 7). Isto é, parece prudente delimitarmos, conforme o autor, que a variação do artigo diante de antropônimos é permitida no PB somente quando estamos nos referindo a um sujeito único e, por isso, não ocorre de maneira irrestrita. As sentenças<sup>10</sup> expressas em (14), a seguir, exemplificam o que é indicado pelo autor<sup>11</sup>:

- (14) a. (O) **João** beijou a menina.  
 b. A menina odeia (o) **João**.  
 c. A menina não gosta **de/do João**.

A delimitação proposta pelo autor introduz, nesta discussão, a ideia de não-opcionalidade que já foi observada também em outros estudos (MÜLLER E NEGRÃO, 1988; SEDRINS, 2017; OLIVEIRA, 2018; AMARAL E SEIDE, 2020). Ora, se é bem verdade que a variação só é verificada em PB quando o artigo definido se refere a um “único indivíduo”, conforme aponta Sedrins (2017, p.7), é válido discutirmos o(s) contexto(s) em que a não-opcionalidade pode também ser conferida em dados do português brasileiro. Isto é, os casos em que o uso variável de artigos definidos frente a nomes próprios de pessoas não é permitido na língua<sup>12</sup>.

---

<sup>8</sup> Como sugere o gerativismo Chomskyano, o termo gramatical refere-se, aqui, ao fato de determinada sentença ser gerada no sistema linguístico (inato) dos falantes da língua portuguesa. Não estamos nos referindo às gramáticas normativas da língua, mas sim, à capacidade que os falantes têm de conhecer a agramaticalidade (estruturas que não são geradas pelo seu sistema linguístico) e a gramaticalidade (estruturas que não são geradas pelo seu sistema linguístico) de determinadas estruturas.

<sup>9</sup> Ver o exemplo (14), a seguir, proposto por Sedrins (2017).

<sup>10</sup> Exemplos extraídos de Sedrins (2017, p. 07).

<sup>11</sup> Como é possível observar, os exemplos de Sedrins (2017) indicam que o uso variável do artigo definido pode acontecer em posições sintáticas distintas no PB. Assim, seu uso pode variar na posição de sujeito (cf.14a), na posição intrassentencial de objeto direto (cf.14b) e na posição objeto indireto (cf.14c).

<sup>12</sup> Nos ocuparemos desses casos nas próximas subseções dessa seção.



Amaral e Seide (2020) trazem uma série de ponderações importantes que merecem ser discutidas neste trabalho, sobretudo, as que se referem à organização sintática dos antropônimos, ambiente analítico de nossa pesquisa. Muitas das considerações que vamos expor tentam alinhar e – quando necessário – contrapor as proposições dos autores citados. Mais do que isso, iremos nos valer dos exemplos extraídos do nosso *corpus* para ilustrar as aceções defendidas pelos autores.

Ainda sobre os aspectos morfossintáticos dos antropônimos, Amaral e Seide (2020) pontuam que a classe dos substantivos (categoria gramatical dos antropônimos) pode ser individualizada em PB pelas propriedades morfológicas que apresenta e pelas funções sintáticas que ocupa. Na concepção dos autores, é correto afirmar que os substantivos geralmente possuem algumas características específicas nas sentenças: “i) ocupam a posição de núcleo dos sintagmas nominais; ii) podem (ou não) ser precedidos por determinantes; iii) são acompanhados (à esquerda ou à direita) por adjetivos” (AMARAL E SEIDA, 2020, p. 114). Apesar disso, os autores atentam para o fato de que, a depender do tipo de determinante que antepõe o antropônimo, o significado da sentença pode ser alterado; em outras palavras, “pode acarretar mudanças na interpretação da sentença e mesmo favorecer uma leitura de alteração de classe” (AMARAL E SEIDA, 2020, p. 114). Para ilustrar suas prerrogativas, os autores apresentam o exemplo, a seguir:

(15) Ele viu **um** / **meu** / **aquele** / **Pedro**

Conforme Amaral e Seide (2020), a sentença (15) representa construções com o uso de determinantes que são recorrentes no PB. Isto é, os determinantes *um*, *meu* e *aquele*, parecem pressupor a existência de outros indivíduos de nome *Pedro*, de modo que a interpretação pode indicar uma leitura semelhante nos três usos: “a seleção de um membro de um mesmo grupo, aquele com indivíduos que têm o mesmo nome próprio” (AMARAL E SEIDE, 2020, p.14). Para os autores, é prudente observarmos, ainda, que não apenas o tipo de antropônimo é importante na sentença (15), mas também os determinantes selecionados e as leituras que eles podem implicar. Todavia, é válido destacarmos que nosso foco de observação é o uso de antropônimos licenciados com (ou sem) artigos definidos. Por essa razão, delimitaremos nossa investigação para essa categoria específica de determinantes.

Assim sendo, é justo investigarmos, ancorados às premissas expostas pelos teóricos, quais são os tipos de elementos que ocupam os sintagmas nominais em que os antropônimos são licenciados como núcleo – o que Amaral e Seide (2020) preferem chamar de *sintagmas*

*antroponímicos*. Desse modo, observaremos, neste capítulo, dois tipos de configurações para esse tipo de sintagma no PB: sintagmas antroponímicos com artigos definidos e sintagmas antroponímicos sem artigos definidos.

A seguir, discorreremos acerca de algumas ocorrências em que o uso do artigo definido diante de antropônimos parece ser, conforme apontam alguns teóricos, anulado em PB.

### • Antropônimos sem artigos

Permeando por vias sintáticas, Müller e Negrão (1988) argumentam que há contextos no PB em que a presença/ausência do artigo pode se alternar e que, em dadas situações, apenas o licenciamento sem o determinante é permitido. Vejamos os exemplos utilizados pelas autoras:

- (16) a. Esta é Rafaela  
b. Esta é a Rafaela

- (17) a. Ela se chama Rafaela  
b. \*Ela se chama a Rafaela

- (18) a. O nome Rafaela é lindo  
b. \*O nome a Rafaela é lindo

As autoras inferem que em (16a) o nome Rafaela não é referencial, ou seja, o nome próprio – da maneira que está colocado na sentença – é equivalente a um batismo ou, ainda, a uma predicação, haja vista que o indivíduo de quem se fala possui a qualidade de se chamar Rafaela. Já em (16b), o indivíduo denotado pelo pronome demonstrativo “esta” é o mesmo indivíduo denotado pelo nome Rafaela; isto é, o uso do artigo, de acordo com as autoras, permite que o ouvinte da oração seja capaz de individualizar o ser referido.

Nos casos (17) e (18), Müller e Negrão (1988) afirmam que o verbo e o substantivo que antepõem os antropônimos induzem a interpretação dele enquanto uma entidade linguística: “estamos aqui falando não do indivíduo, mas de sua etiqueta linguística” (MÜLLER E NEGRÃO, 1998, p. 534); isto é, não estamos nos referindo às pessoas físicas, pessoas tangíveis em matéria e espécie, mas sim aos seus nomes (suas marcas sociais). Dessa maneira, as autoras reafirmam a hipótese de que a estrutura [artigo definido + nome próprio] é utilizada apenas para referenciar o indivíduo designado pelo nome e não o nome em si. Em síntese, de acordo com

as autoras, quando o nome próprio faz referência a entidades linguísticas, a estrutura [artigo + nome próprio] parece ser agramatical no português falado no Brasil (cf. 17 b. e 18 b.).

Oliveira (2018), em dados do português falado no interior do estado de Pernambuco<sup>13</sup>, aponta alguns desses contextos em que parece não haver opcionalidade para o uso variável do artigo definido, uma vez que o licenciamento com o artigo definido resultaria em uma leitura aparentemente agramatical no dialeto estudado. As sentenças<sup>14</sup>, a seguir, ilustram melhor essa afirmação:

- (19) a. Conheço ela **por Aninha**, né?  
b.\*Conheço ela **por a Aninha**, né?

- (20) a. Meu professor de matemática, chamado **Geilson**.  
b.\*Meu professor de matemática, chamado o **Geilson**.

Como podemos verificar, as ocorrências exemplificam de maneira significativa o que é discutido pelos autores: os antropônimos – da forma que estão expostos nas sentenças (19b) e (20b) – não permitem o licenciamento anteposto do artigo definido. Ao analisarmos as sentenças de maneira semântica, podemos inferir que, nos dois exemplos, os antropônimos não tangenciam as pessoas citadas no discurso, mas sim os seus nomes (o que podemos chamar, em consenso com Müller e Negrão (1998), de suas entidades linguísticas<sup>15</sup>).

Vejamos mais um exemplo fornecido por Oliveira (2018)<sup>16</sup> que corrobora com a alegação de Müller e Negrão (1988):

- (21) a. **Bernardo** é um bom nome.  
b.\***O Bernardo** não é um bom nome

É possível perceber que, nas ocorrências exemplificadas acima (21a), o informante parece estar – ainda que intuitivamente – nos apresentando o antropônimo (um nome próprio de pessoa, mas não de alguém específico), por isso, a relevância discursiva parece sobrelevar

<sup>13</sup> É importante ressaltarmos que as sentenças (19b) e (20b) não pertencem aos dados de língua falada no interior do estado. Conforme é possível verificarmos, as sentenças foram utilizadas para expressar a agramaticalidade desse tipo de construção.

<sup>14</sup> Dados extraídos do Oliveira (2018, p. 60).

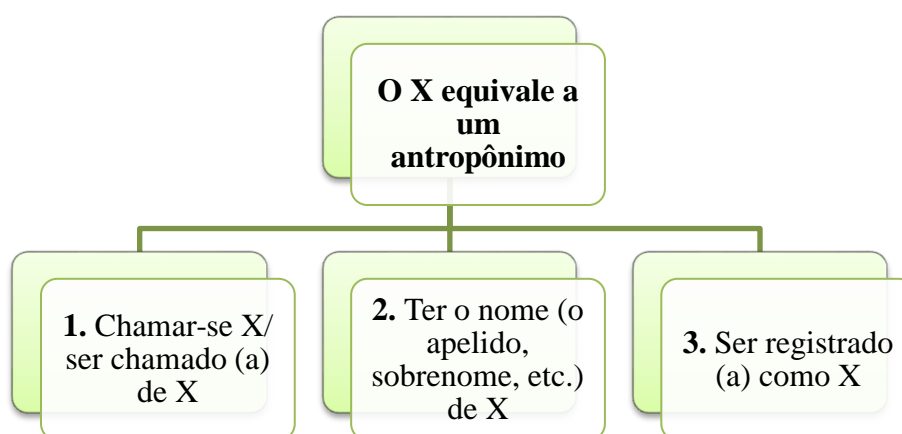
<sup>15</sup> Müller e Negrão (1988) afirmam que o nome próprio pode ser utilizado como uma “entidade linguística”, quando se refere ao nome do indivíduo e não ao indivíduo em si. Ou seja, o nome próprio é empregado não para se referir a um indivíduo, mas ao seu nome.

<sup>16</sup> Dados extraídos do Oliveira (2018, p. 62).

mais o nome do indivíduo do que o indivíduo a quem o nome pertence. Desse modo, em consonância com as autoras, os dados de Oliveira (2018) também sugerem que, nesses contextos específicos, a variação parece não ser permitida no PB, uma vez que o sintagma antroponímico anula o uso do determinante nesse contexto específico – artigo definido (cf. 21b).

Amaral e Seide (2020) também chamam atenção para esse contexto de antropônimos sem determinantes. No entanto, a atualização proposta pelos autores (2020) – ao que já foi observado por Müller e Negrão (1988) e Oliveira (2018) – diz respeito à nomenclatura dada pelos autores às ocorrências desse tipo: construções denominativas. Trata-se, por sua vez, das mesmas ocorrências não referenciais observadas acima, aquelas em que o antropônimo indica o nome que é dado ao indivíduo, e não o ser em si. Ao resgatarmos tais considerações, chamamos atenção para o modelo prático – proposto pelos autores – para identificar os casos em que o sintagma antroponímico é formado apenas pelo nome próprio de pessoa, anulando o artigo definido. No modelo (2), a seguir, organizamos em forma de esquema hierárquico o que é pontuado pelos autores:

**Esquema 1** - Modelo para identificar construções denominativas



**Fonte:** Amaral e Seide (2020, p. 114)

Desse modo, assumindo a designação proposta por Amaral e Seide (2020), vamos considerar esses casos como sendo **construções denominativas**. Mesmo observando que essas ocorrências se referem às mesmas expressas anteriormente por Müller e Negrão (1988) e Oliveira (2018), optamos por adotar a nomenclatura fornecida pelos autores; uma vez que essa

denominação facilita a identificação das ocorrências e torna a leitura mais uniforme em nossas análises (capítulo 5).

Ainda sobre a disposição de antropônimos sem artigos, Amaral e Seide (2020) atribuem relevância aos casos de construções predicativas. Na concepção dos autores, quando esse tipo de ocorrência é observada, o antropônimo não atua como argumento sintático de um verbo e, também, não pode ser considerado uma expressão referencial; assim sendo “trata-se de uma forma para apresentar o nome próprio de um indivíduo, construção que se assemelha a outras denominativas.” (AMARAL E SEIDE, 2020, p. 116). Os autores assumem que sentenças dessa categoria servem para informar o nome civil de um indivíduo (22); no entanto, a inversão entre os elementos e o verbo de cópula parece favorecer a presença do determinante, conforme aponta os exemplos dos autores em (22b).

- (22) a. Pelé é **Edson Arantes do Nascimento**.  
 b. Edson Arantes do Nascimento é (o) **Pelé**.

Em nossa percepção, essas construções podem sofrer influência não apenas da inversão entre os termos; mas também, do tipo de antropônimo que é licenciando após o verbo copulativo com a presença do determinante (**Pelé** – nome artístico do jogador Edson Arantes do Nascimento). Se tal pressuposição se confirmar em nossos dados, poderemos assumir que o critério tipológico (expresso no quadro 1) pode influenciar nas construções predicativas com verbos de cópula. Sob condição de esse ser o único dado – desse tipo – apresentado pelos autores, observaremos como (se) as sentenças predicativas seguem o (mesmo) padrão em nossos dados analíticos (cf. capítulo 5).

Müller e Negrão (1988) expõem, também, que a estrutura [artigo + nome próprio] em contextos tidos como existenciais (23) parece estranha ao PB. As autoras explicam o fato pautando-se na premissa de que, em geral, verbos existenciais não se apresentam acompanhados por essa estrutura, já que eles têm a função de inserir um elemento novo no discurso, ou seja, “introduzir uma nova entidade no velho modelo de mundo que é, assim, feito.” (MÜLLER E NEGRÃO, 1998, p. 534). Vejamos os exemplos contextualizados e utilizados pelas autoras:

➤ Não existe (tem\ há) ninguém aqui que fale russo

- (23) a. Existe (tem\ há) sim! **O carlos**.  
 b. \*Existe (tem\ há) sim! **Carlos**.

No entanto, ao apresentarem um contexto para as ocorrências de verbos existenciais (“não existe/tem/há ninguém aqui que fale russo”), Müller e Negrão afirmam que a omissão do artigo torna “a sentença sem sentido no nosso dialeto” (MÜLLER E NEGRÃO, 1998, p. 535). Isto é, mesmo quando o nome próprio se refere a um ser real, para as autoras, a ausência do artigo parece deixar a sentença agramatical em contextos existenciais no PB (cf. 23 b.).

Sobre essa suposta agramaticalidade sugerida por Müller e Negrão (1988), Oliveira (2018) questiona a generalização proposta pelas autoras. De acordo com estudos recentes acerca do português falado na cidade de Tabira<sup>17</sup>, a autora afirma que o padrão proposto por Müller e Negrão (1998) – a impossibilidade de sentenças existenciais acontecerem sem o artigo – parece não ser recorrente em seus dados. Segundo Oliveira (2018), a omissão do artigo definido em contextos de verbos existenciais não só é aceita na comunidade estudada, como parece ser recorrente na fala de informantes de faixas etárias e sexos distintos. Vejamos os exemplos<sup>18</sup>, a seguir:

➤ Há um amigo de quem você gosta mais?

(24) Tem (existe/ há) [Ø **Lorena**] e mais um monte de amiga.

(25) Tem (existe/ há) [Ø **Rafael**], meu amigo da escola.

Para Müller & Negrão (1988), estes casos acontecem quando os informantes incluem um referente que ainda não havia sido citado, apresentando-o, automaticamente, à inquiridora. Assim, os casos de antropônimos ancorados a verbos existenciais, seriam essencialmente licenciados com artigo; já que, segundo elas, a variante sem artigo não é possível no dialeto verificado pelas autoras. Contudo, Oliveira (2018) ressalta que muitas ocorrências, como as exemplificadas em (24) e (25), foram encontradas em sua pesquisa; fator que parece contrariar o que Muller & Negrão (1988) pontuam; evidenciando, assim, que Pernambuco parece ter um padrão diferente do que foi proposto pelas teóricas em seus estudos.

De modo sumário, Müller e Negrão utilizam o exemplo a seguir, para reafirmarem a hipótese de que a realização do artigo definido diante de antropônimos só é gramaticalmente aceita quando esses nomes se referem a um indivíduo em si. Vejamos que em (26),

---

<sup>17</sup>Comunidade localizada no interior do estado de Pernambuco. De acordo com Oliveira (2018), essa comunidade de fala possui um padrão linguístico que privilegia a ausência de artigo em detrimento de sua presença. Discorreremos acerca desse trabalho na seção 1.3 desse capítulo.

<sup>18</sup> Exemplos extraídos de Oliveira (2018, p. 64)

inversamente, o nome próprio (Carlos) está referenciando a entidade linguística, por isso a agramaticalidade expressa em (26b.):

- Não sei que nome dar ao meu filho. Não tem (existe\ há) nenhum nome que seja bonito e prático

- (26) a. Tem (existe\ há) sim! Carlos  
b. \*Tem (existe\ há) sim! O Carlos

Dizendo de outro modo, para Müller e Negrão (1988), a ausência do artigo definido diante de nomes próprios de pessoas acontece quando, semanticamente, esse nome refere-se à entidade linguística (o próprio nome) e não à pessoa em si, por isso, a sentença expressa em (26 b.) é agramatical no português brasileiro. Esse fator parece evidenciar a importância de distinguir as particularidades que circundam a antroponímia em PB, dado o fato de que o uso variável do artigo definido antepondo nomes próprios não acontece irrestritamente – conforme já delimitamos anteriormente. Assim sendo, discutiremos, a seguir, as ocorrências em que o uso do determinante – artigo definido – parece ser imprescindível para a leitura contextualizada de sentenças com sintagmas antroponímicos.

### • Antropônimos com artigos

Em trabalhos sobre o uso de antropônimos e artigos definidos, Sedrins (2017) aponta que há contextos em que a não-opcionalidade também pode ser observada em casos de marcação obrigatória do determinante. De acordo com o autor, quando o nome próprio é realizado com um modificador (adjetivo ou sintagma preposicionado, por exemplo) que tende a restringir o referente, o licenciamento do determinante parece ser obrigatório no PB. Nesses casos, se presume haver mais de um referente, sendo necessário individualizá-lo por meio de algum elemento especificador. Além do modificador, o recurso utilizado para essa individualização parece ser, justamente, o uso marcado da partícula determinante. Os exemplos<sup>19</sup> em (27) ilustram o que é exposto pelo autor:

- (27) a. **O João** [de camisa vermelha] é meu aluno.  
b. \***João** [de camisa vermelha] é meu aluno.

---

<sup>19</sup> Exemplos extraídos de Sedrins (2017, p. 3).

O autor sugere que, quando há a ocorrência de dois antropônimos iguais, isto é, dois referentes de mesmo nome (João), o uso do artigo parece ser obrigatório diante do antropônimo em que recai a ênfase: (O João de camisa vermelha e não outro João). Em outras palavras, podemos dizer que a utilização de um modificador, o sintagma preposicionado [de camisa vermelha], não só individualiza o antropônimo, como parece exigir uma marca de definição que é conferida, nesse caso, pelo uso do artigo definido (cf. 16a), o que justifica a agramaticalidade<sup>20</sup> apontada pelo autor em (27b). Amaral e Seide (2020) também descrevem ocorrências semelhantes que serão observadas mais adiante.

Em dados do português falado no interior de Pernambuco, observamos algumas ocorrências extraídas de Oliveira (2018, p. 65) que corroboram como o que é discutido pelo autor. Vejamos as construções contextualizadas expressas em (28), a seguir:

➤ Anchieta mudou muito, *parece* duas pessoas *diferentes*.<sup>21</sup>

- (28) a. **O Anchieta** [de São Paulo], marido, só visitante...  
 b. **O Anchieta** [do Logradouro], amigo, namorador.

Sintaticamente, podemos verificar que os dois antropônimos licenciados pelo artigo definido (28a) e (28b), [O Anchieta], estão acompanhados por sintagmas preposicionados modificadores que são responsáveis por situar esse referente em épocas e lugares distintos: [de São Paulo] e [do Logradouro]. Podemos observar que esses antropônimos que são licenciados com a presença dos modificadores estão, assim como no exemplo (27a), com a posição do determinante (nesse caso, o artigo definido) foneticamente concretizada. Tais constatações parecem assegurar a assertiva observada por Sedrins (2017) de que, havendo mais de um referente, a marcação do determinante parece ser obrigatória nos casos em que o antropônimo individualizado é especificado por meio de um modificador (no caso de (28) duas versões individualizadas e diferentes da mesma pessoa).

---

<sup>20</sup> É válido observarmos que a agramaticalidade atribuída à sentença (18b) só pode ser verificada quando traçamos um laço comparativo com a sentença (18a), já que construções como a expressa em (18b), fora do contexto de comparação, poderiam facilmente ocorrer em dialetos que tendem a anular a presença do artigo no PB, como é o caso do português falado na cidade de Tabira.

<sup>21</sup> Segundo Oliveira (2018), no que se refere à semântica das sentenças expressas em (19), o informante estava se reportando ao mesmo referente: Anchieta. Nesse caso, a ideia é que ele havia mudado por intermédio de suas ações no tempo-espaço, transformando-se, assim, em outra pessoa diferente da que já foi um dia. Para tanto, houve uma espécie de individualização, um contraste entre essas “pessoas” por intermédio dos modificadores.



Ainda de acordo com Sedrins (2017), as ocorrências em que o modificador restringe a leitura de um nome próprio e exige a realização do determinante (artigo definido, por exemplo) podem ser observadas, também, em línguas como o inglês que não têm, necessariamente, o uso marcado do artigo com antropônimos. Para justificar essa asserção, o autor lança mão dos exemplos observados em (29)<sup>22</sup>, a seguir:

- (29) a. **John** has left/ \***The John** has left.  
 b. The **tall John** has left. / \***Tall John** has left.

Como é possível verificar, a não exigência do artigo diante de antropônimos no inglês pode ser observada em construções como as expressas em (29a). No entanto, quando há a presença do modificador “*tall*” que restringe a leitura do antropônimo “*John*”, a presença do determinante parece ser obrigatória, como em (29b). Assim sendo, de acordo como os exemplos, podemos inferir que o artigo definido parece não ser exigido junto a antropônimos em inglês, exceto quando o nome é modificado por um elemento que restringe sua leitura, sendo necessária, assim, a realização desse determinante (cf. 29b), semelhante ao que ocorre em PB nos exemplos (27) e (28).

Em se tratando dos antropônimos com artigos, Amaral e Seide (2020) também pontuam sobre as ocorrências em que a aplicação do determinante parece ser exigida no PB, isto é, os casos em que é necessária a anteposição do artigo definido para o sintagma antroponímico obter uma leitura satisfatória. Os autores argumentam que, quando o antropônimo vem acompanhado por um complemento restritivo, um artigo (ou determinante de outra natureza) deve ser inserido para que haja uma interpretação contrastiva entre um ou mais referentes envolvidos na sentença. Em razão dessa afirmação, os autores lançam mão dos seguintes exemplos:

- (30) **O Luís infantil** não existe mais.  
 (31) **A Camila dos anos 1990** não é mais a mesma.  
 (32) **A Priscila que eu conheci** está muito diferente.

Conforme pontuam os autores, os elementos (modificadores) que acompanham os antropônimos nas sentenças têm a capacidade de qualificar restritivamente a leitura desse nome, assim como já verificado por Sedrins (2017), nos exemplos expostos em (27). Podemos

---

<sup>22</sup> Exemplos extraídos de Sedrins (2017, p.3). Grifo nosso.

observar que no exemplo (30), o antropônimo é licenciado com um adjetivo (infantil) e isso – atrelado ao uso do artigo definido – restringe a leitura do nome próprio **Luís**. Ou, dizendo de outra maneira, o uso do modificador parece exigir a leitura definida que o artigo é capaz de denotar.

O mesmo tipo de operação parece ocorrer com o emprego do sintagma preposicionado em (31) e da oração adjetiva em (32) que acompanham os antropônimos **Camila** e **Priscila**, respectivamente. Segundo os autores, o determinante parece ser indispensável para o sentido esperado nesses tipos de construções, em outras palavras: “o artigo é imprescindível e a leitura sempre é de oposição de uma imagem (ou faceta) a outra do indivíduo.” (AMARAL E SEIDE, 2020, p. 121). Isto é, em conformidade com as postulações de Sedrins (2017), os estudos de Amaral e Seide (2020) também apontam para o fato de que o uso de um modificador (que restringe a leitura do nome) parece solicitar o uso marcado do artigo definido.

Até aqui, discutimos acerca da antroponímia em língua portuguesa e algumas de suas características estruturais básicas. De modo sumário, observamos que o português falado no Brasil permite a variação do artigo definido diante de antropônimos; no entanto, essa variação não acontece de maneira irrestrita, haja vista que, conforme Sedrins (2017), ela só é permitida quando estamos nos referindo a um único ser no discurso. À vista disso, investigamos as ocorrências em que o uso variável do artigo parece não ser permitido no português falado no Brasil. Isto é, os casos nos quais a marcação do determinante é exigida para garantir a gramaticalidade das sentenças e os casos em que essa marcação parece ser anulada no sintagma antroponímico: (i) antropônimos com artigos e (ii) antropônimos sem artigos.

Dito isso, na seção seguinte, nos debruçaremos sobre alguns trabalhos sociolinguísticos que estudam a alternância no uso do artigo frente a antropônimos em algumas comunidades de fala brasileiras.

## 2.2 O USO VARIÁVEL DO ARTIGO DEFINIDO: O QUE DIZEM ALGUMAS PESQUISAS SOCIOLINGUÍSTICAS

Magalhães (2012), ao levar em consideração textos portugueses dos séculos XVI e XIX, ressalta que, distintivamente do PB que permite o uso facultativo do artigo definido, o Português Europeu (PE) não permite o mesmo licenciamento variável. De acordo com autora, o PE passou por um processo de mudança ao longo dos séculos até atingir um estágio de não-opcionalidade no que tange à realização variável do artigo. Apoiando-se nos estudos de Castro (2000),

Magalhães (2012) admite que a realização do artigo definido diante de possessivos, por exemplo, é obrigatória em PE atualmente, enquanto o mesmo comportamento se mostrava raro no século XV. Nesse sentido, a autora ressalta não apenas o uso facultativo do artigo no PB, mas também a ideia de mudança linguística ao longo dos séculos (no PE), em que uma variante pode, gradativamente, perder espaço para a outra.

Assim, parece justo inferirmos que esse fenômeno ainda se encontra em processo de variação no Brasil. Os estudos que versam acerca do emprego do artigo definido, tanto no contexto de antropônimos – recorte analítico desta pesquisa – quanto em contextos de pronomes possessivos, possuem uma ampla margem de dados quantificados e descritos, também, no português brasileiro (PB). Deste modo, o uso variável do artigo definido em sintagmas antroponímicos é pauta para inúmeras pesquisas que investigam o comportamento desses sintagmas e as variáveis que condicionam, ou não, alguns usos.

Isso posto, buscaremos, nesta seção, revisitar e discutir os resultados de algumas das pesquisas sociolinguísticas que já atentaram para o comportamento variável do artigo definido em várias comunidades de fala no Brasil.

### **2.2.1 Callou e Silva (1997)**

Entre as pesquisas mais visitadas – para aqueles que se debruçam sobre o fenômeno em questão – está o trabalho realizado por Callou e Silva (1997). Norteadas pelos pressupostos teórico-metodológicos da sociolinguística laboviana e fazendo uso do programa computacional VARBRUL, as autoras analisam o uso do artigo definido diante de pronomes possessivos e nomes próprios em cinco capitais brasileiras (Porto Alegre, São Paulo, Rio de Janeiro, Salvador e Recife). As pesquisadoras adotam como base corpórea entrevistas realizadas pelo Projeto Norma Urbana Oral Culta (NURC), nas quais todos os informantes possuíam curso universitário completo.

Os resultados de Callou e Silva (1997) indicam que a variável *região de origem* exerce grande influência nos dados observados, haja vista que as duas cidades da região nordeste apresentaram um nítido contraste das demais capitais estudadas: a realização do artigo mostrou-se menos expressiva em Recife e Salvador (como é possível verificar no quadro abaixo). As pesquisadoras assumiram a hipótese de que tal distinção, marcada pela baixa frequência do artigo nas duas capitais, pode indicar certo conservadorismo linguístico no falar do nordeste, onde a colonização é mais remota. O quadro (2) a seguir, ilustra melhor essa assertiva:

**Quadro 2** - Realização do artigo definido diante de antropônimos e diante de pronomes possessivos, apresentada no trabalho de Callou & Silva (1997)

<b>Capitais</b>	<b>Antropônimos</b>	<b>Possessivos</b>
Recife	(12/71) 17%	(59/98) 60%
Salvador	(10/24) 32%	(57/87) 66%
Rio de Janeiro	(27/85) 43%	(280/399) 70%
São Paulo	(20/23) 87%	(147/209) 70%
Porto Alegre	(50/63) 79%	(26/33) 79%

**Fonte:** Callou & Silva (1997, p. 21)

As autoras, ao levarem em consideração a variável *presença de preposição*, constataam que tal variável favoreceu constantemente o uso do artigo definido. Seus resultados sugerem que a presença de preposição perante possessivos beneficia a realização do artigo, sobretudo quando as preposições aceitam contrair-se com o artigo, elevando seu peso relativo (P.R.)<sup>23</sup> ao patamar de 0,70. O mesmo parece acontecer nos contextos de anteposição a antropônimos (ex.: *da Maria*), já que o P.R. de 0,77 mostra-se expressivo nos casos em que as preposições são passíveis de contração com o artigo.

Callou e Silva (1997), em seus estudos, também consideram a variável *função sintática* e constataam que as posições de sujeito ou tópico, objeto ou predicativo mostram-se bastante significativas quando o artigo definido antecede um pronome possessivo. A posição de aposto, por sua vez, não foi considerada condicionante nos dados observados. Já, no contexto de antropônimos, a função de adjunto foi a mais expressiva com um P.R. de 0,98. Contudo, é válido inferirmos que adjuntos, em geral, são sempre antecidos por uma preposição, fator que tende a favorecer a realização do artigo, como foi examinado pelas autoras na análise da variável *presença de preposição*. Ainda sobre os antropônimos, as autoras salientam que, assim como nos casos de possessivos, os apostos não são licenciados com a presença do artigo, sendo esta a função sintática com resultados menos expressivos nos dados analisados, com um P.R. de 0,12.

### 2.2.2 Campos Júnior (2013)

Colocando em relevo questões sociais e, sobretudo, identitárias da fala, é válido citarmos a pesquisa realizada por Campos Júnior (2013), que analisou a alternância entre a realização e

---

<sup>23</sup> Vale ressaltarmos que no capítulo metodológico desta dissertação discorreremos mais profundamente acerca do P.R.

não realização do artigo definido diante de antropônimos na cidade de Vitória, capital do Espírito Santo. Ao assumir a hipótese geral de que a comunidade capixaba utilizava uma variedade linguística menos marcada do que as de outras regiões brasileiras, o autor observou que o uso do artigo diante de nomes próprios pode se configurar como uma marca de identidade linguística dos falantes e, por isso, defende a hipótese de que o sexo do falante tem importância nos processos linguísticos da comunidade<sup>24</sup>.

Campos Junior (2013) observou que os falantes da cidade de Vitória tendem a usar menos o artigo definido diante de antropônimos do que informantes oriundos de outras regiões brasileiras. Tal assertiva parece corroborar com a hipótese inicial proposta pelo autor de que existe certo conservadorismo linguístico na comunidade investigada. Campos Junior (2013), no entanto, vai além dessa hipótese e acredita que esse comportamento linguístico se deve ao fato de os moradores utilizarem a língua, ainda que de forma inconsciente, como um recurso de manutenção da identidade linguística da região. O quadro 3 evidencia as realizações observadas pelo autor:

**Quadro 3** – Realização do artigo definido diante de antropônimos de acordo com a origem geográfica

	VIX	BH	RE	SSA	RJ	SP	POA
Freq.	118/300	-	12/71	10/24	27/85	20/23	50/63
%	39	76	17	32	43	87	79

**Fonte:** Campos Jr. (2011, p.26)

Como base no quadro 3, podemos observar as capitais brasileiras que apresentam uma maior frequência no licenciamento do artigo definido diante de antropônimo: São Paulo (SP) com 87%; Porto Alegre (POA) com 79% e Belo Horizonte (BH) com 76%. Na contramão desse seguimento, estão as capitais com menos expressividade no que tange a realização dos artigos definido: Rio de Janeiro (RJ) com 43%; Salvador (SSA) com 32%; Recife (RE) com 17%, e a capital capixaba, Vitória (VIX), observada por Campos Júnior (2011), que apresentou apenas

---

<sup>24</sup> O nativo de Vitória (ES) tem utilizado a língua como um recurso a mais na busca da manutenção de sua identidade, tão fragilizada ao longo de sua história. Primeiro, por décadas de descaso socioeconômico. Segundo – e principalmente – como retorno da reviravolta que a economia local sofreu, pela grande invasão de pessoas oriundas de outros estados. (CAMPOS JR, 2011, p. 26)

39% de ocorrência de artigo, nivelando-se às cidades com menor assiduidade para a marcação do artigo definido diante de nomes próprios.

Levando em consideração a faixa etária e o sexo dos informantes, o autor constatou que os homens usam mais o artigo do que as mulheres. Com um percentual de realização de 54% para os homens e de apenas 30% para as mulheres, o autor aponta duas hipóteses que podem explicar esse resultado. A primeira é de que as mulheres são mais conservadoras que os homens, pois – como mencionado pelo autor – o padrão da comunidade capixaba seria uma tendenciosa ausência do artigo definido diante de antropônimos. A segunda hipótese, proposta pelo autor, baseia-se em Callou e Silva (1997) para indicar que a não realização do artigo pode ter relação com fatores históricos da colonização do país, pois, de acordo com as autoras, quanto mais antiga a colonização, menor a realização do artigo.

Nesse contexto, o autor sugere que a ausência de artigos seria um traço arcaizante da língua difundido pelas mulheres. Em termos gerais, Campos Junior (2011) ressalta que a questão do movimento identitário<sup>25</sup> é uma hipótese instigante que merece maior atenção para pesquisas que podem surgir *a posteriori*.

### 2.2.3 Sedrins, Pereira e Siqueira (2015)

Sedrins, Pereira e Siqueira (2015) analisam a variação na realização do artigo definido diante de antropônimos em dados de língua falada e língua escrita no sertão pernambucano. Os resultados gerais abalizam que a realização do artigo, nessa região, tem baixa expressividade. Contudo, os dados examinados sugerem que a modalidade falada da língua parece favorecer de maneira mais significativa o uso do determinante, do que os dados de língua escrita, isto é, aqueles que foram provenientes de textos escritos pelos informantes e não por meio de entrevistas concedidas oralmente.

O *corpus* que corresponde aos dados escritos atingiu um número percentual pouco significativo nos dados analisados: 3% (7/259) de realização do artigo, contra 97% (252/259) de não realização do artigo. Os dados de língua falada, quando comparados aos dados da modalidade escrita, apresentaram uma leve sensibilidade para a realização do artigo: 9% (26/279) de marcação do artigo, contra 91% (253/269) de não marcação do artigo.

---

<sup>25</sup> Para o autor, esse movimento indica que os falantes da comunidade investigada parecem utilizar a forma sem artigo definido em contextos antroponímicos como uma espécie de manutenção da identidade linguística da região.

Os autores, nesse estudo, levam em consideração a variável *Status informacional*, também controlada no (já citado) trabalho de Silva (1998). Os resultados obtidos pelos pesquisadores indicam que o fato de o antropônimo ser considerado, ou não, um elemento novo no discurso não influencia na realização da partícula definida. Na modalidade escrita, quando o elemento foi considerado uma inserção nova no discurso, o percentual foi de 3% (6/66) de ocorrência do artigo; já, quando o elemento (nome próprio) já havia sido citado no discurso (informação velha), o resultado de ocorrência do artigo foi também de 3% (5/162). No que concerne aos dados de fala, dos 183 sintagmas nominais selecionados para investigação, apenas 18 destes ocorreram com o artigo definido.

Em seus estudos, os autores também analisam a *variável Informação compartilhada* e pautam-se na hipótese de que, assim como trazido em Silva (1998), “a necessidade de definição por parte do falante (o uso do artigo antes do nome próprio) diminuiria à medida que o referente fosse uma personagem também conhecida pelo falante” (SEDRINS, PEREIRA E SIQUEIRA, 2015. p. 211). Tal afirmação parece sugerir que, sendo um referente famoso – informação compartilhada (33a.), a presença do artigo parece ser dispensada, o mesmo parece ocorrer quando esse referente é uma informação não compartilhada; isto é, um referente não famoso (33b.) Seus resultados abalizam que esse fator não se mostra atuante para a realização do artigo, haja vista que, mesmo quando o informante é famoso, o padrão da comunidade tende a não marcar a posição da partícula determinante, como é possível verificar nos exemplos mencionados pelos autores:

- (33) a. “à noite fui ao show de **Luan Santana**”  
 b. “**Jaqueline** levou sua prima e suas amigas para sua casa”

Os autores mostram que, mesmo tratando-se de um referente de natureza compartilhada (33a), a tendência é de não realização do artigo definido, uma vez que o percentual foi de 100% (11/11) de não marcação. Nos casos em que o referente (nome próprio) não era de natureza compartilhada (33b) o percentual foi de 10% (26/267) de realização do artigo definido. Com isso, verifica-se que tal variável parece não exercer influência nos dados obtidos e discutidos pelos pesquisadores, a julgar pela baixa expressividade de aparição do artigo.

Sobre as variáveis extralinguísticas *faixa etária* e *sexo*, os pesquisadores indicam que estas podem influenciar a realização do fenômeno. De acordo com os autores, os informantes da primeira faixa etária realizaram mais o artigo definido: 15% (23/153) de realização, contra 85% (130/153) de não realização. A segunda faixa etária apresentou um percentual menos

expressivo de 5% (3/66) de realização, contra (63/66) de não realização. A terceira faixa etária, por sua vez, não apresentou nenhum caso de ocorrência do artigo. Desse modo, os autores verificam que a variável *faixa etária* influencia no comportamento dos dados, tendo em vista que os adultos parecem não utilizar a variante que preenche a posição do determinante na sentença. No que tange ao sexo dos informantes, os autores aclaram que as mulheres realizam mais o artigo do que os homens. As mulheres apresentaram um percentual de 16% de realização, contra 84% de não realização; ao passo que os homens apresentaram um resultado percentual de 2% de realização, contra 98% de não realização.

#### 2.2.4 Pereira (2017)

Em sua dissertação de mestrado, Pereira (2017) volta a se debruçar sobre o fenômeno de variação do artigo definido no português falado no sertão do Pajeú pernambucano. O autor, adotando como modelo teórico e metodológico a Sociolinguística Variacionista e o programa o GoldVarb X – como aporte computacional e estatístico, verifica como se dá a realização do artigo definido diante de pronomes possessivos e diante de antropônimos em dados provenientes das cidades de Carnaíba e Serra Talhada. Em termos gerais, o autor constata “o que já havia sido obtido por Pereira (2011) no que se refere à tendência à não-realização do artigo: os municípios de Carnaíba e Serra Talhada, curiosamente, obtêm o mesmo percentual de realização do artigo: 17%, apresentando-se como característica local.” (PEREIRA, 2017, p. 191.).

Com base em seus resultados, Pereira (2017) averiguou que os dois municípios estudados na pesquisa carregam a propriedade similar de inibir a presença do artigo definido, principalmente, diante de antropônimos. O resultado percentual indica que para o contexto de pronomes possessivos a ocorrência do artigo foi de 28% na cidade de Serra Talhada e de 23% na cidade de Carnaíba. Já, em relação aos antropônimos, com uma realização bem menos expressiva, a cidade de Serra Talhada teve um percentual de apenas 9% de uso do artigo, ao passo que a cidade de Carnaíba obteve um percentual de 15% de uso do artigo definido.

O autor descreve a influência das variáveis de acordo com as cidades analisadas. No município de Carnaíba a variável *tipo de preposição* mostrou-se expressiva, a julgar pela preposição *de* que, com um resultado de 43% realização, foi a que mais condicionou o uso do artigo definido. Pereira (2017) acredita que a expressividade obtida pela preposição *de* “seja efeito de seu traço de [+aglutinante], o que acaba favorecendo a realização do artigo” (PEREIRA, p. 191, 2017), isto é, deve-se ao fato dessa preposição permitir se contrair com



artigos definidos. No que diz respeito ao contexto de pronomes possessivos, o autor analisou que a variável tipo de preposição também se mostrou relevante, com P.R. de 0,82, é a preposição que mais favorece a realização do artigo,

No que diz respeito à variável *função sintática*, o autor constatou que a posição de *adjunto* (P.R. 0,84) foi o fator que mais privilegiou a realização do artigo definido diante de antropônimos na cidade de Carnaíba. Outro fator que se mostrou atuante foi a função sintática de *tópico*, com P.R. de 0,64, seguida da função sintática de *predicativo* que obteve um P.R. de 0,60. Já em relação ao contexto de pronomes possessivos, o autor verificou que os fatores *objeto indireto*, *sujeito* e *objeto direto* foram os que mais colaboraram para a realização do artigo, com P.R. de 0,75, 0,59 e 0,70, respectivamente.

A variável *referência do antropônimo*, que diz respeito à natureza do referente, ou seja, se a entidade (nome próprio) referida no discurso é um personagem ou uma pessoa real, também exerceu influência nos dados observados pelo autor. Segundo Pereira (2017), no município de Carnaíba, o fator *personagem* favorece de maneira considerável a realização do artigo definido. Desse modo, com um P.R. de 0,78, houve um favorecimento para realização do artigo quando o nome se referia a um ser de natureza ficcional (personagem de televisão). Já os referentes de natureza real obtiveram um P.R. de 0,41.

O autor destaca, ainda, as duas variáveis de ordem extralinguísticas que se mostram significantes nos resultados obtidos na cidade de Carnaíba: a) o *sexo* do informante (homens, ao contrário das mulheres, produzem mais o artigo diante de possessivos, com P.R. de 0,62) e b) a faixa etária “(a faixa 1, ao contrário da faixa 3, é a que menos favorece a realização do artigo (P.R. de 0,37))” (PEREIRA, 2017, p. 191). Desse modo, assim como acontece nos estudos de Sedrins, Pereira e Siqueira (2015), os informantes mais velhos (terceira faixa etária) tendem a não fazer a realização do artigo definido na região do sertão do Pajeú pernambucano.

No que diz respeito aos resultados obtidos na cidade de Serra Talhada, o autor constatou que as variáveis linguísticas *referência do antropônimo*, *função sintática* e *tipo de preposição* foram as que mais obtiveram significância. Assim como no município de Carnaíba, a variável *referência do antropônimo* parece ter condicionado de forma expressiva a realização do artigo definido em Serra Talhada. Quando o nome próprio se referia a um *personagem* de televisão, filme ou séries o P.R. foi de 0,76, já quando o nome próprio era do tipo *real*, o P.R. foi de 0,43.

Pereira (2017) observou que a *função sintática* mostrou-se atuante em Serra Talhada nos dois contextos de análise perquiridos no seu estudo: diante de antropônimos e diante de pronomes possessivos. De acordo com os resultados obtidos, a função sintática de *adjunto* foi a que mais favoreceu a realização do artigo definido com um P.R. de 0,87. Conquanto o autor

destaca que, pelo fato de o adjunto vir acompanhado por preposição, esta variável – *tipo de preposição* – pode exercer influência na variável *função sintática*. Já que a preposição parece favorecer a realização do artigo, as funções sintáticas que aparecem acompanhadas delas serão, conseqüentemente, favorecidas. Diante do contexto de possessivo, os resultados não foram muito distintos, mas, ao contrário do que aconteceu com os antropônimos, o programa GoldVarbX selecionou a variável *contexto preposicionado* e mostrou que o fato de se ter preposição é algo significativo com um P.R. de 0,67, uma vez que a função de *objeto indireto* foi uma das que mais favoreceram a realização do artigo, com P.R de 0,79.

### 2.2.5 Oliveira (2018)

Em trabalho mais recente acerca do uso variável do artigo definido em sintagmas antroponímicos, Oliveira (2018) também investiga dados provenientes do Sertão do Pajeú pernambucano. Em sua pesquisa, a autora averigua se há variação desse fenômeno na cidade de Tabira e discute como se dá esse processo nos dados da comunidade de fala pesquisada. A hipótese geral da pesquisa era de que, em conformidade com o estado a qual pertence, a comunidade de fala registraria um uso menos marcado do artigo definido diante de antropônimos. Dentre as principais motivações do trabalho, a pesquisadora ressalta o fato de sua investigação ser pioneira, dado que nenhum registro linguístico da cidade havia sido estudado, até então, acerca do português falado na comunidade.

Norteadas pela base epistêmica da Sociolinguística Variacionista e valendo-se do programa computacional GoldvarbX, que proporciona dados estatísticos, Oliveira (2018) constatou que, assim como pontuam Callou e Silva (1997) e Marroquim (1996) [1945], a cidade de Tabira segue em uniformidade com o falar de Pernambuco, uma vez que opta pela variante sem artigo, o que corrobora com a hipótese inicial da pesquisa. Com isso, a autora pondera que comunidades localizadas no nordeste (Recife e Salvador) tendem a licenciar menos artigos do que cidades localizadas no sul e sudeste do país (cf. CALLOU E SILVA, 1997).

Com base nos dados divulgados pela autora, foi possível observar que a comunidade de fala emprega as duas variantes possíveis para o fenômeno – com artigo definido e sem artigo definido. Com um percentual global de 79,9% (167/209) de não realização do artigo definido frente a um total de 20,1% (42/209) de realização do artigo definido, Oliveira (2018) aponta que o padrão da comunidade é expressivo quanto à ausência do artigo definido diante de nomes próprios de pessoas, uma vez que a variante sem artigo é a mais empregada na cidade.

Segundo Oliveira (2018), os resultados estatísticos mostraram que as variáveis linguísticas Tipo de informação e Referência do antropônimo foram apontadas pelo programa computacional como sendo as mais relevantes para a variação. De tal modo, a autora ressalta que o fator *personagem* – da variável referência do antropônimo – condicionou em 88% o uso do artigo definido. Tal significância pode ser observada, também por meio do P.R. fornecido pela autora: 0.972; no que concerne ao outro fator da variável referência do antropônimo, o fator Real obteve um percentual pouco significativo de 11% (20/184) de ocorrência, com um P.R. de 0.382. Pondo em destaque a variável *tipo de informação*, Oliveira (2018) observa que o fator que mais condicionou a presença do artigo foi *informação compartilhada*, com um percentual de 43,7% (28/87) de ocorrência da partícula e um P.R. de 0.856; já o fator informação não compartilhada obteve o percentual baixíssimo de 3.3% (4/122) de ocorrência do artigo, e um pouco expressivo de P.R. de 0.219.

Oliveira (2018) destaca que, apesar de não serem apontados pelo programa GoldvardX como relevantes para a variação, as seguintes variáveis também foram consideradas e descritas em sua pesquisa: *Status informacional*, *Contexto preposicionado*, *Faixa etária* e *Sexo*. A descrição dessas variáveis se mostra importante, haja vista que também são consideradas nesta pesquisa.

Mesmo não realizando o cruzamento entre todas as variáveis selecionadas, o trabalho de Oliveira (2018) deixa algumas considerações e apontamentos importantes que valem a pena serem retomados e discutidos aqui. Na concepção da autora, a variável *função sintática* deve ser considerada em trabalhos futuros acerca do português falado na comunidade de fala pesquisada. A autora justifica que a ausência de variação (para alguns fatores da variável) impossibilitou o andamento do programa computacional, de modo que a variável *função sintática* teve que ser desconsiderada na rodagem oficial dos dados. Um fator que parece influenciar para esse tipo de acontecimento é a quantidade reduzida de dados e as numerosas funções sintáticas escolhidas. Nessa linha argumentativa, destacamos – conforme Oliveira (2018) – que tal variável será considerada no andamento de nossa pesquisa, a julgar pela importância de uma descrição mais estrutural do fenômeno.

Ademais, a autora aponta para uma questão importante que cabe ser discutida aqui. Conforme Oliveira (2018), o artigo definido parece desempenhar um papel sociolinguístico diferente no público LGBTQ+ da cidade. Isto é, alguns informantes pertencentes a esse grupo social parecem utilizar o que a autora optou por chamar de *não concordância de gênero* em seus dados. Ou seja, isso acontece quando os informantes empregam, de maneira proposital, a

marcação do artigo feminino diante de antropônimos comumente conhecidos como substantivos masculinos; como é possível observar em (33), a seguir:

(34) Tem **a Pablo Vittar**, que agora gosto de ouvir.

As discussões dessas ocorrências merecem evidência em nossa pesquisa, haja vista a recorrência de dados desse tipo em nosso *corpus* de análise. Para Oliveira (2018), como o artigo carrega alguns traços particulares no que se refere ao gênero e ao número do antropônimo, a sentença expressa em (34) apresenta uma inversão marcada pelo uso da determinante feminino “a” em detrimento do masculino “o” antes do nome “Pablo”. Segundo a autora, “tal ocorrência parece marcar a identidade de uma comunidade específica de falantes que utilizam a língua e toda sua heterogeneidade como forma de expressão e inserção social” (OLIVEIRA, 2018, p. 70). Dizendo de outro modo, percebemos uma influência de questões sociais e identitárias que parecem interferir diretamente na organização canônica da concordância no interior dos sintagmas antroponímicos.

Outro apontamento deixado por Oliveira (2018) é a hipótese de que os nativos de Tabira parecem utilizar a variante *ausência do artigo* como forma de manutenção de uma identidade linguística local, conforme também observou Campos Júnior (2013). Para a autora, alguns colaboradores da pesquisa se mostraram resistentes à variante inovadora (com artigo definido) na comunidade de fala pesquisada. Ao consultar a ficha social desses informantes, Oliveira (2018) infere que esse comportamento se mostrou bem mais recorrente em informantes que reafirmavam veementemente sua identidade local e uma afinidade mais acentuada com a cultura sertaneja (música, expressões teatrais e poesia popular). Nosso interesse nesses apontamentos observados pela autora é investigar se a mesma hipótese se sustenta na capital do estado. Assim sendo, observaremos se o comportamento linguístico identitário, sugerido por Campos Junior (2013) e retomado por Oliveira (2018), será recorrente também na cidade do Recife, capital do estado.

Até esse momento, revisitamos algumas das pesquisas sociolinguísticas relevantes no que se refere à realização variável do artigo definido em sintagmas antroponímicos no PB. Analisamos trabalhos mais amplos que explanam sobre comunidades brasileiras mais distantes, como é o caso da pesquisa de Callou e Silva (1997) – que sobreleva algumas capitais do Brasil – e o trabalho de Campos Junior (2013), que põe em relevo o português falado na capital do Espírito Santo, Vitória.

Situando ainda mais nossa ótica teórica, discutimos trabalhos mais locais que investigam o comportamento do artigo no estado de Pernambuco (SEDRINS, 2013; SEDRINS, PEREIRA e SIQUEIRA, 2015; PEREIRA, 2017; OLIVEIRA, 2018). Desse modo, após revisarmos o estado da arte acerca do nosso fenômeno analítico, acreditamos na contribuição de tais trabalhos para o andamento de nossa pesquisa, uma vez que a nossa investigação acerca do português falado no Recife dará continuidade ao mapeamento linguístico que já vem sendo produzido no estado.

Assim sendo, no próximo capítulo, discutiremos todos os procedimentos metodológicos e éticos que nortearam essa pesquisa.

### 3 BASE EPISTÊMICA: A SOCIOLINGÜÍSTICA VARIACIONISTA

Ao eleger um fenômeno linguístico como objeto de estudo, seja ele de qualquer nível (fonológico, morfológico ou sintático), é importante discutir quais são os seus desdobramentos estruturais e sociais na língua a ser estudada. Isto é, quando adotamos que os elementos linguísticos desempenham papéis específicos dentro do sistema linguístico, precisamos compreender, também, quais são os valores comunicativos que esses elementos assumem na esfera social da comunicação humana. Nessa linha de raciocínio, evidenciamos que, mesmo partindo da estrutura linguística de um dado fenômeno<sup>26</sup>, estamos associados às concepções mais funcionalistas<sup>27</sup> da linguística, àquelas que elegem as relações semântico-pragmáticas desenvolvidas pelos sujeitos na sociedade como ponto de partida para a investigação das línguas humanas.

Assim, concordamos que o movimento de diligência aos aspectos mais sociais das línguas ganha força em meados do século XX, com a Virada Pragmática. As mudanças consideráveis<sup>28</sup> provocadas por esse período fazem com que as ciências linguísticas inaugurem focos de estudos divergentes, abrindo espaço para o surgimento e a evolução de novas vertentes linguísticas. Nessa linha argumentativa, podemos citar – a título de exemplo – o cognitivismo Chomskyano, e as, já citadas, teorias de base social que apontam o sujeito como centro das operações linguísticas, a saber: Pragmática, Análise do Discurso, e a Sociolinguística, base epistêmica com a qual nos filiamos e que será o ponto de partida para as discussões propostas neste capítulo.

Isso posto, este capítulo está organizado em três seções específicas que abarcam alguns dos principais conceitos da Sociolinguística. Na seção 3.1, apresentaremos alguns dos trabalhos clássicos publicados por William Labov e suas principais contribuições para a Teoria Variacionista. Além disso, de maneira adjacente, discutiremos alguns aspectos específicos observados pelo autor<sup>29</sup>, por acharmos que eles serão fundamentais para o nosso capítulo de

---

<sup>26</sup> O uso variável da estrutura [artigo definido + antropônimo] no PB.

<sup>27</sup> Mais especificamente, a Sociolinguística Variacionista.

<sup>28</sup> Como pontua Araújo (2004, p. 109): “o esquema triádico (signo, objeto e interpretante) leva em conta usuários e o contexto de fala, portanto, leva em conta os fatores pragmáticos da comunicação[...] O centro da linguagem não é mais a proposição assertórica. O mundo é visto como síntese de possíveis fatos, para uma comunidade linguística, para uma comunidade de interpretação, cujos membros têm condições de entender-se entre si, acerca de algo no mundo.”

<sup>29</sup> O conceito de afirmação linguística identitária, por exemplo, que será importante para nossas análises.

análise. Na seção 3.2, por sua vez, apresentaremos algumas considerações sobre a Sociolinguística no Brasil. Por fim, na seção 3.3, discutiremos alguns conceitos fundamentais (metodológicos e éticos) da teoria presentes, sobretudo, em manuais brasileiros de Sociolinguística.

### 3.1 WILLIAM LABOV E A SOCIOLINGUÍSTICA VARIACIONISTA

William Labov – associado às premissas funcionalistas<sup>30</sup> – é um dos principais estudiosos a evidenciar a noção de que linguagem e sociedade estão intrinsecamente ligadas e, por isso, formam a base da comunicação humana que se exterioriza por meio da língua. Em suas pesquisas, publicadas em 1963, acerca do inglês falado na ilha de *Martha's Vineyard* – no estado de *Massachusetts*, o autor associa fatores de ordem social, como idade, sexo, ocupação, origem étnica e atitude, aos comportamentos linguísticos dos nativos da ilha, no intuito de investigar como se dava a pronúncia de determinados fones do inglês. É por meio destes estudos que Labov consegue, através de uma metodologia refinada, sistematizar a heterogeneidade da língua falada, ganhando, a partir disso, grande notoriedade dentro dos estudos linguísticos.

Em termos específicos, o autor observou uma alternância sonora inerente ao contexto social da comunidade de fala. Trata-se da variação fonética dos elementos iniciais dos ditongos /ay/ e /aw/. Ao desenvolver uma criteriosa metodologia de análise – que levava em consideração não apenas fatores de ordem linguística (ambiente fonético), mas métodos baseados, também, na sociologia e antropologia da comunidade – Labov tentou reorganizar a história contemporânea da mudança sonora. Com isso, foi possível perceber que o complexo padrão linguístico da comunidade admitia a influência direta de fatores sociais passíveis de serem isolados e observados.

A grande questão em pauta parecia ser a explicação e possível reconstrução do processo de mudança linguística em curso que, segundo o autor, se segmentava em três questões basilares: “origem das variações linguísticas; a difusão e propagação das mudanças linguísticas; e a regularidade da mudança linguística” (Labov, 1972 [2008] p. 19). Assim, o autor argumenta que não pode haver um estudo de mudança linguística que desconsidere o contexto

---

<sup>30</sup> A gramática funcional leva em consideração o uso das expressões linguísticas na interação verbal; inclui na análise da estrutura gramatical toda a situação comunicativa: o propósito do evento da fala, os participantes e o contexto discursivo. (PETTER, 2008, p. 22).

socioeconômico da comunidade na qual ele está imerso. É dessa proposição que depreendemos a afamada assertiva proposta pelo autor: de que as pressões sociais operam como uma força que atua no presente vívido das línguas.

A escolha da comunidade de fala em investigação é justificada, sobretudo, pela vantajosa distância que a ilha de *Martha's Vineyard* tem do continente, haja vista sua localização cercada pelo Oceano Atlântico. É válido ressaltarmos que a ilha é composta por pequenas localidades e seus respectivos habitantes (cerca de 6.000): descendentes de ingleses, ascendentes de portugueses e imigrantes, remanescentes indígenas (*Gay Head*), além de grupos formados por uma miscelânea de ingleses, franco-canadenses, irlandeses, alemães e poloneses e um elevado número de turistas em dadas épocas do ano.

**Figura 1** - Localização da Ilha de Martha's Vineyard no mapa



Fonte: <http://www.guiageo-eua.com/massachusetts/mapa-politico.htm>

O objetivo da pesquisa era entender as estruturas do inglês dos falantes da ilha de *Martha's Vineyard* e as diferenças internas que existiam, além de observar os processos de mudança em curso. Para o andamento da pesquisa, um modelo de entrevista que contemplasse as ocorrências do ambiente fonético (os ditongos /ay/ e /aw/) foi desenvolvido e aplicado em diversos contextos de uso: espontâneo, monitorado, estilo de leitura e na fala emocionalmente carregada. Além das entrevistas formais, foram feitas, também, observações em ocasiões diversas (nas ruas, lanchonetes, lojas etc.), a fim de circunstanciar ainda mais o *corpus* da pesquisa. Assim sendo, as informações foram agrupadas em um *corpus* com 69 entrevistas



advindas de falantes nativos da ilha<sup>31</sup>, com cerca de 3.500 ocorrências de /ay/ e 1.500 ocorrências de /aw/ no total global.

Em torno das discussões propostas pelo autor acerca dos padrões fonéticos da ilha de *Martha's Vineyard*, podemos inferir dois pontos que consideramos ser destacáveis de observação para o nosso trabalho: i) tanto fatores extralinguísticos (faixa etária), como fatores linguísticos (ambiente segmental) podem influenciar na centralização dos ditongos; ii) em contextos de afirmação cultural, histórica, geográfica e social, a língua pode ser utilizada como um elemento de manutenção identitária. Tais premissas ficam evidentes quando somos apresentados aos principais resultados da pesquisa proposta pelo autor.

Sobre os papéis sociais que podem influenciar a língua falada, é possível verificar que, de certo modo, a faixa etária dos informantes, “distribuição por idade e tempo” (LABOV, 1972 [2008], p. 41), foi importante para a centralização dos ditongos, a julgar pelo aumento sucessivo que se acentua, sobretudo, no grupo de 31 a 45 anos. Tais considerações podem ser verificadas no quadro 4, a seguir:

**Quadro 4 -** Centralização dos ditongos [ay] e [aw] por grupos faixa etária.

<b>Idade</b>	<b>(ay)</b>	<b>(aw)</b>
75-	25	22
61-75	35	37
46-60	62	44
31-45	81	88
14-30	37	46

**Fonte:** Labov (1972 [2008], p. 46)

Acerca dos resultados obtidos sobre a faixa etária, foi possível verificar um aumento regular entre as faixas etárias intermediárias (31 a 45 e 46 a 60). A partir do que se coloca, o autor questiona se é um caso de mudança sonora em curso ou caso de mudança regular relacionado à idade. Desse modo, Labov avalia que o percentual de centralização mostra que o efeito da idade não deve ser descartado, sendo, assim, um possível desencadeador de variação.

---

<sup>31</sup> “Os grupos ocupacionais mais importantes da ilha estão representados na amostra: 14 na pesca, 8 na agricultura, 6 na construção, 19 no ramo dos serviços, 3 profissionais liberais, 14 estudantes, 5 donas do lar. Os grupos étnicos, os três principais, também: 42 descendentes de ingleses, 16 de português e 9 descendentes de índios.” (LABOV, 1972 [2008], p. 32).

No que tange ao ambiente segmental dos ditongos, contexto em que fatores de base linguística são controlados, Labov observou que a posição das consoantes (consoante subsequente e consoante antecedente) poderia influenciar no fenômeno da centralização dos ditongos analisados, de modo que, para o autor, as consoantes iniciais que mais favorecem as sílabas centralizadas nos ditongos são /h/, /l/, /r/, /w/, /m/, /n/.

Nesse ponto, podemos destacar que um dos principais legados postulados pela teoria de base Variacionista é, justamente, a ideia de que fatores sociais e linguísticos influenciam de maneira significativa no processo de variação e mudança linguística. Esse processo se dá pela concatenação de variáveis que, ao se imbricarem, são responsáveis pelo fenômeno variável das línguas. Isto é, o social é capaz de interferir no funcionamento interno das estruturas linguísticas, assim como questões inerentes ao próprio sistema das línguas (fonéticas, morfológicas e sintáticas) também fomentam inúmeras discussões e são passíveis de isolamento e análise.

Outro aspecto importante destacado pelo trabalho de Labov, que vale a pena ser discutido nos limites deste trabalho, é a ideia de afirmação identitária manifestada por meio da língua. Ao observar que a economia local estava ligada ao turismo, o autor ressalta a resistência de muitos nativos ao movimento de ocupação irrestrita da ilha. Assim sendo, esses moradores mais conservadores podem ser, *grosso modo*, os responsáveis por manterem determinados traços linguísticos como forma de manutenção de uma identidade local.

Labov acrescenta que há nativos que acreditam que a ilha de *Martha's Vineyard* pertence, sobretudo, aos descendentes das antigas famílias e não aos veranistas responsáveis pelo turismo local. Nesse sentido, parece justo afirmarmos que os moradores parecem usar a língua como forma de reafirmação identitária. Sobre isso, o autor pontua que “o estudo dos dados mostra que a alta centralização de (ay) e (aw) está intimamente correlacionada a expressões de grande resistência as incursões dos veranistas” (LABOV, 1972 [2008], p. 48).

Os principais nativos resistentes aos veranistas são os que habitam nas áreas rurais da ilha alta, em especial na comunidade de *Chilmark*. Para o autor, os *chilmarkenses* se orgulham do seu modo de vida diferenciado e isso se mostra evidente, também, na língua utilizada por eles: “Acho até que nós usamos um tipo de língua inglesa totalmente diferente... pensamos diferente aqui na ilha... é quase uma língua separada dentro da língua inglesa” (LABOV, 2008, p. 48). Assim, como pontua a frase utilizada para exemplificar a sua hipótese, também concordamos que as distinções fonéticas marcadas em *Martha's Vineyard* podem (sim) estar relacionadas a uma luta (inconsciente ou não) pela busca de identidade linguística, social e geográfica.

Essas discussões serão importantes para a nossa análise à medida que servem de base para estudos posteriores que tomam o uso variável do artigo definido como objeto de estudo. Como observamos no capítulo anterior, Campos Júnior (2011) e Oliveira (2018) chamam atenção para o que Labov já havia observado na Ilha de *Martha's Vineyard*: usos linguísticos podem ganhar ênfase à proporção que carregam traços linguísticos identitários de um povo. Uma das hipóteses com a qual trabalhamos nesta pesquisa gira em torno da mesma premissa. Acreditamos que o uso pouco expressivo do artigo definido diante de antropônimos na cidade do Recife (CALLOU e SILVA, 1997) pode ser um traço de manutenção da identidade linguística local. Tais discussões serão retomadas no capítulo 5, quando discutiremos – de maneira mais pormenorizada – os dados da capital pernambucana.

Ainda sobre os estudos labovianos, podemos citar outro trabalho de fôlego realizado pelo autor acerca da estratificação social do (r) em lojas de departamento na cidade de Nova York. A pesquisa, de mesmo nome, se insere na lista dos mais importantes trabalhos realizados pelo teórico sobre o inglês falado em Nova York e, sobretudo, acerca do papel importantíssimo da estratificação social para a Sociolinguística de base Variacionista. O autor ressalta que estudar uma língua em uso demanda do pesquisador um criterioso método de coleta de dados. Assim, uma das maneiras de obter dados mais confiáveis é por meio de entrevistas individuais gravadas, método, aliás, que é adotado até os dias atuais em estudos com linguística de *corpus*.

O objetivo central da pesquisa era compreender como se dava a estratificação social no que diz respeito à pronúncia do (r) e suas variações ocorridas em posição pós-vocálica. Antes da coleta sistemática dos dados, Labov levou em consideração observações preliminares que o fizeram crer que a presença ou ausência do (r) pós-vocálico é um parâmetro importante de estratificação social na cidade de Nova York. Desse modo, o critério de estratificação utilizado foi a observação da fala dos vendedores de três grandes lojas de departamento com *Status* sociais diferentes: *Status* superior – *Saks Fifth Avenue*; *Status* intermediário: *Macy's*; *Status* inferior: *S. Klein*. Como hipótese geral do trabalho, o autor acreditava que quanto maior o *Status* social da loja mais os vendedores iriam usar a forma de prestígio do (r).

A coleta dos dados seguiu uma metodologia específica que visava à obtenção de falas anônimas e casuais, isto é, falas pouco monitoradas:

[...] o entrevistador se aproximava do informante no papel de um cliente que pedia informações sobre um departamento. O departamento em questão ficava no quarto andar. Quando o entrevistador perguntava: 'por favor, onde ficam os sapatos femininos?', a resposta geralmente era: 'Fourth floor' ('Quarto andar') (LABOV, 1972 [2018], p. 70).

Assim, o *corpus* da pesquisa dispõe de um quadro de 264 informantes com aproximadamente 6 horas e meia de gravação. O número de entrevistas por loja não foi uniforme, dado que na *Saks Fifth Avenue* foram ouvidos 68 falantes, 125 na *Macy's* e 71 na *S. Klein*. Confirmando a hipótese levantada por Labov, os resultados abalizaram uma evidente estratificação para o uso do (r) nas três lojas de departamento. Conforme exposto nos dados, os empregados da *Saks*, da *Macy's* e da *S. Klein* obtiveram os respectivos percentuais para o uso do (r): 62%, 51% e 21%. Frente à distinção nítida entre as lojas de *status* inferior e superior, o autor acredita que “a pronúncia do *r* é norma que a maioria dos empregados da loja *Macy's* almejam alcançar [...] Em outras palavras, os empregados de *Saks* têm mais segurança num sentido linguístico” (LABOV, 1972 [2008], p. 73)

Outra questão apontada pelo estudo sugere que a forma menos prestigiada do (r) foi mais utilizada na loja *S. Klein*, de *status* inferior, socioeconomicamente falando, a qual possuía – quando comparada às outras duas lojas de departamento – a maioria de funcionários negros. Esse resultado põe em relevo uma questão importante no que diz respeito à variável raça<sup>32</sup>, controlada na pesquisa. Parece que, *grosso modo*, a posição social ocupada por determinados falantes sugere padrões de fala diferentes, dando margem para a disseminação de preconceitos culturais e linguísticos.

Esse tipo de julgamento de valor ocorre, sobretudo, em casos de fenômenos linguísticos estigmatizados que costumam ser responsáveis por repartições sociais e econômicas, como pontua Lucchesi (2015)<sup>33</sup>. Com essa postura, a questão primordial que queremos exteriorizar aqui é indiscutível: fatores sociais e seus estigmas são capazes de atravessar de maneira latente a língua em uso. No entanto, alguns fenômenos parecem não sofrer esse atravessamento socioeconômico. Em nosso caso específico, estamos lidando com uma forma variável que parece não sofrer essas pressões sociais, a julgar pelo uso facultativo do artigo definido antepondo nomes próprios de pessoas (apontado inclusive, em vários manuais de gramática), isto é, uma variação geográfica que sofre influência do fator região de origem dos informantes (CALLOU E SILVA, 1997).

Além desses fatores apresentados, o autor faz um detalhamento minucioso acerca de outras variáveis extralinguísticas (ocupação e faixa etária) que também exercem influência sob o fenômeno variável investigado. No entanto, não iremos nos aprofundar nessas questões nesse

---

<sup>32</sup> Termo utilizado pelo autor.

<sup>33</sup> “Ao longo dos tempos, a língua tem sido um poderoso instrumento de dominação e construção da hegemonia ideológica que legitima o poder das classes dominantes.” (LUCCHESI, 2015, p. 18)

momento, haja vista que o nosso principal interesse em detalharmos esse estudo reside, justamente, nos métodos de estratificação utilizados por Labov. Assim, podemos inferir que formas diferentes de falar podem se alternar mediante critérios econômicos, sociais, geográficos e identitários – como os aludidos no estudo anterior sobre o inglês falado na ilha de *Martha's Vineyard* – e que cabe ao linguista desenvolver um método de coleta e de análise que seja capaz de controlar todas as variáveis, linguísticas e extralinguísticas, que influenciam a variação.

De modo sumário, concordamos que, entre as inúmeras contribuições dos estudos labovianos, sua empreitada metodológica para observação da língua falada é, sem dúvidas, um dos mais importantes subsídios para os estudos que surgiram *a posteriori*. Pensar em um modelo teórico-metodológico que desse conta da heterogeneidade das línguas fez com que uma gama de pesquisadores se interessasse pelo estudo e pela sistematização de fenômenos variáveis nas línguas naturais. No Brasil, não foi diferente, inúmeras pesquisas de cunho sociolinguístico ganharam força e acabaram solidificando a teoria no país<sup>34</sup>.

Na próxima seção, nos ocuparemos de expor algumas considerações acerca da Sociolinguística Variacionista no Brasil.

### 3.2 A SOCIOLINGUÍSTICA NO BRASIL: BREVES INCURSÕES

Nesta seção, iremos expor algumas discussões acerca do desenvolvimento da Sociolinguística no Brasil, desde a sua estabilização até os dias atuais. Daremos enfoque, sobretudo, às observações expostas por Freitag (2016) acerca da Sociolinguística no âmbito nacional e do que a autora pontua como um modelo teórico produzido – sob uma série de circunstâncias – especialmente no país. Nosso objetivo nessa discussão é apresentar os desdobramentos da teoria e suas principais contribuições para os estudos da linguagem no país.

Apoiando-se nos estudos de Scherre e Roncarati (2008), Freitag (2016) assume que a gênese da teoria no Brasil se deu a partir de uma disciplina ministrada por Anthony Julius Naro, em 1976, no curso de mestrado da PUC- RJ. No entanto, a autora ressalta que, mesmo se inserindo muito cedo no cenário da Sociolinguística, estudos importantes – como o livro *Padrões Sociolinguísticos* (1972) de Labov – só ganharam traduções muitos anos depois, na primeira década dos anos dois mil.

---

<sup>34</sup> Conforme Freitag (2016, p. 445), “a Sociolinguística de orientação variacionista é uma das áreas de pesquisa mais abrangentes e produtivas no Brasil.”

Conforme pontua Freitag (2016), a Sociolinguística só se consolida, enquanto área de pesquisa no Brasil, a partir de uma série de circunstâncias que envolviam a implementação de um programa de incentivo a educação para jovens e adultos – O Movimento Brasileiro de Alfabetização (Mobral) e a articulação de alguns pesquisadores brasileiros. Essa mobilização possibilitou um engajamento mais enérgico dos teóricos que já tinham afinidade com a teoria, conforme exposto a seguir:

[...] Segundo relata Anthony Naro, em entrevista a Sebastião Votre, Claudia Roncarati e Rebeca Nascimento (2008), Miriam Lemle o procurou para articularem a proposta sob encomenda do Mobral que tinha que cumprir dois requisitos: ser pesquisa e ter o uso do computador. Naro disse que sabia o que fazer, pois tinha assistido anteriormente a uma palestra de Labov, em Chicago; entrou então em contato com ele, que indicou Gregory Guy, que veio ao Brasil realizar coleta de dados e trazer os programas computacionais para o projeto. E assim a sociolinguística variacionista se torna uma linha de pesquisa no Brasil. Não foi algo planejado; foi, segundo Naro, um “conjunto de circunstâncias” (e interesses) que levou a este empreendimento. (FREITAG, 2016, p. 452).

Na visão de Freitag (2016), a Sociolinguística no Brasil possui características específicas, especialmente, no que diz respeito a sua inclinação para trabalhar com constituição de bancos de dados extensos. Dentre os principais projetos que adotaram esse modelo de operacionalização, podemos citar alguns dos mais significativos: o Projeto Censo da Variação Linguística do Estado do Rio, O projeto a Variação Linguística Urbana da Região Sul do Brasil (Varsul) e projeto Norma Urbana Culta – NURC. A autora admite que esse tipo de abordagem trabalha com um número relevante de dados, contribuindo de forma significativa para os estudos linguístico no Brasil<sup>35</sup>.

Indo além nas suas considerações históricas, Freitag (2016) questiona quais são os motivos que fizeram com que a Sociolinguística no Brasil priorizasse as interfaces teóricas ao invés de trabalhar com os papéis mais sociais da língua. Para autora, esse movimento pode ser causado por dois fatores específicos: “produção de dados para alimentar modelos teóricos de língua e a força de políticas de financiamento e nucleação da pesquisa na pós-graduação.” (FREITAG, 2016, p. 453). Ao justificar sua hipótese, a autora assevera que a produção científica no país está inteiramente ligada às instituições de ensino superior e que grande parte dos cursos não se alinha a uma determinada língua teórica específica, fazendo com que os

---

<sup>35</sup> “A abordagem a partir de bancos de dados sociolinguísticos trouxe subsídios para a descrição do português brasileiro, com a padronização da amostragem e coleta de dados, que permite, de certa forma, a comparação de resultados, e, assim, traz contribuições para uma norma brasileira, com descrições sociolinguísticas em interface teórica tanto com abordagens formais (como sociolinguística paramétrica, gramáticas em competição), como com abordagens funcionais (sociofuncionalismo, gramaticalização)” (FREITAG, 2016, p. 453)

bancos de dados sociolinguísticos nutram pesquisas linguísticas de outra natureza. Dessas diretrizes, resulta o fato de a Sociolinguística ser conceituada, muitas vezes, como uma metodologia e não como uma vertente teórica – como, de fato, é<sup>36</sup>.

Acreditamos que o processo metodológico da Sociolinguística é importante na execução da teoria – como bem pontuamos na seção anterior acerca dos trabalhos de Labov. No entanto, concordamos com as postulações de Freitag (2016) ao sustentarmos a premissa de que a teoria não se limita a isso. Admitir assertivas que resumam a teoria a um mero modelo metodológico de coleta – utilizado para fomentar discussões linguísticas de outra natureza – é incongruente para concepção teórica que adotamos nessa pesquisa. Dessa forma, ressaltamos que a associação entre os fatores linguísticos e sociais é fundamental para o processo de análise estrutural de uma língua e que essa correlação se torna imprescindível para a explicação dos fenômenos variáveis inerentes às línguas naturais.

Sobre os rumos da Sociolinguística no Brasil, a autora ressalta que o que tem sido feito até os dias atuais “uma trajetória de construção de bancos de dados” (FREITAG, 2016, p. 456) deve ser mantido. A autora propõe que esse tipo de mapeamento sociolinguístico que visa à geração de dados (fomentando descrições e testagens linguísticas) deve ser chamado de documentação sociolinguística. Além disso, Freitag (2016) discute a importância de aproximar a teoria Variacionista às questões de ensino de língua materna, fator que ainda é motivo de polêmica e preconceitos linguísticos no país. Sobre isso, autora ressalta a necessidade de: “[...] uma postura ativa dos pesquisadores da Sociolinguística em mobilizar os resultados de seus estudos para reverter-los em tecnologia social: materiais didáticos e de difusão no suporte, reconhecimento e valorização das variedades.” (FREITAG, 2016, p. 451)

De maneira mais prognóstica, Freitag (2016) chama atenção para o que ela conceitua como senso de ativismo linguístico indispensável aos (socio)linguistas, isto é, um resgate ao lado mais social da sociolinguística. Para tanto, a teórica ressalta a importância de processos mais amplos de coleta de dados que levem em consideração critérios mais abrangentes de estratificação social. A autora sugere, sobretudo, uma mudança de foco necessária para os atuais

---

<sup>36</sup> Como afirma Naro, *apud* Freitag (2016, p. 454) “A Sociolinguística é sobretudo linguística. A Teoria da Variação é uma metodologia. Tem que ter por detrás uma teoria linguística. [...] a Sociolinguística é essa: o que ela faz é estudar a gramática em ação. Mas ela tem que ter uma teoria da gramática por detrás, para orientar as grandes perguntas. Também existe um ramo da Sociolinguística mais voltado para a Sociologia. Mas, pelo menos até agora [2008], no Brasil, ninguém se interessou por ele [...] É, por exemplo, no papel das mulheres e dos homens na sociedade ou em um grupo qualquer. Lá nos Estados Unidos, o foco era principalmente nos afro-americanos. Mas existe todo um ramo da Sociolinguística mais voltado para questões sociais do que para questões linguísticas... Mas, aqui no Brasil, não temos esse campo de estudos.” (VOTRE & RONCARATI 2008. p. 14)

pesquisadores, haja vista a necessidade de sobrelevar questões – cada vez mais latentes – de identidade, estilo e atitude linguística.

Nesse sentido, verificamos que, mesmo sendo uma área de pesquisa relativamente antiga no Brasil, pesquisadores atuais se mobilizam no intuito de apontar os rumos pelos quais a pesquisa sociolinguística deve caminhar no país.

Dando continuidade às nossas discussões, iremos apresentar, na próxima seção, os principais conceitos que tangenciam a Sociolinguística Variacionista. Nessa linha argumentativa, daremos crédito, também, às questões relacionadas à ética nas pesquisas linguísticas, principalmente no cenário nacional.

### 3.3 SOCIOLINGUÍSTICA VARIACIONISTA: CONCEITOS BÁSICOS

Nesta seção, iremos expor os conceitos basilares da Sociolinguística Variacionista, a fim de sistematizar tais diretrizes teóricas com o plano metodológico e analítico utilizado no curso deste trabalho. Como já foi possível verificar, iniciamos nossas discussões regressando aos postulados iniciados por William Labov (2008) [1972], dado o fato de que foi ele quem, mais efetivamente, reiterou a ideia de indissociabilidade entre linguagem e sociedade. Num segundo momento, apresentamos algumas discussões acerca da pesquisa sociolinguística no Brasil. Assim sendo, nosso objetivo nesta seção é relacionar as diligências, expostas até aqui, às concepções discutidas por alguns manuais teóricos e pesquisadores(as) da Sociolinguística Variacionista: Tarallo (2007); Santos de Paula (2011); Coelho *et al* (2012) Freitag (2014, 2018) e Mollica (2015).

Fazendo coro com as ideias de Mollica (2015), concordamos que a Sociolinguística Variacionista é um dos ramos investigativos da Linguística que focaliza o estudo da língua enquanto prática social. Ou seja, a língua sem juízos valorativos, aquela que se manifesta e se molda no interior das comunidades de fala. Em outras palavras, dizemos que esse domínio das ciências linguísticas se volta, sobretudo, para um método de análise que prioriza a correlação inegável entre aspectos sociais e aspectos linguísticos. Sobre isso, Santos de Paula (2011) acrescenta que:

a Sociolinguística é uma ciência interdisciplinar que se ocupa das relações entre língua e sociedade, por meio de uma investigação empírica e foi definida como campo de estudo [...] em meados da década de 1960. Embora investigando aspectos mutuamente interligados entre estrutura linguística e sociedade, a sociolinguística enfoca,



prioritariamente, os usos linguísticos concretos, especialmente os de caráter heterogêneo (SANTOS DE PAULA, 2011, p. 2).

Assim sendo, assumimos, nos limites deste texto, que a Sociolinguística evidencia a natureza heterogênea dos dados linguísticos e busca delimitar métodos de coleta e análise que deem conta da pluralidade que é a língua falada no âmago das comunidades.

Como já pontuamos anteriormente, Labov, em seu livro *Padrões Sociolinguísticos* (2008 [1972]), admite que a sua resistência em aderir ao termo Sociolinguística deve-se ao fato de que, para ele, não há uma teoria linguística que seja efetivamente bem-sucedida sem levar em consideração o fator social. Nessa linha argumentativa, podemos inferir que, conforme as premissas labovianas, a língua não deve ser investigada dissociada de seu contexto social. Logo, o interesse básico da Sociolinguística é estudar a língua mediante dados reais, que sejam capazes de corresponder a um conjunto representativo de falantes da comunidade investigada. Desse modo, esse modelo teórico elege **a variação linguística** como seu principal objeto de estudo, já que, segundo Mollica (2015), a variação é “um princípio geral e universal, passível de ser descrita e analisada cientificamente” (MOLLICA, 2015, p. 10).

Ao considerar a variação linguística como ponto principal das análises realizadas na sociolinguística, é válido circunscrever que há, no domínio desses estudos, diversas denominações para o termo variação<sup>37</sup>. Contudo, de maneira geral, podemos induzir – em conformidade com Coelho *et al* (2012) – que a variação linguística é o processo pelo qual duas formas coexistentes podem ocorrer com mesmo valor referencial dentro de uma comunidade de fala. Além disso, de maneira hegemônica, os teóricos desse segmento concordam que a variação é um processo ao qual todas as línguas naturais estão sujeitas, uma vez que o processo de variação é inerente às línguas “e não compromete o bom funcionamento do sistema linguístico nem a possibilidade de comunicação dos falantes” (COELHO *et al* 2016, p. 16).

Fenômenos variáveis são constantes e, por isso, dizemos que as formas linguísticas equivalentes de se falar – sejam elas de qualquer natureza: fonética, morfológica, sintática – concorrem no âmbito de todas as línguas naturais. Essas formas alternativas de articular a língua são chamadas de **variantes linguísticas**: “as diversas maneiras de se dizer a mesma coisa em

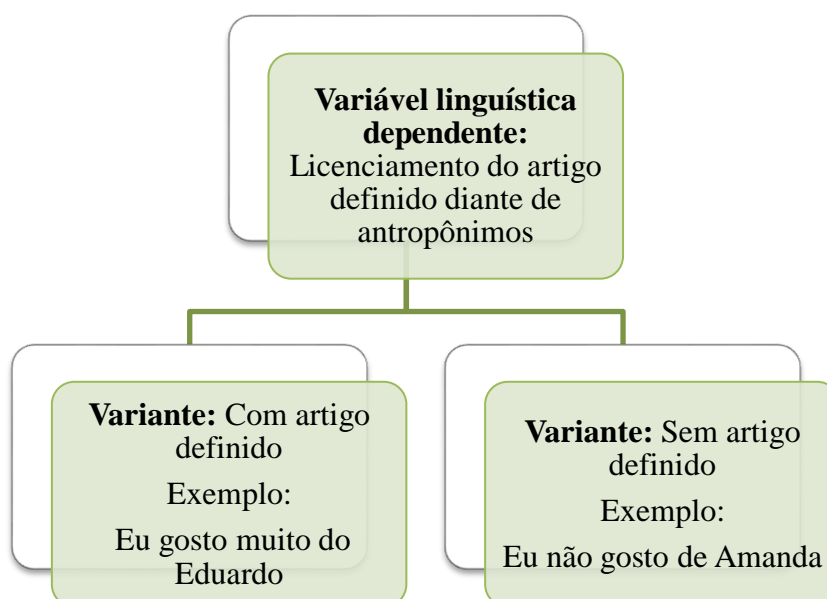
---

<sup>37</sup> De acordo com Mollica (2015), o fenômeno da variação pode ocorrer em segmentos distintos; nos eixos diatópicos e diastráticos: “no primeiro, as alternâncias se expressam regionalmente, considerando-se os limites físico-geográficos; no segundo, elas se manifestam de acordo com os estratos sociais, levando-se em conta fronteiras sociais.” (MOLLICA, 2015, p. 12)

um mesmo contexto, e com um mesmo valor de verdade. A um conjunto de variantes dá-se o nome de **variável linguística**” (TARALLO, 2007, p. 8).

Assim sendo, como o objetivo central desta pesquisa visa a analisar uma possível alternância entre o uso variável do artigo definido antepondo nomes próprios de pessoas na capital pernambucana, dizemos que o licenciamento do artigo definido, no contexto citado, é a nossa **variável linguística dependente** e as duas variantes possíveis para o nosso fenômeno investigativo são: (i) realização do artigo definido e (ii) ausência do artigo definido. Para uma melhor compreensão, o esquema, a seguir, ilustra melhor essa assertiva:

**Esquema 2** - Variável linguística dependente e variantes possíveis para o fenômeno.



**Fonte:** Exemplos retirados do nosso *corpus* de análise

Conforme expomos, as formas alternativas que constituem um fenômeno variável são teoricamente denominadas de **variável linguística dependente**, dado que seu comportamento está sujeito à influência de variáveis independentes (linguísticas e/ou extralinguísticas). Dessa maneira, cabe admitir que o emprego das variantes não é aleatório, já que ele pode ser influenciado pelo grupo de fatores que constituem as **variáveis** de natureza social e/ou estrutural. Dizendo de outro modo, estas “variáveis independentes ou grupo de fatores podem ser de natureza interna ou externa à língua e podem exercer pressão sobre seus usos, aumentando ou diminuindo sua frequência de ocorrência” (MOLLICA, 2015, p. 11). Nessa esteira de pensamento, somos capazes de inferir que a Sociolinguística assume – conforme já observamos anteriormente nos trabalhos de Labov – que os mecanismos estruturais da língua

se correlacionam, efetivamente, aos fatores que são extrínsecos a ela; de tal modo, essa relação deve ser controlada por meio da observação do grupo de variáveis adotadas pelo pesquisador.

Conforme mencionado, ao observarmos uma língua em uso, somos apresentados a múltiplas realidades linguísticas, isso acontece pelo fato de estarmos lidando com um conjunto variado de falantes que carregam características específicas, especialmente no que diz respeito aos seus perfis sociais. Como indica Tarallo (2007), todos os artifícios que desempenham condicionamento de uma variante linguística em detrimento de outra devem ser considerados como agentes da variação. Sobre esse mesmo aspecto, Santos de Paula (2011) também faz coro com os demais autores ao consentir que a variação está intrinsecamente ligada à heterogeneidade dos usos linguísticos e que:

existem condições que favorecem o uso variável, coexistindo regras em competição que possibilitam a ocorrência de uma ou outra variante. Mesmo nesta situação, o uso é regulado por regras que atuam no sentido de favorecer ou não determinadas formas em seus múltiplos contextos. Determinar a relação entre os usos linguísticos de uma determinada comunidade e os fatores sociais que os desencadeiam é uma das tarefas da pesquisa sociolinguística, sendo fundamental identificar e quantificar os fatores que são internos ao sistema ou externos ao mesmo. (SANTOS DE PAULA, 2011, p. 2).

É prudente admitirmos que, de acordo com as premissas sugeridas pelo autor, cabe ao sociolinguista investigar quais são os grupos de fatores que podem agenciar o fenômeno da variação. Isto é, conforme pontuamos no início desse capítulo, reconhecer os aspectos sociais que influenciam no objeto de estudo é tão importante quanto o conhecimento dos aspectos estruturais que ele exerce no sistema da língua.

Dizendo de outro modo, podemos inferir que, assim como pontua Mollica (2015), “agentes como escolarização alta, contato com a escrita, com os meios de comunicação em massa, nível socioeconômico alto, origem social alta” (MOLLICA, 2007, p. 27) também são fatores comumente associados ao fenômeno da variação linguística. Esse grupo de fatores elencados pela autora está, geralmente, ligado às formas mais prestigiadas, de modo que a língua é rotulada, *grosso modo*, como culta e coloquial. Essa última, no entanto, acaba sendo vista como desprestigiada, já que se associa erroneamente, aos grupos sociais menos favorecidos.

Apesar desses julgamentos equivocados em relação aos usos linguísticos prestigiados e desprestigiados na língua, Tarallo (2007) lembra que não há algo exclusivo nas variantes que as definam como “boas, ruins, corretas ou incorretas”. Trata-se, então, de uma questão de

atitude linguística<sup>38</sup> dos membros da comunidade de fala. Essa postura de conferir julgamento de valor às formas em uso fortalece a concepção inverídica de que existe maneira “certa ou errada” de se falar uma língua. Citando um caso análogo, o autor vale-se dos estudos de Labov<sup>39</sup> para observar que a ausência do fonema /r/, por exemplo, é estigmatizada socialmente no inglês falado em Nova Iorque (já que a sua presença está relacionada a pessoas de *status* social mais elevado), mas que, no segmento oposto, a presença do mesmo /r/ pós-vocálico é estigmatizada socialmente, como uma variante desprestigiada na Inglaterra. Desse modo, o autor ressalta a ideia de que cada comunidade de fala seleciona, ainda que intuitivamente, suas formas de prestígio, e estas podem coincidir, ou não, com as formas gramaticalizadas do fenômeno:

[...] as variantes de uma comunidade de fala encontram-se sempre em relação de concorrência: padrão *vs.* não-padrão; conservadora *vs.* inovadora; de prestígio *vs.* Estigmatizadas. Em geral, a variante considerada padrão é, ao mesmo tempo, conservadora e aquela que goza de prestígio sociolinguístico na comunidade [...] nem sempre, no entanto, a coincidência entre os três pares é verificada. (TARALLO, 2007, p. 12)

Parece justo afirmarmos que o arcabouço teórico-metodológico da Sociolinguística Variacionista disponibiliza contribuições imprescindíveis às ciências linguísticas, haja vista seu caráter metodológico de lidar com o fenômeno da variação. No âmbito desses estudos, somos capazes de delimitar, também, as marcas linguísticas que individualizam a fala de determinadas comunidades linguísticas. Se associarmos tais premissas ao fenômeno que estamos investigando nesta dissertação (o licenciamento do artigo definido antepondo nomes próprios de pessoas), somos levados a crer que a variável social *região de origem*, por exemplo, exerce influência para o comportamento do fenômeno variável no Brasil<sup>40</sup>, fator que ratifica a importância de controlar variáveis desse tipo.

Ao concebermos a importância dos fatores sociais para estratificação dos nossos informantes, optamos por selecionar algumas variáveis extralinguísticas que, segundo pesquisas já realizadas, são capazes de exercer influência sobre o fenômeno em análise: (i) sexo do informante e (ii) faixa etária. De acordo com Coelho *et al* (2016), a variável social *sexo* é

---

<sup>38</sup> Conceito abordado por Freitag (2016) como um objeto de análise importante para os próximos passos da Sociolinguística no Brasil.

<sup>39</sup> Ver seção 2.1 deste capítulo

<sup>40</sup> Essa variação pode ser considerada, *grosso modo*, enquanto uma variação regional (cf. SILVA (1997); CALLOU e SILVA (1998)).

sempre levada em consideração em estudos sociolinguísticos. Algumas pesquisas<sup>41</sup> apontam que, em geral, mulheres são mais conservadoras na língua, por isso, tendem a ser mais sensíveis às variantes valorizadas socialmente. Ainda de acordo com os autores, a variável faixa etária, que leva em consideração a idade dos informantes, suscita inúmeras discussões, haja vista que é através dela que os pesquisadores conseguem inferir se há um fenômeno de mudança linguística em curso na língua; ou seja, a “mudança linguística implica no processo de substituição gradual de uma forma por outra” (COELHO *et al*, 2016, p. 44), podendo, assim, se manifestar através das faixas etárias.

Na visão de Freitag (2018), os exercícios metodológicos da sociolinguística no Brasil debruçam-se em técnicas de amostragens aleatórias e estratificadas. Esse método fundamenta-se na divisão da população em células sociais, isto é, em grupos relevantes para a pesquisa. No nosso caso de pesquisa, optamos por separar os informantes em células que distinguem o sexo e a idade dos informantes. Nas palavras de Freitag (2018, p. 668), as amostragens aleatórias e estratificadas impõem “confiabilidade e replicabilidade às análises”, uma vez que todos os falantes têm a mesma probabilidade de serem selecionados.

Adjacentes aos fatores externos, Mollica (2007) delega que as forças internas ao sistema linguístico (fonológicas, morfológicas, sintáticas e semânticas) também são atuantes frente à variação. Variáveis dessa natureza estão diretamente ligadas à variável linguística dependente eleita para investigação, isto é, ao selecionarmos um fenômeno de ordem sintática – o licenciamento do artigo definido diante de nomes próprios, por exemplo – dispomos de uma amálgama de fatores sintático-semânticos que podem, ou não, condicionar um uso em detrimento de outro<sup>42</sup>. A título de exemplo, podemos citar algumas das variáveis linguísticas independentes que se mostraram atuantes em trabalhos que já se debruçaram sobre o fenômeno em análise e, por isso, também foram selecionadas por nós no andamento desta pesquisa: Contexto preposicionado, Função sintática, *Status* informacional, Tipo de informação, Referência do antropônimo e Tipologia do antropônimo. Essas variáveis serão caracterizadas no próximo capítulo, quando voltaremos nosso olhar para os processos metodológicos que compuseram esta pesquisa.

---

<sup>41</sup> Em consonância com o autor, Campos Junior (2015) – em estudos sobre o mesmo fenômeno na capital Capixaba – também atenta para o conservadorismo apresentado pelas mulheres que resistem ao uso do artigo definido diante de antropônimos, uma vez que esta parece ser a forma inovadora para a comunidade de fala pesquisada.

<sup>42</sup> Na pesquisa em tela, as formas em variação são: a realização *versus* a não realização do artigo definido diante de antropônimos. Para melhor absorção dessas variáveis, sugerimos a leitura do segundo capítulo desta dissertação.

Outro aspecto importante que se refere à pesquisa sociolinguística – e vem sendo pauta de muitos debates atualmente – diz respeito aos processos éticos na coleta de dados linguísticos que envolvem seres humanos. Conforme aponta Freitag (2014), os procedimentos que orientam as coletas de dados para a constituição de *corpus* (tanto nas ciências humanas e sociais quanto nas pesquisas em Linguística) estão subordinados às mesmas regras que norteiam as ciências médicas. A autora argumenta que esse “sufocamento” (FREITAG, 2014, p. 7) nasceu após a Segunda Guerra Mundial, época em que a comunidade científica internacional concentrou inúmeros esforços para regulamentar os procedimentos científicos envolvendo seres humanos – especialmente após a barbárie dos governos nazistas – acerca disso a autora pontua que:

A decisão de intervir, tutelando os direitos das pessoas submetidas às pesquisas científicas e, principalmente, a necessidade de estabelecer obrigações e responsabilidades para instituições e seus pesquisadores, colocou-se como algo urgente no cenário do pós-guerra. Assim, o Código de Nuremberg (1947), a Declaração Universal dos Direitos Humanos (1948) e a Declaração de Helsinque (1964) surgem, no âmbito do Direito Internacional, como os principais alicerces nos quais se apoiaram as constituições e legislações infraconstitucionais no mundo para fins de regramento ético nas pesquisas envolvendo seres humanos. (FREITAG, 2014, p. 7)

De acordo com Freitag (2014), tais premissas éticas começaram a ganhar força, no âmbito nacional, mediante uma resolução do Conselho Nacional de Saúde, órgão responsável pela normatização dos parâmetros que abrangem as pesquisas com humanos<sup>43</sup>. Todavia, a autora pontua que o modelo altamente criterioso e biomédico das resoluções vem recebendo críticas efetivas dos pesquisadores vinculados às ciências humanas, dado que muitas questões pertinentes à área não foram consideradas. Sobre essas questões, Freitag (2014) destaca a ideia de “ética igual, para pesquisas diferentes” (FREITAG, 2014, p. 7) no Brasil.

A autora também argumenta que grande parte das associações voltadas para o controle das pesquisas alertam que esse tipo de procedimento, excessivamente ligado às ciências médicas, pode dificultar o trabalho dos pesquisadores, sobretudo, em Linguística<sup>44</sup>. Assim,

---

<sup>43</sup>“A via de entrada desses parâmetros ético-normativos foi o Conselho Nacional de Saúde que, primeiramente através da Resolução n. 196, de 10 de outubro de 1996, e após, com a Resolução n. 466, de 12 de dezembro de 2012 – vem estabelecendo os parâmetros éticos da pesquisa envolvendo seres humanos.” (FREITAG, 2014, p.7)

<sup>44</sup>Atualmente, a Associação Brasileira de Antropologia (ABA), com o aval da Associação Nacional de Pós-Graduação e Pesquisa em Ciências Sociais (ANPOCS) e de várias associações representantes das diversas áreas da pesquisa em ciências sociais e humanas, a exemplo da Associação Nacional de Pós-Graduação e Pesquisa em Letras e Linguística (ANPOLL), vem, através de um Grupo de Trabalho constituído também por membros da CONEP/MS, discutindo e delineando o documento que complementará a Resolução n. 466/12 nos aspectos que lhes forem pertinentes.” (FREITAG, 2014, p. 11).

como forma paliativa, conselhos voltados para áreas de conhecimento específicas (Linguística, por exemplo) foram criados e podem avaliar e autorizar a efetivação de projetos de pesquisa. Essa medida facilita a execução das pesquisas; mas, de acordo com Freitag (2014), o ideal seria completar o documento da Resolução com vistas a corrigir todos os aspectos pertinentes às demais áreas de pesquisas.

Como foi possível observar nas discussões propostas até aqui, a maioria dos trabalhos em Sociolinguística caminham em direção à coleta de dados de fala, fator que pressupõe lidar com entrevistas e seres humanos. Freitag (2014) destaca que, entre as muitas questões que compõem essas coletas, a aceitação do participante e a preservação de sua identidade – e dados pessoais – são imprescindíveis para garantir a ética nas pesquisas linguísticas. Todas essas medidas nos levam a crer que os processos que concebem uma pesquisa em Linguística devem ser (sim) guiados por parâmetros éticos e normativos. Todavia, fazemos coro com a autora ao nos questionarmos até que ponto a subordinação às ciências médicas é relevante para as ciências humanas e sociais, haja vista a natureza diferentemente complexa de cada uma das áreas.

Conforme podemos verificar, nossas discussões teóricas apontam para uma vertente específica e fortemente ligada à Sociolinguística Variacionista: a ética na coleta de dados. De modo sumário, é possível perceber que, ao fazermos uma inclinação teórica para as questões metodológicas da Sociolinguística, não deixamos de lado a natureza teórica que tangencia nossa vertente de estudo. De tal modo, as discussões propostas aqui serão retomadas com mais robustez no próximo capítulo desta dissertação, quando detalharemos a metodologia adotada nos cursos dessa investigação.

Endossando as discussões propostas nos limites deste capítulo, concordamos com os autores – citados neste capítulo – ao assumirmos que a língua manifesta-se de maneira heterogênea e a variação, que é inerente a ela, pode ser condicionada por dois grandes grupos de fatores: linguísticos e sociais. À vista disso, conforme propõe Tarallo (2007), a Sociolinguística firma-se através de seu encadeamento entre fatores linguísticos e extralinguísticos; e é só por meio desse paralelismo que podemos obter uma melhor compreensão de como a língua é utilizada pelos falantes e quais são os elementos estruturais que constituem esses usos.

De modo geral, apresentamos alguns dos principais estudos de Labov com o objetivo de descrever os primeiros empreendimentos do autor acerca da Sociolinguística de base Variacionista. Em um segundo momento, expusemos algumas considerações acerca da efetivação dessa área de pesquisa no Brasil. Em seguida, com base em alguns manuais teóricos, delimitamos alguns dos principais conceitos teóricos da área. Por fim, discutimos acerca dos

processos metodológicos e éticos que garantem legitimidade à pesquisa sociolinguística no país.

Assim sendo, esperamos descrever, como base nessas discussões teóricas, os fatores motivadores da variação na comunidade de fala em estudo: Recife. Para tanto, no capítulo seguinte apresentaremos as incursões metodológicas que nortearam nossa pesquisa.



## 4 OS PROCESSOS METODOLÓGICOS

Sabemos que os procedimentos metodológicos que regulamentam uma investigação científica encontram-se sempre vinculados a uma vertente teórica específica; com a Sociolinguística não é diferente. Na percepção de Naro (2015, p. 25), os processos metodológicos da Teoria da Variação representam “uma ferramenta poderosa e segura que pode ser usada para o estudo de qualquer fenômeno variável nos diversos níveis e manifestações linguísticas”. Em outras palavras, podemos assegurar que a Sociolinguística Variacionista dispõe de uma base investigativa já consolidada nas ciências linguísticas, especialmente no que concerne à sua robustez metodológica, haja vista que – conforme pontua Tarallo (2007) – teoria e método conservam em si uma associação tão lógica, quanto necessária.

É adequado salientarmos, também, que as discussões metodológicas da Sociolinguística não se limitam tão somente ao que iremos expor aqui. Existem outras múltiplas abordagens que também se mostram relevantes, dada a natureza particular de cada fenômeno linguístico em variação. Por esse motivo, conforme pontua Naro (2015), cabe ao linguista delimitar quais são os fatores e ferramentas mais relevantes para cada tipo de fenômeno em investigação. Dito isso, descreveremos, no curso deste capítulo, os métodos utilizados para verificarmos como se dá o licenciamento do artigo definido em sintagmas antroponímicos na cidade do Recife.

Em nosso delineamento analítico, priorizamos trabalhar com a metodologia quantitativa que elege a estatística como um dos principais parâmetros de análise. Como já observamos, a Sociolinguística Variacionista é uma vertente científica e empírica que se dedica a observar dados reais de fala, por isso, o trabalho com números se mostra extremamente relevante. Todavia, não desconsideraremos a análise qualitativa dos dados, já que essa também se mostra imprescindível para a interpretação do *corpus* de pesquisa. Isto é, “o progresso da ciência linguística não está nos números em si, mas no que a análise dos números pode trazer para nosso entendimento das línguas humanas” (NARO, 2015, p. 25).

No capítulo anterior, observamos os principais conceitos que circunscrevem a teoria Variacionista para que, nesta seção, possamos retomar, na prática, algumas dessas concepções teóricas. Isso posto, este capítulo se ocupará de discutir alguns dos processos metodológicos que constituíram a nossa pesquisa, a saber: o contexto sócio-histórico da comunidade de fala; os procedimentos éticos; a seleção dos informantes; a coleta dos dados; a transcrição e a

codificação dos dados analíticos; as variáveis e as hipóteses levantadas para esse estudo; e, por fim, o uso do programa GoldVarbX como aporte analítico dos nossos dados.

#### 4.1 A COMUNIDADE DE FALA

Nesta seção, dedicaremos especial atenção ao contexto sócio-histórico da comunidade de fala em estudo: Recife. Para tanto, levaremos em consideração alguns dados oficiais do IBGE sobre a capital pernambucana. Além disso, de modo mais amplo, revisitaremos os postulados apresentados pelo Atlas Linguístico de Pernambuco – **ALiPE** (2013), a fim de delinear as principais características culturais e históricas do estado de Pernambuco, assim como aspectos importantes para a formação da capital pernambucana.

Conforme aponta o ALiPE (2013), a cidade do Recife teve sua formação iniciada logo após a divisão do território – até então português – em capitanias hereditárias. Ainda de acordo com o ALiPE (2013), a criação da Vila Capital de Olinda, em 1537, viabilizou o crescimento de uma pequena zona portuária que servia para o desembarque de grandes embarcações. Por questões geográficas, essas navegações não eram capazes de atracar na região da Vila Capital de Olinda, que hoje compreende o município de Olinda. Dessa maneira, nas imediações desse pequeno porto, deu-se início a cidade vizinha do Recife.

Assim como ocorreu em várias regiões do território brasileiro, a história de Pernambuco também é marcada pela grande exploração das riquezas naturais, pela massiva opressão e genocídio dos povos indígenas e originários e pelo tráfico de pessoas negras e escravizadas vindas da África; tal fator pode explicar as múltiplas influências culturais presentes no estado até os dias atuais.

Nesse mesmo alinhamento histórico de invasão e colonização, as cidades de Olinda e Recife – ainda sob o domínio da coroa portuguesa – foram invadidas por Holandeses que tinham o intuito de aumentar os seus domínios na região nordeste. Tal empreitada designou o cargo de comandante da colônia a Maurício de Nassau, que desempenhou atividades de melhoria econômica nas duas cidades. A retomada do território para Portugal, por sua vez, só aconteceu em 1645, conforme dados presentes no ALiPE (2013).

A história de Pernambuco seguiu sendo marcada por grandes revoluções, como é o caso da Revolução Pernambucana que se deu, posteriormente, contra os abusos da coroa portuguesa em terras brasileiras, já no ano de 1817. Tal movimento é considerado como um dos precursores da posterior Independência do Brasil.

Além desses feitos históricos, Recife também é conhecida como uma das mais antigas capitais do Brasil – fator que é apontado por alguns linguistas (CALLOU e SILVA, 1997) como sendo responsável por um possível traço de conservadorismo linguístico presente no português falado no estado.

A cidade do Recife, capital pernambucana, situa-se na região nordeste do país e é banhada pelo Oceano Atlântico. Estabelecida na Mesorregião Metropolitana, Recife localiza-se, especificamente, na microrregião Metropolitana do Recife e tem cerca de 1.653.461 habitantes, segundo dados oficiais do IBGE de 2020<sup>45</sup>.

Segundo o ALiPE (2013), grande parte da população pernambucana e, consequentemente, recifense resulta de uma vasta mistura de etnias que é responsável por diversos aspectos culturais e sociais da cidade. Conhecida também pelo turismo, Recife atrai visitantes de todo mundo, especialmente pelas famosas praias e pelas diversas manifestações culturais, como é o caso do Carnaval, por exemplo, conhecido por ser um dos maiores e mais tradicionais do Brasil.

No que diz respeito à Literatura e aos aspectos culturais, Recife é berço de diversos escritores e intelectuais importantes para a história do país. Destacamos alguns artistas admiráveis e de renome internacional: João Cabral de Melo Neto, Ascenso Ferreira, Chico Science, Manoel Bandeira, Gilberto Freyre, dentre muitos outros.

Por fim, no que diz respeito à economia local, segundo dados oficiais do IBGE em 2020, o PIB do Recife é de R\$ de 31.994,3. Conforme dados disponíveis do site da Prefeitura Municipal do Recife<sup>46</sup>, a economia da capital se baseia, sobretudo, na prestação de serviços, comércio, atividades especializadas em tecnologia da informação e comunicação (polo digital), economia criativa, atividades médico-hospitalares e turismo. Além disso, a cidade está situada numa região de polos industriais; com destaque, também, para a zona portuária de Suape, um dos principais portos da região Nordeste do país.

Isso posto, na próxima seção, discorreremos sobre os procedimentos éticos que regulamentaram esta pesquisa.

---

<sup>45</sup> <https://www.ibge.gov.br/cidades-e-estados/pe/recife.html>. < Acessado em 19 de janeiro de 2021, às 13:39>

<sup>46</sup> <http://www2.recife.pe.gov.br/index.php>. < Acessado em 19 de janeiro de 2021, às 13:45>

## 4.2 OS PROCEDIMENTOS ÉTICOS: SUBMISSÃO DO PROJETO DE PESQUISA AO COMITÊ DE ÉTICA EM PESQUISA

Conforme observarmos no capítulo anterior, os procedimentos que norteiam as pesquisas com seres humanos são guiados por uma série de regulamentações éticas que legitimam o teor científico desses estudos. De acordo com Freitag (2014, p. 9), um ponto importante no avanço dessa área “foi a criação da Comissão Nacional de Ética em Pesquisa (CONEP/MS), uma instância colegiada, de natureza consultiva, deliberativa, normativa, educativa, independente e vinculada ao Conselho Nacional de Saúde”. Assim sendo, esse domínio é responsável por analisar e guiar todas as cláusulas legais que envolvem as pesquisas com seres humanos no Brasil.

Ainda de acordo com Freitag (2014), é válido observarmos que, além da criação do CONEP/MS, foram instituídos, no Brasil, Comitês de Ética locais nas Instituições de Ensino Superior que realizam pesquisas envolvendo seres humanos. Esses órgãos são os responsáveis por analisar e aprovar (ou desaprovar) os projetos de pesquisa submetidos para apreciação nas instituições. É pertinente observarmos que, mesmo dispondo desses comitês de ética locais, os projetos devem ser submetidos por meio da Plataforma Brasil<sup>47</sup>: um ambiente eletrônico unificado, disponibilizado pelo Governo Federal para coordenar a submissão de todos os projetos de pesquisa enviados aos Comitês de Ética do país. Todo o processo de submissão ao Comitê de Ética segue uma rigorosa metodologia que será pormenorizada ao longo desta seção.

Para iniciarmos o processo de submissão, foi necessária a criação de um cadastro na Plataforma Brasil; na qual disponibilizamos, de maneira digitalizada, todos os dados pessoais e institucionais necessários para autenticar o vínculo dos(as) pesquisadores(as) envolvidos(as) com suas respectivas Instituições de Ensino. No caso da nossa pesquisa, especificamente, o Comitê de Ética em Pesquisa (CEP) – da Universidade Federal de Pernambuco (UFPE) – disponibiliza o Manual do Pesquisador<sup>48</sup> que se incumbe de apresentar, de maneira mais didática, todos os procedimentos burocráticos que envolvem a submissão do projeto de pesquisa.

---

<sup>47</sup> Endereço digital: <http://plataformabrasil.saude.gov.br/>

<sup>48</sup> Endereço digital: <https://www.ufpe.br/documents/619819/801156/ManualPesquisador.pdf/7805a81f-38ff-4fe0-a0cc-191a8f7c77a5>

Após a criação do cadastro pessoal, foi possível inserir o projeto de pesquisa e os documentos exigidos pelo comitê de ética<sup>49</sup> na plataforma Brasil. Essa fase do procedimento exige que o(a) pesquisador(a) discrimine todas as partes que substanciam o projeto em seções diferentes e rigorosamente especificadas (introdução, metodologia, riscos, orçamento etc.). Posteriormente, foi gerada uma folha de rosto que, depois de ser preenchida pelo(a) pesquisador(a) responsável, foi submetida ao Coordenador geral da Pós-Graduação para ser assinada e carimbada. Por fim, a folha de rosto foi entregue, pessoalmente, no prédio do Comitê de Ética em Pesquisa da UFPE que – após a inspeção de todos os documentos – emitiu um parecer de aceite para análise no dia seis (06) de maio de 2019. Por meio de um novo parecer consubstanciado, o projeto de pesquisa foi aprovado para execução, pelo relator do CEP (UFPE), no dia um (01) de julho de 2019.

É válido evidenciarmos, também, que todos os procedimentos metodológicos que nortearam nosso trabalho foram executados de acordo com o cronograma<sup>50</sup> de pesquisa, submetido e aprovado pelo CEP (UFPE). Isto é, os processos que envolvem a seleção dos informantes e a coleta dos dados só foram iniciados após o consentimento do Comitê de Ética em Pesquisa e seguiram as datas e as documentações regulamentadas pelos(as) pesquisadores(as) envolvidos no ato de submissão do projeto de pesquisa. Nesse sentido, a realização da presente pesquisa obedeceu aos preceitos éticos da Resolução 510/16 do Conselho Nacional de Saúde, que garante a legitimidade de todos os dados que serão apresentados nesta dissertação.

Dito isso, na seção seguinte, nos ocuparemos em descrever os processos que regulamentaram a seleção dos informantes.

#### 4.3 A SELEÇÃO DOS INFORMANTES

Após a escolha da comunidade de fala, outro aspecto imprescindível para a execução de uma análise Sociolinguística é a escolha do número de informantes que irá integrar a pesquisa. Na concepção de Silva (2015), cabe ao pesquisador a responsabilidade de delimitar quais os

---

<sup>49</sup> Carta de Anuência; Currículo Lattes de TODOS os pesquisadores; TCLE maiores de 18 anos e/ou TCLE para pais e responsáveis e/ou TALE de 7 a 18 anos (dependendo da idade dos participantes da pesquisa) ou outros adequados para a população em estudo; Projeto detalhado; Termo de Confidencialidade; Declaração de Vínculo do curso. Todos esses documentos estão anexados nessa pesquisa.

<sup>50</sup> Ver anexo.

critérios de amostragem serão utilizados na seleção dos informantes e, principalmente, quais serão os desdobramentos necessários para que as generalizações (uniformidade dos informantes) sejam observadas na comunidade de fala. Isto é, dada a amplitude populacional de grandes centros urbanos (especialmente uma capital, como é o caso desta pesquisa), o número de informantes deve corresponder a uma parcela representativa de pessoas, uma vez que “se pretende obter uma pesquisa sobre a comunidade e não apenas sobre a amostra específica” (SILVA, 2015, p. 119).

A probabilidade pode ser utilizada, até certo ponto, como aliada para apresentar resultados mais proporcionais ao número de amostras realizadas. Sobre isso, Silva (2015) apresenta alguns critérios importantes para a seleção de informantes; dentre os quais destacamos o *critério de homogeneidade do fenômeno*, que foi utilizado no andamento da nossa investigação. De acordo com a autora, a língua é mais homogênea para “alguns fenômenos do que para outros” (SILVA, 2015, p. 191), e tal critério poderia justificar o uso de amostras menores para fenômenos específicos. Conforme pontuamos anteriormente, o uso pouco expressivo do artigo definido no estado de Pernambuco<sup>51</sup> parece apontar para uma homogeneidade do fenômeno no estado, fator que justifica o fato de termos trabalhado com uma amostra não tão extensa de informantes.

É válido ressaltarmos, também, que parte da coleta idealizada para esta investigação foi inviabilizada, haja vista o cenário pandêmico iniciado em março de 2020<sup>52</sup>. Em virtude disso, houve uma redução significativa no número de informantes que iriam compor o *corpus* inicial desta pesquisa. No projeto de pesquisa preambular submetido ao Comitê de Ética em Pesquisa, seriam coletados dados linguísticos provenientes de 24 informantes nativos de Recife (subdivididos em *sexo* – masculino e feminino – e em três *faixas etárias* distintas). No entanto, em razão da conjuntura atual, trabalhamos com dados de língua falada oriundos de apenas 12 informantes, subdivididos somente em duas faixas etárias: primeira faixa etária (18 a 35 anos) e segunda faixa etária (acima de 35 anos).

Nessa linha argumentativa, assumimos a premissa de que toda comunidade de fala carrega traços particulares capazes de imprimir uma identidade aos seus falantes. De tal modo, acreditamos que a cidade de Recife também possui especificidades linguísticas que a

---

<sup>51</sup> Fator apontado pelas pesquisas expostas na seção 2.3 do capítulo 2.

<sup>52</sup> Em março de 2020, a Organização Mundial da Saúde (OMS) declarou cenário de pandemia para o Covid-19, infecção causada pelo novo coronavírus. Tal infortúnio inviabilizou significativamente a coleta de dados linguísticos prevista para esta pesquisa, uma vez que a recomendação prevista pela OMS é de que a população pratique quarentena absoluta (prática de distanciamento social) até que haja uma campanha de imunização adequada.

caracterizam. No intuito de pormenorizar essas particularidades, obtendo – assim – uma amostra representativa da fala dessa comunidade, adotamos *o método de amostragem aleatória estratificada*<sup>53</sup>. Conforme aponta Silva (2015), esse modelo assegura que uma parcela representativa da comunidade de fala participe da seleção, uma vez que cada informante teve a mesma probabilidade de ser escolhido(a) e inserido(a) na amostra. Assim sendo, nossas Variáveis Extralinguísticas foram distribuídas em células<sup>54</sup> que se organizaram da seguinte maneira:

**Quadro 5 - Estratificação dos informantes**

<b>FAIXA ETÁRIA</b>	<b>SEXO</b>	<b>NÚMERO DE INFORMANTES</b>
18 a 35	Mas.	3
18 a 35	Fem.	3
Acima de 35	Mas.	3
Acima de 35	Fem.	3

**Fonte:** Projeto de pesquisa submetida ao comitê de ética

Como podemos verificar no quadro 5, selecionamos um total de 12 informantes oriundos do município de Recife. Todos os informantes entrevistados carregavam a particularidade de não terem se afastado da localidade pelo período mínimo de um ano. Além disso, já residiam na região por um longo período. Desses 12 informantes (divididos nas duas faixas etárias distintas), 6 pertenciam ao sexo masculino e os outros 6 ao sexo feminino; somando um quadro total de 3 informantes de cada sexo para cada faixa etária.

Como seguimos orientações regulamentadas pelo CEP (UFPE), foi necessária a criação de critérios de inclusão e exclusão dos informantes, com vistas a promover uma amostragem mais representativa dos dados. Esse processo demanda que o(a) pesquisador(a) responsável faça projeções acerca dos nativos que podem ou não colaborar com a pesquisa, levando em conta o perfil social e linguístico dos informantes. Nessa conjuntura, o perfil social<sup>55</sup>

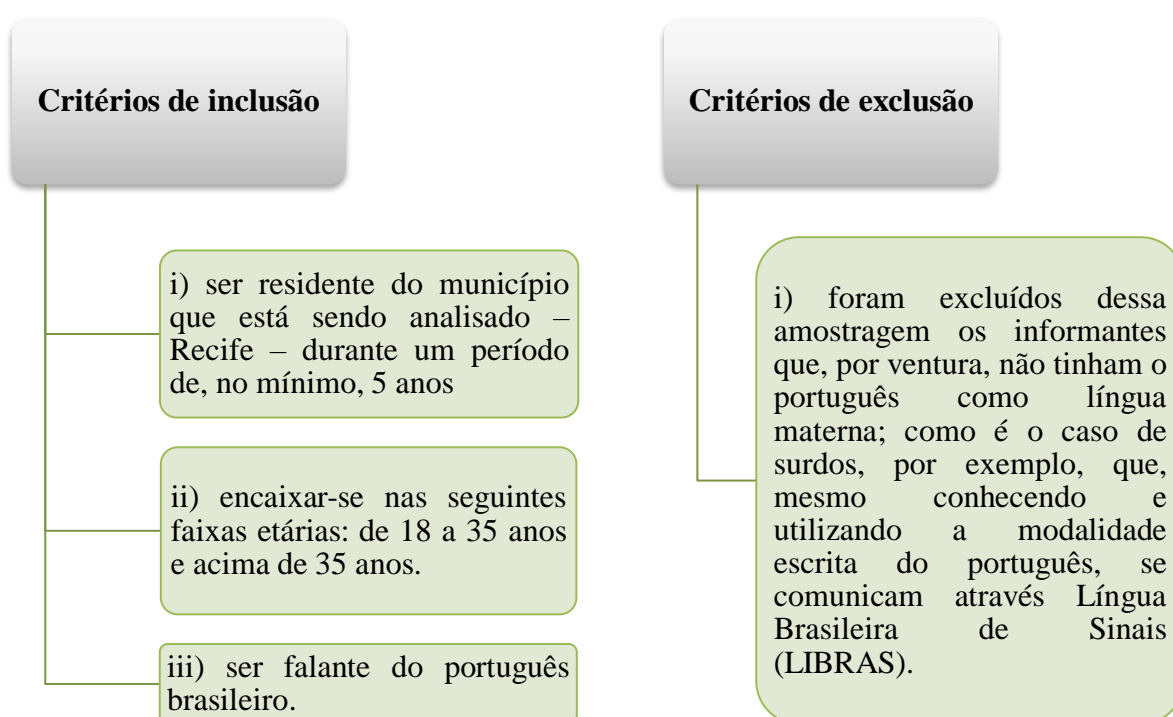
<sup>53</sup> De acordo com Silva (2015, P. 121), a amostragem aleatória estratificada divide a população em células compostas por indivíduos com as mesmas características sociais. Em nosso caso específico, homens e mulheres subdivididos por faixa etária.

<sup>54</sup> Coelho *et al* (2015) consideram que a célula social é um conjunto de indivíduos agrupados pelas mesmas características sociais relevantes para a análise de fenômenos de variação e mudança linguística.

<sup>55</sup> É válido apontarmos que muitos estudos sobre diversos fenômenos linguísticos investigados no PB apontam para a influência da variável escolaridade, contudo, considerando que tal variável não exerceu um papel

selecionado em nossa investigação diz respeito unicamente à faixa etária e ao sexo dos informantes, já que todos deviam residir obrigatoriamente na cidade do Recife. O perfil linguístico, por sua vez, foi selecionado para garantir que todos os informantes que participaram do estudo tivessem o português brasileiro como língua materna. Desse modo, para ser informante dessa pesquisa o(a) entrevistado(a) teria que atender aos seguintes critérios sociais:

**Esquema 3 - Critérios de inclusão e exclusão de informantes**



**Fonte:** Projeto de pesquisa submetido ao comitê de ética

O recrutamento dos participantes voluntários desta pesquisa foi efetivado em forma de convite. Isso indica que, após sinalizarem o interesse em fazer parte do projeto científico, os informantes que aceitaram participar do estudo autorizaram (por meio dos documentos solicitados pelo comitê de ética) suas respectivas contribuições nas seguintes atividades: entrevista e coleta dos dados (via gravação de áudio) e, posteriormente, a análise do *corpus* linguístico.

Como se trata de uma pesquisa de campo, na qual a inquiridora precisou ir até os

---

significativo em alguns trabalhos que já se debruçaram sobre o fenômeno investigado nesta pesquisa (CAMPOS JUNIOR (2011); SEDRINS, PEREIRA e SIQUEIRA (2015); OLIVEIRA, (2018)), a escolaridade dos informantes não foi considerada nesta estratificação, assim como o perfil socioeconômico.



informantes, criamos uma rede de indicações na comunidade estudada. Ou seja, por meio de consultas a pessoas já conhecidas na comunidade de fala, pedimos indicações aos primeiros colaboradores (informantes já conhecidos dos(as) pesquisadores(as) e identificados como nativos das comunidades), criando, assim, uma amostragem representativa de participantes. Além disso, conforme pontuamos anteriormente, os informantes precisavam se encaixar nos critérios de inclusão e exclusão descritos no esquema (3).

É válido destacarmos que os colaboradores foram informados com antecedência (através de ligação prévia) sobre a natureza da pesquisa. Tendo ciência e autorizando a participação, os informantes foram entrevistados individualmente pela inquiridora responsável pelo projeto de pesquisa. Como algumas projeções de riscos aos colaboradores são exigidas pelo CEP (UFPE) elencamos, abaixo, alguns dos possíveis riscos aos quais os participantes estariam, supostamente, expostos:

#### **4.3.1 Riscos**

Os riscos existentes para os participantes na execução desta pesquisa foram amenizados com adoção das seguintes práticas:

- (i) a identidade dos participantes foi preservada, uma vez que trabalhamos com dados de fala que foram transcritos, analisados e, possivelmente, divulgados;
- (ii) a coleta dos dados foi obtida por meio de entrevistas semiestruturadas, realizadas em um local sugerido pelo próprio entrevistado, a fim de possibilitar o conforto do informante;
- (iii) os informantes foram convidados a assinar um termo de consentimento que garantia a aceitação – por parte do colaboradores – ao tratamento que os pesquisadores dariam aos dados;
- (iv) os informantes foram assegurados, antes das entrevistas, que poderiam desistir da colaboração a qualquer momento, sem nenhum dano para o próprio informante e/ou para a pesquisa.

Assim sendo, na subseção seguinte, damos continuidade à descrição dos nossos procedimentos metodológicos, apresentando algumas discussões acerca da nossa coleta de dados.

#### 4.4 A COLETA DOS DADOS

Como bem sabemos, a Sociolinguística Variacionista investiga a língua vinculada à sua conjuntura social de (re)produção, isto é, “enfoca, prioritariamente, os usos linguísticos concretos, especialmente os de caráter heterogêneo” (SANTOS, 2011, p. 3). Dizendo de outro modo, é papel do Sociolinguista analisar a maneira como a língua é empregada por seus falantes, em situações reais de língua falada<sup>56</sup>. De acordo com Santos (2011), um dos pressupostos básicos do estudo Variacionista é garantir que o aparente “caos”, observado em uma análise superficial do fenômeno linguístico, seja regulado por meio de procedimentos metodológicos norteadores. Nesse sentido, Labov (2008 [1972]) aponta uma alternativa para que haja a sistematização dos dados a serem analisados. Segundo o autor, um *corpus* pode ser instituído por meio de entrevistas gravadas<sup>57</sup> com os informantes da comunidade de fala pesquisada, com o intuito de viabilizar o manejo dos dados coletados.

Segundo Silva (2015), para executar uma pesquisa que demanda coleta de dados linguísticos, é indispensável ter um arcabouço metodológico criteriosamente esquematizado, uma vez que imprevistos podem vir a aparecer. Na percepção da autora, existem três tipos de interações que podem ser adotadas pelos pesquisadores para facilitar o contato com os informantes: “interações livres, entrevistas e testes” (SILVA, 2015, p. 124). Conforme observado em nosso recorte metodológico, optamos por trabalhar com entrevistas semiestruturadas que priorizavam uma interação guiada pela inquiridora, mas espontânea por parte dos informantes.

Como propõe Tarallo (2007), é importante que a inserção dos(as) pesquisadores(as) na comunidade de fala seja feita através de terceiros, ou seja, de pessoas já devidamente aceitas e familiarizadas com a comunidade. Por isso, buscamos adequar os nossos comportamentos sociais e linguísticos aos da comunidade investigada, uma vez que esse procedimento tenta minimizar o efeito negativo da presença de pesquisadores sobre “o comportamento sociolinguístico natural da comunidade” (TARALLO, 2017, p. 21). A adesão desse tipo de precaução se deve ao fato de prezarmos, essencialmente, por dados espontâneos. Além do mais,

---

<sup>56</sup> “A língua falada a que nós temos referido é o veículo linguístico de comunicação usado em situações naturais de interação social, do tipo comunicação face a face. É a língua que usamos em nossos lares ao interagir com os demais membros de nossas famílias...” (TARALLO, 2007, p. 18)

<sup>57</sup> “As entrevistas consistem na gravação da conversa entre o pesquisador e o falante e devem girar em torno de um assunto previamente planejado pelo pesquisador, sendo indispensável que a conversa seja informal ou casual para que o colaborador da pesquisa não se sinta intimidado ou observado, de forma que o vernáculo seja registrado” (SANTOS, 2011, p. 10)

em consonância com Santos (2011), acreditamos que a coleta de dados envolve questões intrínsecas aos colaboradores da pesquisa e, para garantir uma confiabilidade aos dados levantados, os informantes devem estar confortáveis e minimamente condicionados.

Para que essa fase se consolidasse, organizamos um questionário<sup>58</sup> com vistas a nortear as entrevistas realizadas com os informantes. Nossas entrevistas foram direcionadas da maneira mais informal possível, uma vez que a nossa intenção era obter uma fala não monitorada, pois, como já citamos – e em concordância com Labov (2008 [1972], p. 83), “(...) nosso objetivo é observar o modo como as pessoas usam a língua quando não estão sendo observadas”. Somando todas as doze (12) entrevistas realizadas, obtivemos um total de 3 (três) horas e 40 (quarenta) minutos de áudios provenientes dos informantes, com interferência de perguntas e diálogos conduzidos pela inquiridora e pelos informantes.

Após concederem suas entrevistas, os informantes foram instruídos a responderem uma ficha social<sup>59</sup>, confeccionada antecipadamente<sup>60</sup>, que continha perguntas de cunho individual: dados profissionais, dados pessoais, contato com os meios de comunicação, redes de relacionamento, atividades de lazer, viagens, entre outras questões. Mesmo com todas as Variáveis Extralinguísticas já delimitadas, com esses inquéritos tínhamos acesso a dados particulares de cada informante; pois, como afirma Santos (2009, p. 70), “para a Sociolinguística, o social não pode estar separado da língua; há, além dos linguísticos, fatores externos à língua que a influenciam”. Dessa maneira, a consulta às fichas sociais facilitou o nosso trabalho na hora de traçar um perfil mais pormenorizado dos informantes que fizeram parte da pesquisa.

As entrevistas foram gravadas durante o período de janeiro a março do ano 2020, com a autorização dos participantes que assinaram o Termo de Consentimento Livre e Esclarecido (TCLE)<sup>61</sup>. Esse termo garante, de certa maneira, a validação da pesquisa, posto que é através dele que os informantes asseguram que estão de acordo com os critérios da pesquisa e com os seus fins analíticos. Foi garantido aos informantes que, caso optassem, eles teriam amplo direito à desistência. Além disso, ratificamos que o produto final desta coleta (a pesquisa em questão)

---

<sup>58</sup> O modelo de questionário utilizado no andamento das entrevistas encontra-se no apêndice deste trabalho.

<sup>59</sup> “A ficha social contém dados referentes à idade, grau de instrução, local de nascimento, religião, sendo ainda possível perguntar se o falante morou em outro lugar e por qual período ou outras informações que o pesquisador julgar relevantes.” (SANTOS, 2011, p. 13),

<sup>60</sup> O modelo de ficha social utilizado no andamento das entrevistas encontra-se no apêndice desse trabalho.

<sup>61</sup> O modelo de Termo de Consentimento Livre e Esclarecido (TCLE) utilizado nesta pesquisa encontra-se no apêndice deste trabalho. Além disso, foi preparado com base nas adequações exigidas pelo comitê de ética. Disponível em: <<https://www.ufpe.br/cet>>

estaria à disposição, para leitura ou consulta, de todos os participantes que ajudaram consideravelmente a construí-lo.

Vale ressaltar que todos os dados coletados na comunidade de fala fazem parte do banco de dados dos(as) pesquisadores(as) e podem ser consultados (com a autorização dos coordenadores e por meio de uma nova submissão ao comitê de ética), *a posteriori*, para fins de investigação linguística. A seguir, detalhamos, com base nas exigências do CEP (UFPE), os dados pessoais da pesquisadora responsável, acerca do armazenamento do *corpus* linguístico.

#### 4.4.1 Armazenamento dos dados coletados

Os dados coletados nesta pesquisa (dados de língua falada e questionário de identificação pessoal – ficha social) foram armazenados em computador pessoal, e-mails e HD externo sob a responsabilidade da pesquisadora Marília Adrielle Siqueira de Oliveira pelo período de, no mínimo, 5 anos.

Concluída a descrição acerca de fase de coleta e armazenamentos dos dados, passamos para a transcrição e seleção dos dados analisados. Na subseção seguinte, descrevemos de maneira pormenorizada esta etapa da pesquisa.

#### 4.5 TRANSCRIÇÃO E CODIFICAÇÃO DOS CONTEXTOS SELECIONADOS

Após a gravação das entrevistas e a seleção dos contextos de análise, a próxima fase da pesquisa diz respeito à transcrição dos dados de fala. Na concepção de Paiva (2015, p. 135), “o objetivo básico da transcrição é transpor o discurso falado, da forma mais fiel possível, para registros gráficos mais permanentes”, isto é, o manejo com os dados se torna mais funcional mediante a transcrição, uma vez que analisar os dados somente oralmente seria incongruente. Ainda de acordo com a autora, toda e qualquer transcrição de dados linguísticos pressupõe uma teoria que a norteia. Isso indica que essa etapa da pesquisa pode ser considerada, até certo ponto, como uma análise prévia dos dados, especialmente se consideramos que ela já aponta para o nível linguístico a ser investigado (fonético, sintático etc.), sobre isso Paiva (2015) ressalta que:

Mesmo quando se procura fazer uma transcrição que possa ser útil para trabalhos futuros e diversificados, assume-se uma postura teórica. E, além disso, é a orientação teórica do pesquisador e seus objetivos que modelam previamente um conjunto de convenções (um sistema de transcrição) que norteará a transposição dos registros orais para uma forma gráfica. Esse sistema de convenções se faz necessário para garantir

um mínimo de consistência no processo de transcrição de dados de fala. (PAIVA, 2015, p. 135).

Desse modo, Santos (2011, p. 13) também pontua que “seríamos, com efeito, incapazes de estudar a língua falada somente através de sua oralidade”, por isso, os nossos dados de fala foram materializados mediante uma transcrição autêntica dos áudios coletados. Por se tratar de uma variação de ordem sintática, elegemos para a nossa análise a transcrição ortográfica. Ainda de acordo com Santos (2011), grande parte dos projetos linguísticos empregam a transcrição grafemática ou ortográfica como recurso para sistematizar a língua falada, todavia, salientamos que o tipo de transcrição depende da natureza particular de cada fenômeno<sup>62</sup>. Nessa mesma linha argumentativa, Paiva (2015) assume que o grau de detalhamento da transcrição se relaciona com os objetivos preestabelecidos pelo(a) pesquisador(a) responsável.

No que diz respeito à seleção dos dados, extraímos para transcrição apenas os casos em que presenciamos a aparição do fenômeno, isto é, nos dedicamos apenas aos contextos de antropônimos (ambiente analítico em que estudamos o comportamento do artigo definido) e, por esse motivo, as entrevistas não foram transcritas integralmente. Essa fase exigiu um tempo considerável no andamento da nossa pesquisa, a julgar pelo manuseio pormenorizado que tivemos com os dados<sup>63</sup>. Paiva (2015) ressalta que a transcrição dos dados analíticos é, sobretudo, uma convenção “entre aquilo que percebemos e aquilo que reconstruímos pela interpretação” (PAIVA, 2015, p. 146). Isto é, a subjetividade que reside em cada fenômeno linguístico e a autonomia de cada pesquisador também devem ser ponderadas no andamento do trabalho.

Nessa linha argumentativa, no andamento das nossas transcrições, optamos por utilizar o recurso de digitação por voz, disponível na versão online do *Google docs*. Essa modalidade de criação de documentos – acessível para usuários das contas *Google* – permite que a digitalização das sentenças seja feita por meio da leitura, em voz alta, dos contextos analíticos selecionados. Tal fator otimizou o processo de transcrição, a julgar pela praticidade e eficácia do processo<sup>64</sup>.

---

<sup>62</sup> Alguns recursos especiais de notação podem também ser utilizados, como, por exemplo, “a fonética/fonologia, pode-se utilizar da transcrição fonética, e, neste caso, é recomendada a utilização do alfabeto fonético internacional.” (SANTOS, 2011, p. 14)

<sup>63</sup> Dedicamos 2 (dois) meses, no curso de nossa pesquisa, para a seleção e transcrição dos dados analíticos.

<sup>64</sup> É importante ressaltarmos que todo o processo de digitação por voz foi revisado, assegurando a autenticidade dos contextos transcritos.

Com a transcrição do *corpus* finalizada, passamos à etapa da codificação dos dados analíticos. Como utilizamos o programa computacional GOLDBARBX no processamento e análise dos nossos dados, o sistema de codificação se torna essencial para a obtenção de resultados estatísticos, uma vez que tal algoritmo trabalha com a leitura de códigos. A etapa de codificação não só antecede a análise propriamente dita, como é responsável por regular a organização dos dados a serem analisados. De acordo com Santos e Vítório (2011, p. 9), “é preciso atribuir símbolos não só para a variável dependente, mas também para todos os fatores arrolados como condicionantes do fenômeno em estudo”. Ou seja, cada letra do código gerado corresponde a um fator das variáveis independentes selecionadas, além de referenciar, obrigatoriamente, um dos fatores da variável dependente estudada<sup>65</sup>. Vejamos, a seguir, no quadro 6, o sistema de códigos<sup>66</sup> elaborado para a rodagem da pesquisa em tela:

**Quadro 6 - Sistema de codificação para a qualificação**

<b>Variável dependente</b>	
s = Sem artigo	c = Com artigo
<b>Variáveis independentes extralinguísticas</b>	
SEXO:  F= Feminino M = Masculino	FAIXA ETÁRIA:  1 = Primeira faixa etária 2 = Segunda faixa etária
<b>Variáveis independentes linguísticas</b>	
CONTEXTO PREPOSICIONADO: p = com preposição q = sem preposição	FUNÇÃO SINTÁTICA:  f (= sujeito) m (= adjunto) y (= Comp. Nominal) e (= Antitópico/aposto) t (= tópico)

<sup>65</sup> Em nosso caso, a variável dependente é o licenciamento do artigo definido diante de antropônimos; e os fatores são a realização do artigo *versus* a não realização do artigo.

<sup>66</sup> Em nosso sistema de códigos, tomamos por base a mesma proposta assumida por Pereira (2017, p. 97). No entanto, como controlamos a variável tipologia do antropônimo – não considerada por Pereira (2017) – alguns códigos não foram os mesmos adotados pelo autor.

	d (= Objeto Direto) i (= Objeto Indireto) w (= predicativo do sujeito)
<i>STATUS</i> INFORMACIONAL:  a (= Informação nova) v (= Informação antiga)	TIPO DE INFORMAÇÃO:  o (= compartilhada/famoso) n (= não compartilhada/não famoso)
REFERÊNCIA DO ANTROPÔNIMO:  R (= real) P (= personagem)	TIPOLOGIA DO ANTROPÔNIMO:  A(= Prenome) B(= Sobrenome) C(= Agnome) D(= Apelido (ou alcunha, ou cognome)) G(= Hipocorístico) H(= Pseudônimo) I(= Codinome) J (= Heterônimo) K (= Nome artístico (e nome de palco) L(= Nome religioso) N(= Nome social) O(= Nome de urna) P(= Nome parlamentar)

**Fonte:** Projeto de pesquisa submetido ao comitê de ética

Como selecionamos, 9 (nove) variáveis, incluindo a variável dependente binária desta pesquisa (que reúne os fatores com artigo e sem artigo), o nosso sistema de codificação foi disposto no programa computacional com o código antecedendo a sua respectiva sentença (cf. 34). Esse tipo de disposição foi selecionado porque a leitura dos códigos é feita da esquerda para a direita, respeitando o espaçamento entre a sentença e o código.

(35) (sF1panRiA Eu gosto de Sofia

Como é possível verificar no exemplo (35), o programa reconhece para o processamento dos dados válidos apenas o material que se encontra após o parêntese. Nessa configuração, as letras são dispostas mediante ausência de espaçamento o que as caracterizam como matéria passível de leitura pelo programa, como o código em si. Já os espaços em branco – utilizados após a codificação – indicam que a sentença a qual o código se refere não deve ser processada pelo *Software*, uma vez que sua exibição é inserida no programa apenas para facilitar o manejo dos dados pelo pesquisador.

Outro fator que merece ser destacado nesta seção diz respeito às análises sintáticas realizadas após a transcrição dos contextos analíticos. Como é possível verificar no nosso sistema de codificação (cf. quadro 6) e na seção (3.6.2) desta dissertação, selecionamos a função sintática dos antropônimos como uma das variáveis linguísticas independentes a serem controladas durante o processamento dos dados. Essa atividade foi realizada antes do processo de codificação dos dados e cada uma das 470 sentenças que abarcavam sintagmas antroponímicos foi analisada separadamente.

Isso posto, discutiremos, na seção seguinte, as variáveis selecionadas para a análise preliminar dos dados.

#### 4.6 VARIÁVEIS SELECIONADAS

Mesmo já tendo levado em consideração as incursões apresentadas no capítulo anterior acerca dos conceitos de variáveis linguísticas e extralinguísticas, podemos acrescentar ainda que – incorporada a cada uma destas variáveis – existem os fatores que são, evidentemente, capazes de atuar como condicionadores da variação. De acordo com Mollica (2017), os fatores linguísticos correspondem às características sistêmicas da língua, isto é, aquelas inerentes ao sistema linguístico<sup>67</sup>; já os fatores de ordem extralinguística, de modo diferente, referem-se às especificidades sociais, particulares e contextuais de cada indivíduo<sup>68</sup>.

Na subseção anterior, apresentamos os códigos (alusivos aos fatores de cada variável) utilizados para a codificação, para que nesta seção pudéssemos apresentar as variáveis adotadas no andamento desta pesquisa.

##### 4.6.1 Variáveis extralinguísticas

O quadro 7, a seguir, ilustra as variáveis extralinguísticas selecionadas e os seus respectivos fatores:

---

<sup>67</sup> “No conjunto de variáveis internas, encontram-se os fatores de natureza fono-morfo-sintática, os semânticos, os discursivos e os lexicais” (MOLLICA, 2007, p. 11).

<sup>68</sup> Com base em Mollica (2007), citamos, respectivamente, exemplos das especificidades mencionadas: etnia, sexo; escolarização, classe social; grau de formalidade, tensão discursiva.



**Quadro 7** - Variáveis e fatores extralinguísticos

<b>Variáveis extralinguísticas</b>	<b>Fatores extralinguísticos</b>
SEXO	Masculino Feminino
FAIXA ETÁRIA	Primeira faixa etária (7 – 10) Segunda faixa etária (20 – 39) Terceira faixa etária (acima de 50 anos)

**Fonte:** Projeto de pesquisa submetido ao comitê de ética

Sobre a faixa etária dos informantes, Oliveira (2018) destaca que, mesmo não sendo apontada pelo programa GoldVarbX como sendo uma das variáveis significativas, os falantes mais jovens carregavam tendência de serem mais inovadores na língua e, por isso, usavam mais a variante com artigo definido do que os falantes da faixa etária mais avançada. De tal modo, vamos considerar que tal hipótese também se repetirá nos dados da capital pernambucana.

No que tange à variável sexo dos informantes, Campos Junior (2011), em seus estudos sobre o emprego do artigo definido na capital Capixaba, observou que as mulheres usam menos o artigo do que os homens na cidade do Espírito Santo. Nesse sentido, o autor argumenta que, como postulado por Silva (1996), ainda permanece desconhecido o comportamento inverso entre os sexos, no entanto, ressalta que tal distinção merece ser investigada. Assim sendo, resolvemos controlar tal variável com a hipótese de que, assim como exposto por Campos Júnior (2011), mulheres realizam menos o artigo definido diante de antropônimos.

Com base nos estudos mencionados, acreditamos que as variáveis extralinguísticas escolhidas podem influenciar o comportamento do fenômeno em análise. Por isso, fez-se necessário o enquadramento dos informantes na estratificação mencionada acima. A seguir (3.6.2), discutiremos acerca das variáveis linguísticas que, conforme estudos que investigam a variação do artigo definido em sintagmas antroponímicos, podem influenciar no comportamento do fenômeno.

#### 4.6.2 Variáveis linguísticas

Do mesmo modo que fatores externos influenciam a língua falada no âmbito das comunidades, fatores inerentes à estrutura linguística também podem induzir e/ou condicionar determinados usos. Em vista disso, elegemos para análise as seguintes variáveis linguísticas: informação compartilhada, *status* informacional, contexto preposicionado, referência do antropônimo, função sintática do antropônimo e tipologia do antropônimo. Para melhor ilustrarmos tais apontamentos, apresentamos, a seguir, as variáveis linguísticas selecionadas:

**Quadro 8** - Variáveis e fatores linguísticos

Variáveis linguísticas	Fatores linguísticos
CONTEXTO PREPOSICIONADO	Com preposição Sem preposição
FUNÇÃO SINTÁTICA	Sujeito Adjunto Complemento Nominal Aposto Tópico Objeto Direto Objeto Indireto Predicativo do sujeito
<i>STATUS</i> INFORMACIONAL	Informação nova Informação antiga
TIPO DE INFORMAÇÃO	Compartilhada/famoso Não compartilhada/não-famoso
REFERÊNCIA DO ANTROPÔNIMO	Real Personagem
TIPOLOGIA DO ANTROPÔNIMO	Prenome Sobrenome Agnome Apelido (ou alcunha, ou cognome) Hipocorístico

	Pseudônimo Codinome Heterônimo Nome artístico (e nome de palco) Nome religioso Nome social Nome de urna Nome parlamentar
--	---

**Fonte:** Projeto de pesquisa submetido ao comitê de ética

A variável *informação compartilhada*, considerada por Pereira (2017), diz respeito ao nível de conhecimento que os membros da entrevista (inquiridor e entrevistado) têm do referente. Callou e Silva (1997) lançam mão da hipótese de que, quanto mais conhecida for a pessoa de quem se fala, menos se empregará o artigo definido, ou seja, o fato de o nome próprio referenciar uma figura de domínio público tenderia a diminuir a necessidade do determinante, o que anularia a presença do artigo nestes contextos. Contudo, Pereira (2017), ao analisar os dados do município de Carnaíba-PE, apontou que a realização do artigo definido se mostrou mais recorrente quando a informação era compartilhada (com referentes famosos) do que quando a informação não era compartilhada. Desse modo, tal variável foi considerada para a análise dos nossos dados, com a hipótese de que, assim como no município de Carnaíba, o fator *informação compartilhada* condiciona a presença do artigo na comunidade de fala investigada por nós.

Já a variável *Status Informacional*, que leva em consideração se o antropônimo foi inserido pela primeira vez ou retomado no discurso, se mostrou estatisticamente relevante nos estudos de Callou e Silva (1997). As autoras observaram que o fator *informação nova* estimulou constantemente a presença do artigo, tanto em possessivos quanto em patronímicos. Em face disso, com a hipótese de que o *fator informação nova* favorece a realização do artigo diante de antropônimos, tal variável também foi levada em consideração na análise dos nossos dados.

Com base nos estudos já citados de Callou e Silva (1997), notamos que a variável que leva em consideração se o *contexto é regido ou não por preposição* se mostrou muito atuante em seus trabalhos. As autoras indicam que a presença de algumas preposições favorece constantemente o uso do artigo definido diante de antropônimos. Desse modo, controlamos tal variável com a hipótese de que a presença da preposição atua como elemento linguístico condicionante para a realização do artigo definido antepondo nomes próprios.

A variável *referência do antropônimo* leva em consideração a natureza do referente, ou seja, se o indivíduo que é referenciado no discurso é uma pessoa real ou um ser de natureza

ficcional (personagem de novelas, filmes ou desenhos animados). Em seus estudos, Pereira (2017) destaca que o fator personagem também favorece de maneira considerável a presença da partícula determinante, desse modo, em concordância com o autor, tal variável foi considerada na análise dos nossos dados, com a hipótese genérica de que referentes de natureza ficcional condicionam a realização do artigo.

Callou e Silva (1997) e Pereira (2017) também levaram em consideração *a função sintática* dos antropônimos. Os autores atestaram que essa variável exerce influência no comportamento dos dados, desse modo, tal variável merece ser controlada para sabermos se as relações sintáticas desempenhadas pelos antropônimos condicionam ou não a presença do artigo definido.

Outro critério de ordem linguística selecionado no andamento desta pesquisa foi a variável *tipologia dos antropônimos*. Mesmo não sendo considerada, até então, em trabalhos de cunho sociolinguístico, acreditamos que a demarcação tipológica dos antropônimos – abalizada por Amaral e Seide (2020) – apresenta um compilado substancial de tipos de antropônimos que devem ser controlados (enquanto fatores linguísticos) mediante codificação (cf. quadro 6). A ideia de verificar a tipologia dos antropônimos é fortalecida, sobretudo, quando percebemos, nos exemplos extraídos do nosso *corpus* de análise<sup>69</sup>, que tipos específicos de antropônimos são mais licenciados com determinantes do que outros. Essas informações serão mais bem delimitadas no nosso capítulo de análise.

Assim sendo, acreditamos que determinados tipos de nomes próprios de pessoas podem ser, ou não, mais sensíveis ao uso marcado dos determinantes. Dadas as variáveis e seus respectivos fatores, além das hipóteses selecionadas para cada uma delas, passamos, por fim, para a primeira fase da nossa análise: as rodagens teste e o processamento dos dados.

#### 4.7 *GOLDVARBX*: PROCESSAMENTO E RODAGEM DOS DADOS

Conforme Santos e Vitória (2011), uma análise quantitativa, como a que adotamos nessa investigação, envolve um número elevado de dados e, por isso, a adesão de programas computacionais subsidia o(a) pesquisador(a) da área com o tratamento estatístico dos dados

---

<sup>69</sup> Observar os exemplos exposto do capítulo um (1) dessa dissertação: O objeto de estudo. Na seção Antroponímia brasileira, apresentamos uma sequência tipológica de antropônimos apresentada por Amaral e Seide (2020) e exemplificada com os dados do nosso *corpus*.

sociolinguísticos. De tal modo, adotamos como suporte analítico da nossa pesquisa o programa GOLDVARBX<sup>70</sup> (versão atualizada do programa VARBRUL<sup>71</sup>) que nos auxiliou durante o estágio de manipulação estatística dos dados, elegendo, assim, o agrupamento de fatores mais expressivos para a variação.

Após a codificação dos nossos dados, executamos duas rodadas teste, nas quais foi preciso excluir a variável *sexo*, haja vista que estamos lidando com uma análise preliminar. Posteriormente, concretizamos mais duas rodadas, das quais extraímos todos os resultados que substanciam esta pesquisa até o momento.

Por meio da quantificação, o programa gerou dígitos precisos e específicos, isto é, além de revelar os fatores mais significativos para a variação, a amálgama de números gerados facilitou o processo de criação de gráficos e, posteriormente, o procedimento de descrição analítica dos dados.

Isto posto, no próximo capítulo, apresentaremos – com base nas postulações expostas até aqui – a descrição e análise quantitativa dos nossos dados parciais.

---

<sup>70</sup> O programa “seleciona o(s) grupo(s) de fator(es) mais significativo(s) para realização da variação linguística, que deverão ficar com um peso relativo acima de 0.51, o que demonstraria a efetividade na aplicação da regra variável. Valores abaixo de 0.51 podem ser interpretados como desfavorecedores a sua aplicação. Deste modo, são identificados os fatores linguísticos e sociais que favorecem ou não o uso de uma ou outra variante, que está diretamente relacionada à variável dependente em análise” (SANTOS, 2011, p. 17).

<sup>71</sup> De acordo com Guy e Zilles (2007, p. 105), “um conjunto de programas computacionais de análise multivariada, especificamente estruturado para acomodar dados de variação sociolinguística” (*apud* SANTOS e VITÓRIO, 2011, p. 2).

## 5 DESCRIÇÃO E ANÁLISE QUANTITATIVA DOS DADOS

Tomando por base as discussões realizadas no curso deste trabalho, este capítulo se incumbirá de apresentar os resultados gerais da nossa pesquisa. Conforme observamos no capítulo anterior, utilizamos o programa *GoldVarbX* como ferramenta analítica que auxilia, sobretudo, na obtenção de dados estatísticos mais precisos. A principal atribuição do *Software*, nessa fase da pesquisa, foi a de designar se as variáveis escolhidas foram significativas para um comportamento variável do fenômeno e (se sim) quais os seus principais desdobramentos linguísticos e extralinguísticos, especialmente no que concerne à atuação da variável linguística dependente. Além disso, utilizamos o programa *Visual Studio Code* que viabilizou a leitura dos resultados, já que o seu pacote de ferramentas – voltadas para a área de desenvolvimento e tecnologia – inclui um controle mais funcional sobre dados numéricos e linguagens de programação complexas.

Salientamos que nosso processo de análise será conduzido mediante retomadas às hipóteses – gerais e específicas – elaboradas para a pesquisa, de modo amplo, e para cada variável em particular. É válido destacarmos, também, que o retorno ao capítulo teórico será feito, sempre que necessário, com vistas a promover uma saturação fundamental entre os capítulos (teórico e analítico) deste trabalho. Ressaltamos, por fim, que os resultados estatísticos apresentados aqui estarão agregados aos contextos analíticos específicos, isto é, às variáveis selecionadas para o andamento desta pesquisa. Nessa conjuntura, será possível descrevermos tanto as variáveis significativas, quanto aquelas que forem descartadas pelo *GoldVarbX*.

Para uma melhor compreensão das discussões que iremos propor nos limites deste capítulo, nossos tópicos de análise estão estruturados de modo a garantir uma leitura ampla dos resultados. Na seção 5.1, expomos os resultados globais de aplicação e não aplicação do artigo definido em sintagmas antroponímicos na cidade do Recife. Na seção 5.2, apresentamos as variáveis apontadas como significativas pelo *GoldvarbX*. De modo correlato, discutiremos – na seção 5.3 – as variáveis apontadas como não significativas pelo programa. Na seção 5.4, iremos apresentar o cruzamento de algumas variáveis. Já na seção 5.5 iremos discutir algumas questões relevantes acerca da antroponímia na capital pernambucana.

## 5.1 VARIÁVEL LINGUÍSTICA DEPENDENTE: RESULTADOS GLOBAIS DE APLICAÇÃO E NÃO APLICAÇÃO DO ARTIGO DEFINIDO EM SINTAGMAS ANTROPONÍMICOS NA CIDADE DO RECIFE

Na execução deste trabalho, levamos em consideração – conforme já mencionamos – dados provenientes de doze informantes selecionados mediante os critérios expostos no capítulo 4. Na rodagem oficial dos dados no programa *GoldVarbX*, tivemos um somatório de 734 ocorrências, nas quais o objeto de análise (sintagma antroponímico) estava expresso uma única vez. Isto é, nos casos em que os informantes elencavam mais de um antropônimo em uma mesma sentença (ex.: *João e o Daniel foram grandes amigos meus*<sup>72</sup>), consideramos duas ocorrências nessa única sentença. De tal modo, esse tipo específico de sentença<sup>73</sup> recebeu dois códigos distintos na quantificação dos dados, um para cada ocorrência de antropônimo: [João]; [o Daniel].

A variável linguística dependente investigada nesta pesquisa de dissertação é do tipo binária e compreende as seguintes variantes: realização do artigo definido *versus* a não realização do artigo definido diante de antropônimos. Nos exemplos<sup>74</sup> a seguir, ilustramos melhor essa assertiva:

### ➤ Realização do artigo definido diante de antropônimos

(36) “**O Guilherme Boulos** fez uma visita ao terminal do Barro” (inf. MRMT. F. 2).

(37) “Eu gosto muito do **Joey Tribbiani**” (inf. HDM. F. 2).

(38) “Atuação **da Mary Streep** é incrível” (inf. VMO. M. 2).

(39) “**O Dérick**, eu detesto ele” (inf. JPS. M. 2).

---

<sup>72</sup> Exemplo real retirado do nosso *corpus* de análise (inf. HDM. F. 2).

<sup>73</sup> Ao todo, o nosso *corpus* analítico apresentou 13 sentenças contendo mais de um antropônimo, conforme exposto no exemplo.

<sup>74</sup> Dados extraídos do banco de dados da pesquisa.

➤ Não realização do artigo definido diante de antropônimos

(40) “Eu tinha uma colega, **Letícia**, a gente era muito amiguinha...” (inf. LSM. F. 2)

(41) “**Bethânia** é tudo!” (inf. MSON. M. 2)

(42) “Eu levo **Lizandra** comigo.” (inf. MRMT. F. 2)

(43) “**Lula** é muito sensato” (inf. JPS. M. 2).

Como é possível observar, os exemplos expressos acima sugerem que a variação do artigo definido pode ser verificada em várias posições sintáticas no português falado na capital pernambucana. Seja na posição de sujeito (cf. 36, 41 e 43), na posição intrassentencial de objeto direto (cf. 42) e de objeto indireto (cf. 37 e 38) e até deslocado à esquerda da sentença (cf. 39), o artigo pode ser omitido sem grandes prejuízos semânticos nas sentenças observadas acima.

Ao nos debruçarmos sobre esses casos, percebemos que os moradores da comunidade de fala Recife empregam as duas variantes possíveis para esse fenômeno: tanto a realização quanto a não realização do artigo definido antepondo nomes próprios. No entanto, indo além nessas observações, verificamos que os falantes da capital pernambucana – mesmo tendo nítida preferência pela variante sem artigo – também utilizam a variante com artigo, especialmente se compararmos com os dados do interior do estado<sup>75</sup>. Apesar disso, os resultados corroboram com a hipótese geral desta pesquisa, de que falantes pernambucanos têm (sim) preferência pela não marcação dos determinantes em sintagmas antroponímicos.

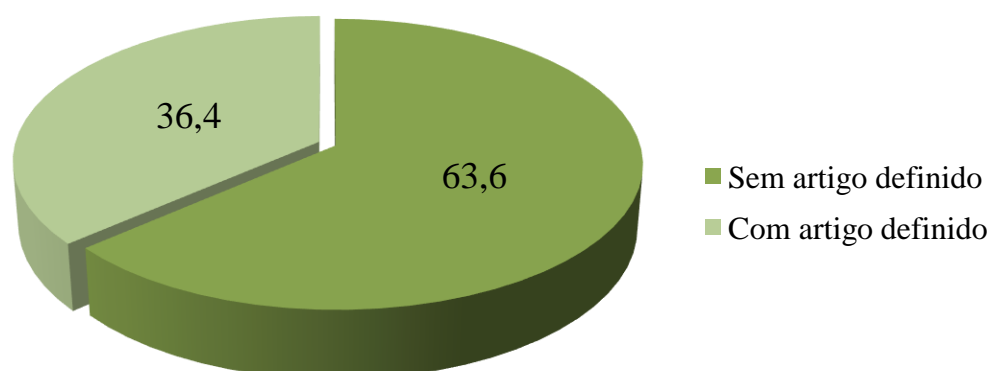
A análise dos dados de língua falada nos permitiu observar que houve um percentual global de 63,6% (467/734) de não realização do artigo, ao passo que o número de ocorrência de sintagmas antroponímicos com artigos definidos foi de 36,4% (267/734). Essa comparação fica mais evidente no gráfico (1), a seguir:

---

<sup>75</sup> Expressos em Pereira (2017) e Oliveira (2018).



**Gráfico 1** - Panorama geral de realização e não realização do artigo definido diante de antropônimos



**Fonte:** Autoria própria

Tomando por base o gráfico acima, percebemos que o não cumprimento da partícula determinante antepondo nomes próprios de pessoas, como já era previsto na hipótese desta pesquisa, é mais acentuado na cidade do Recife. Contudo, é pertinente destacarmos que alguns aspectos podem apontar para um uso ligeiramente expressivo das duas formas na comunidade de fala; o número elevado de sentenças codificadas (734) confirma a preferência pelos dois usos na cidade. Isto é, ao nos depararmos com um total de 267 (36,4%) ocorrências do artigo em sintagmas antroponímicos *versus* 467 (63,6%) de não ocorrências, conseguimos perceber que as variantes seguem, de fato, em variação contínua na capital pernambucana. Em outras palavras, podemos supor que os falantes recifenses também se mostram bastante sensíveis à variante com artigo. Essas explanações ficarão mais nítidas quando apresentarmos os resultados percentuais de acordo com a faixa etária dos informantes, por exemplo.

No decorrer desta pesquisa, mostraremos evidências de que o uso da variante com artigo (cf. 36 a 39) parece sofrer forte influência de três variáveis linguísticas importantes – Tipo de informação, Referência do antropônimo e Tipologia do antropônimo – na comunidade de fala; sobretudo porque essas variáveis legitimam a realização do artigo apenas em contextos bastante específicos. Além disso, as duas variáveis sociais (sexo e faixa etária) – selecionadas para essa empreitada analítica – também apresentaram uma sugestiva relevância na análise dos dados percentuais.

Esses apontamentos revelam que o padrão da comunidade parece, de fato, optar pela variante sem artigo. Contudo, de acordo com as nossas investigações, parece justo inferirmos,

de antemão, que fatores sociais e linguísticos estão (sim) interferindo para que a variação do artigo definido em contexto de prenomes aconteça na cidade do Recife.

Após duas rodadas no *GoldVarbX*, o programa elegeu cinco variáveis como sendo as mais significativas para o comportamento do fenômeno variável investigado nessa pesquisa, a saber: *Sexo*, *Faixa etária*, *Tipo de informação*, *Referência do antropônimo* e *Tipologia do antropônimo*, respectivamente. Assim sendo, apresentaremos, a seguir, os resultados preliminares referentes a essas variáveis e a seus respectivos fatores.

## 5.2 VARIÁVEIS APONTADAS COMO SIGNIFICATIVAS PELO PROGRAMA *GOLDVARBX*

Serão expostos aqui os resultados e as discussões das variáveis abalizadas como relevantes pelo programa computacional *GoldVarbX*. É pertinente evidenciarmos, também, que as variáveis selecionadas são de ordem Linguística e Extralinguística e que seus resultados serão apresentados, ao curso desta análise, tanto em números percentuais, quanto através do peso relativo (P.R.) fornecido pelo programa.

É adequado realçarmos que os contextos analíticos utilizados na rodada oficial dos dados foram processados em função da variante *presença do artigo*. Com isso, queremos dizer que o programa apresentou os fatores mais significativos para a presença do artigo definido antepondo antropônimos, e não para a omissão dele. Na ocasião das rodadas, apontamos o código “c” (com artigo) como fator de aplicação; já que, como apontam Santos e Vítório (2008), o programa solicita um valor de aplicação: “a variante dentro da variável dependente usada para gerar o valor de aplicação dos resultados” (SANTOS & VITÓRIO, p. 11, 2018).

À vista disso, os P.R. disponibilizados pelo programa computacional apresentaram os fatores mais significativos para a marcação do artigo definido, conforme já mencionamos. Isso será relevante à medida que, uma vez confirmado que o perfil da comunidade de fala tende para a não marcação da partícula determinante, isolaremos as variáveis responsáveis por condicionar, justamente, o uso da variante inovadora, isto é, a variante com artigo (cf. 36 a 39). De tal modo, apresentaremos, a seguir, o resultado da primeira variável apontada como significativa para o uso do determinante: a variável *Sexo dos informantes*.

### 5.2.1 Sexo

Tomando por base os resultados obtidos pelo *GoldVarbX*, observamos que o sexo dos informantes foi uma das variáveis de caráter social apontadas como significativa nas nossas investigações. Conforme já mencionamos, para essa variável, levamos em consideração apenas o sexo biológico dos informantes. É importante ressaltarmos que, mesmo considerando a importância de se ponderar sobre questões de gênero e suas possíveis implicações linguísticas, optamos por controlar, nesse momento, apenas o sexo dos nossos informantes. Nessas circunstâncias, os resultados angariados a partir das rodadas do programa *GoldVarbX* serão apresentados no quadro (9) a seguir:

**Quadro 9** - Resultado da variável Sexo do informante

<b>Sexo</b>	<b>Presença</b>	<b>Ausência</b>	<b>P.R.</b>
Feminino	25.4% (93/366)	74.6% (273/366)	0.412
Masculino	47.3% (174/368)	52.7% (194/368)	0.587

**Fonte:** Autoria própria

Conforme exposto acima, existe uma distinção entre os fatores (masculino e feminino) no que diz respeito à preferência pela marcação do artigo antepondo antropônimos. De acordo com os nossos resultados, é possível conferir que os homens de Recife parecem ser mais sensíveis à presença da partícula determinante. Confirmando nossa hipótese para essa variável, percebemos que, assim como divulgado por Campos Júnior (2011) sobre a capital capixaba, as mulheres recifenses também tendem a realizar menos o artigo definido diante de antropônimos do que os homens na comunidade de fala.

Ao nos debruçarmos sobre os resultados percentuais, verificamos que o sexo masculino apresentou uma porcentagem expressiva de 47.3% (174/368) de realização, contra 52.7% (194/368) de ausência do artigo diante de nomes próprios; fator que parece ratificar que os homens têm um percentual de marcação superior ao das mulheres na comunidade. Ou seja, os homens utilizam a partícula determinante diante de antropônimos de maneira bastante acentuada, haja vista a quase equivalência entre os usos. Já o sexo feminino apresentou os dados

percentuais de 25.4% (93/366) de marcação do artigo definido, contra 74.6% (273/366) de não marcação da partícula definida em contextos de antropônimos.

A distinção entre os sexos fica evidente, também, quando voltamos nosso olhar para os resultados obtidos por meio do P.R. É possível constatar que, após o cruzamento de todos os fatores, o fator *masculino* apresentou um P.R. de 0.587, ao passo que o *feminino* foi menos acentuado: 0.412.

Todavia, é prudente destacarmos um fato interessante que merece ser discutido no âmbito deste trabalho: parte considerável dos informantes homens que compunham as células sociais observadas nesta investigação sinalizou (em ficha social) como sendo pertencente à comunidade LGBTQ+ da cidade. *Grosso modo*, podemos dizer que essa especificidade parece influenciar significativamente para os resultados obtidos, uma vez que o número mais expressivo de aplicação atribuído ao sexo masculino está atrelado a outro fenômeno sintático – já mencionado nos estudos de Oliveira (2018) – que merece ênfase aqui: a não concordância de gênero e/ou concordância inversa.

Em outras palavras, queremos dizer que é, justamente, nesse grupo de informantes que encontramos sentenças com configurações atípicas no que tange à concordância nominal de gênero. Os exemplos, a seguir, ilustram melhor tais ocorrências:

(44) “**A Vinícius** tem medo das héteros!” (inf. VMO. M. 2).

(45) “A poc, **a Epitácio** também me apoiou bastante.” (inf. JPS. M. 2).

(46) “Apesar de todos os surtos ainda estou com **a Ryan**.” (inf. JPS. M. 2).

(47) “O que eu acho mais interessante é que **a Luan** só me assumia na Metrópole.” (inf. JPS. 2).

De acordo com Oliveira (2018), ocorrências como as expressas acima foram recorrentes também nos dados de Tabira, no interior do estado de Pernambuco. Na concepção da estudiosa, aqueles informantes pertencentes ao grupo social LGBTQ+ também utilizaram o que agora optamos por chamar – em consonância com Oliveira (2018) – de **concordância inversa de gênero**. Nesse prisma, é possível asseverarmos que os falantes recifenses, também pertencentes a esse público, parecem repetir o mesmo padrão dos falantes sertanejos: empregar a concordância inversa de gênero com nomes próprios masculinos no âmbito dos sintagmas (cf. 44 a 47).

Nos dados de Oliveira (2018), esse tipo de não concordância de gênero se deu apenas entre falantes do sexo masculino, ou seja, as mulheres da comunidade de fala – mesmo aquelas

identificadas como pertencentes à comunidade LGBTQ+ – não apresentaram tal padrão de aplicação do artigo definido em sintagmas antroponímicos. Assim como nos dados de Oliveira (2018), a pesquisa em tela mostra que o mesmo arquétipo parece se repetir na capital pernambucana. Isto é, todos os homens identificados enquanto LGBTQ+ entrevistados nesta pesquisa (3 informantes dos 6 homens colaboradores; sendo 2 da primeira faixa etária e 1 da segunda faixa etária) apresentaram ocorrências de concordância inversa.

Um fato que merece destaque é que, mesmo (re)produzindo sentenças como as expressas nos exemplos de (44) a (47), esses mesmos informantes também apresentaram ocorrências de concordância de gênero tidas como padrão pela gramática normativa (Ex.: *O João não gosta de mim.*<sup>76</sup>). Nessa linha argumentativa, optamos por considerar os dois casos como sendo sim variação do artigo definido em sintagmas antroponímicos, mesmo aqueles que apresentavam concordância inversa (cf. (44) a (47), por exemplo). Assim sendo, os casos de concordância inversa foram calculados enquanto ocorrências de uso marcado do artigo definido; mais adiante veremos o porquê dessa escolha.

Nossa primeira justificativa para quantificarmos esse tipo de ocorrência como uso variável reside, justamente, no fato de observarmos que os mesmos falantes fazem uso das três formas já catalogadas aqui: (i) sem o artigo definido; (ii) com artigo definido e (iii) com artigo definido, porém com concordância inversa; provando que as três formas existem e competem entre si, enquanto variantes. Ainda assim, como sabemos que esta última ocorrência diz respeito a uma parcela muito específica de uso<sup>77</sup>, entendemos que ela, *grosso modo*, pertence à segunda, ou seja, é apenas um sintagma antroponímico com a posição de determinante marcada<sup>78</sup>; ainda que esse determinante – nos moldes da gramática tradicional – não concorde com o nome próprio que o acompanha na sentença.

A segunda justificativa põe em relevo questões (inteiramente) sintáticas e é, seguramente, a que julgamos mais relevante para este trabalho que elege sintagmas antroponímicos como objeto de análise. Mesmo após considerarmos que o uso do artigo feminino antes de antropônimos masculinos põe em relevo questões semânticas, comunicativas e sociais, entendemos que há um importante critério sintático que devemos obedecer: o critério

---

<sup>76</sup> Exemplo real retirado do nosso *corpus* de análise (inf. VMO. M. 2)

<sup>77</sup> Mesmo observando essas três ocorrências, acreditamos que não haveria motivos para criarmos uma nova variante para a variável dependente desta pesquisa (cf. (iii)). Conforme observado, sabemos que as ocorrências de concordância inversa só parecem ocorrer sob a influência do sexo e da orientação sexual dos informantes, mas seguem sendo uma forma com a posição do determinante (artigo) marcada (cf.(ii)).

<sup>78</sup> [determinante + antropônimo]

da posição marcada do determinante no sintagma. Como é possível verificar nos exemplos (44), (45), (46) e (47), a posição do determinante está preenchida nas sentenças e isso garante que, mesmo que de maneira inversa e distante do que prescreve a gramática tradicional, ainda temos um artigo definido antecedendo um nome próprio de pessoa no interior do sintagma antroponímico. Tal fator, para o curso desta pesquisa, ratifica a necessidade de quantificação de dados como esses.

Todavia, acreditamos que só estudos futuros – que tenham como objetivo o aprofundamento apenas nesse tipo de ocorrências – serão capazes de delinear com robustez a variação do artigo definido em contextos antroponímicos em comunidades LGBTQ+ do Recife. Por ora, optamos por não nos aprofundarmos nessas questões, haja vista que isso fugiria do nosso objetivo geral de pesquisa.

Ainda assim, achamos válido debatermos sobre algumas ponderações significativas. Tomando por base as sentenças ilustradas nos exemplos acima (cf. de (44) a (47)), observamos que os informantes aplicam propositalmente o artigo definido feminino [a] antes de antropônimos que são habitualmente admitidos como masculinos, como é possível verificar em: “a Vinícius”, “a Epitácio” “a Ryan” e “a Luan”. Dizendo de outro modo: parece haver uma consciência linguística e, sobretudo, social regulando tais escolhas.

**Consciência linguística** porque os informantes utilizam seu conhecimento da língua para induzir que construções como [a João] – uma inversão apenas na concordância de gênero – parecem ser permitidas no PB; mas não produzem, por exemplo, construções com inversões estruturais, do tipo [\*João a], porque têm noção, ainda que inconscientemente, de que estariam violando a gramaticalidade da nossa língua. A **consciência social**, por sua vez, reside no fato de os informantes terem discernimento na hora de eleger o determinante feminino na saturação da sentença. Essa suposição se mostra mais receptível à medida que percebemos que esse comportamento é mais recorrente quando os dois indivíduos envolvidos no discurso são integrantes da comunidade LGBTQ+: tanto o informante entrevistado quanto o ser referenciado no seu discurso.

Nesse enquadramento analítico, parece prudente conjecturar – indo mais além nas ponderações propostas por Oliveira (2018) – que sentenças como as expressas em (44), (45), (46) e (47) são menos usuais (e/ou inexistentes) entre falantes que tenham perfis sociais dessemelhantes ao do público investigado, isto é, pessoas que não sejam membros da comunidade LGBTQ+. Por conseguinte, entendemos que há uma potencial influência de questões sociais e identitárias que parecem interferir diretamente na organização canônica da concordância nominal no interior dos sintagmas antroponímicos. Não obstante, podemos inferir

que também há, por parte desses informantes, uma forte percepção linguística que os permite inovar na língua – fazendo uma inversão de gênero (cf. 44 a 47) – mas não a ponto de torná-la agramatical.

Vale ressaltar que as discussões sobre essa temática serão mais aprofundadas à proporção que tivermos acesso a dados mais complexos, possivelmente, em trabalhos futuros acerca da temática em questão.

Na seção seguinte, apresentaremos os resultados da próxima variável extralinguística apontada como significativa pelo programa computacional: *Faixa etária*.

### 5.2.2 Faixa etária

Nesta subseção, iremos avaliar de que maneira a variável *faixa etária* dos informantes influenciou no comportamento estatístico dos nossos dados, haja vista que ela foi considerada pelo programa computacional GoldVarbX como sendo a segunda variável extralinguística significativa para o emprego do artigo definido diante de antropônimos na cidade do Recife. É prudente salientarmos que, conforme já o fizemos no capítulo metodológico desta pesquisa, consideramos apenas duas faixas etárias durante nossa coleta de dados (Primeira faixa etária: 18 a 35 anos; segunda faixa etária: acima de 35 anos). Conforme justificamos, a primeira faixa etária levaria em consideração a fala das crianças da comunidade, fator que, provavelmente, forneceria uma amostragem mais representativa de dados linguísticos da comunidade de fala<sup>79</sup>. Todavia, para esta análise, consideramos apenas a fala de jovens e adultos recifenses e essas duas faixas etárias escolhidas se mostraram relevantes, assim como veremos a seguir.

De modo geral, podemos inferir que a primeira faixa etária condicionou de maneira bastante significativa a realização do artigo definido diante de antropônimos e isso se mostra perceptível à medida que observamos os resultados percentuais expressos no quadro 10, a seguir:

---

<sup>79</sup> Conforme ressaltamos no capítulo metodológico desta pesquisa, a nossa coleta de dados linguísticos foi severamente prejudicada pela pandemia mundial do novo Corona Vírus. Tal infortúnio nos obrigou a diminuir drasticamente o número de informantes e faixas etárias que iriam compor o corpus geral desta pesquisa. De tal modo, consideramos apenas a fala de 12 jovens e adultos divididos em duas faixas etárias distintas, deixando de fora – por questões de segurança dos envolvidos na pesquisa – crianças e pessoas acima de 65 anos e/ou com comorbidades.

**Quadro 10** - Resultado da variável Sexo do informante

<b>Faixa etária</b>	<b>Presença</b>	<b>Ausência</b>	<b>P. R</b>
Primeira Faixa etária (18 a 35 anos)	47.7% (226/474)	52.3% (248/474)	0.635
Segunda Faixa etária (acima de 35 anos)	15.8% (41/260)	84.2% (219/260)	0.267

**Fonte:** Autoria própria

Ainda se considerarmos que o número de ausência de artigos definidos é maior do que o percentual de presença em ambas as faixas etárias; não podemos, contudo, ignorar o expressivo número de ocorrências atribuídas a primeira faixa etária de informantes (18 a 35 anos). Conforme exposto acima, a *primeira faixa etária* apresentou um percentual de 47.7% (226/474) de marcação do determinante, contra 52.3% (248/474) de não aplicação da partícula. Se observarmos o número de ocorrências por sentenças, e não percentualmente, podemos inferir que há, de fato, uma quase equivalência entre as duas variantes da variável dependente (com artigo e sem artigo, respectivamente). Em outras palavras, 226 ocorrências é um número quase análogo a 248, sobretudo se levarmos em consideração o número expressivo de sentenças contabilizadas (734) na pesquisa.

Ao colocarmos em perspectiva os resultados percentuais da *primeira faixa etária* com os da *segunda faixa etária*, podemos perceber que, de fato, a *primeira* tem uma expressiva preferência pela variante com artigo. Isso porque apenas 15.8% (41/260) dos informantes da *segunda faixa etária* utilizaram antropônimos licenciados com a presença de artigos definidos, ao passo que o número de ausência do determinante para essa faixa etária é ligeiramente categórico: 84.2% (219/260).

O P.R., que leva em consideração o cruzamento de todos os fatores, aponta, mais uma vez, para importância da variável *faixa etária* na variação do artigo definido em contextos antroponímicos na comunidade de fala estudada: 0.635 para a *primeira faixa etária* e apenas 0.267 para a *segunda*.

O P.R se mostra fundamental para esse tipo de análise uma vez que é capaz de apontar, justamente, para os desencadeadores da variação. Isso quer dizer que – ao sinalizarmos a variante com artigo (c) como a inovadora e como o fator de aplicação no programa computacional – os resultados são gerados em função da presença do artigo, e não da ausência



deste. Desse modo, o P.R mais elevado aponta sempre para os fatores que fomentam a variação na comunidade de fala e a *primeira faixa etária* dos informantes foi um desses fatores significativos.

Tomando por base os estudos de Oliveira (2018), decidimos controlar a variável faixa etária sob a hipótese de que – conforme aponta a estudiosa em dados do Sertão do Pajeú pernambucano – falantes da faixa etária mais avançada usariam menos o artigo do que falantes mais jovens. Com isso, percebemos que tal hipótese se concretizou em nossos dados.

Ora, se levarmos em consideração que os falantes mais velhos são mais resistentes às variantes inovadoras, somos levados a perceber que o mesmo movimento que aconteceu no interior do estado (OLIVEIRA, 2018) parece se repetir em Recife: os falantes mais velhos são mais conservadores e utilizam bem menos o artigo precedendo nomes próprios de pessoas do que os falantes mais jovens. Em contrapartida, os colaboradores da primeira faixa etária se mostram bem mais sensíveis à variante inovadora do fenômeno.

Outro fator que merece ser discutido nos limites desta subseção é a importância do controle da variável *faixa etária* para localizarmos e/ou projetarmos possíveis processos de mudança linguística em curso. Ainda mais, quando nos referimos a fenômenos linguísticos específicos que se atrelam a questões dialetais da fala, como é o caso da variação do artigo definido em sintagmas antroponímicos.

Conforme ressaltamos no trabalho de Callou e Silva (1997), a variação do artigo definido diante de antropônimos sofre forte influência do fator região de origem dos informantes. Isso nos diz que, a depender da região do falante, o uso da partícula determinante será mais ou menos acentuado. Ainda de acordo com as autoras, duas capitais do nordeste (Recife e Salvador) têm uma tendência mais acentuada para anular o uso de artigo definido, especialmente quando comparadas com outras capitais do sudeste. Nesses moldes, podemos observar que a variante padrão na capital pernambucana, assim como no estado de Pernambuco (PEREIRA, 2017; OLIVEIRA, 2018), parece – de fato – ser a forma sem artigo.

Todavia, ao nos debruçarmos mais profundamente sobre essas constatações, podemos inferir que – conforme mostram nossos resultados – temos uma proeminência de uso do artigo definido, justamente, entre os falantes mais jovens, o que nos faz suspeitar, ainda que ligeiramente, de um possível processo de mudança linguística. No entanto, como bem sabemos, tal processo é bastante complexo e precisaríamos de muitas outras evidências para afirmar que se trata, realmente, de mudança linguística em curso. Além disso, necessitaríamos, também, de uma vasta análise diacrônica acerca da antroponímia em Pernambuco para avaliarmos a concretude de tais possibilidades. Contudo, acreditamos que esses apontamentos são bastante

relevantes, sobretudo no que diz respeito ao desenvolvimento necessário de pesquisas futuras sobre o falar da capital pernambucana.

Assim sendo, na próxima subseção, iremos expor e discutir os resultados percentuais disponibilizados pelo programa *GoldVarbX* acerca da primeira variável linguística tida como significativa: *Tipo de informação*.

### 5.2.3 Tipo de informação

A variável *tipo de informação*, conforme exposto no capítulo 4, diz respeito ao nível de conhecimento que os componentes da entrevista (inquiridora e entrevistado) têm dos referentes citados no discurso; assim, seus respectivos fatores são: *informação compartilhada (famoso)* e *informação não compartilhada (não famoso)*.

É válido observarmos que há, no âmbito dessa variável, uma bifurcação que poderia implicar na criação de uma nova variável, haja vista a necessidade de especificar, também, se o indivíduo referenciado é – ou não – uma figura de domínio público. Isto é, a título de exemplo, poderíamos encontrar um referente que fosse, simultaneamente, não famoso e conhecido pelos membros da entrevista; da mesma maneira, o contrário também poderia acontecer: um referente famoso (de domínio público), mas não conhecido por um dos membros da entrevista (logo, uma informação não compartilhada). Nessa ótica argumentativa, precisaríamos investir na criação de uma nova variável para controlar, exclusivamente, essa particularidade. Todavia, é prudente ressaltarmos que, em nossos dados, todas as ocorrências assinaladas como *informação compartilhada*, os referentes eram, justamente, famosos.

Isso posto, não houve, em nosso *corpus* analítico, nenhuma ocorrência na qual o referente (antropônimo) era, ao mesmo tempo, de conhecimento dos envolvidos e não famoso, e/ou vice-versa. Assim sendo, não existiu a necessidade de instituir uma nova variável que considerasse apenas os episódios nos quais os referentes eram figuras de domínio público, isso porque o nosso *corpus* de pesquisa consentiu a integração, isto é, a equivalência das variáveis.

De antemão, é apropriado observarmos que, na grande parte das ocorrências licenciadas com a presença do artigo definido, os nomes próprios de pessoas – alocados no interior dos sintagmas antroponímicos – se referiam a indivíduos de domínio público, ou seja, os referentes eram famosos (mesmo que de natureza ficcional), fator que já pressupõe o cruzamento necessário entre as duas variáveis: *Tipo de informação e Referência dos antropônimos*. Os

exemplos (48), (49), (50) e (51), extraídos do nosso *corpus* analítico, exemplificam melhor tal assertiva:

(48) “Eu gosto muito **da Madonna**” (inf. VMO. M. 2).

(49) “Mas deixa **o Harry Potter**, né?” (inf. MRMT. F. 2)

(50) “Há quem curte **o Graciliano Ramos**” (inf. VMO. M. 2).

(51) “Eu *tava* ouvindo **o Belchior**” (inf. JPS. M. 2)<sup>80</sup>.

Com o auxílio dos exemplos ilustrados acima, podemos observar que, das variáveis investigadas nessa pesquisa, essa foi a terceira apontada como sendo bastante significativa pelo programa GoldVarbX. De modo abrangente, podemos sancionar que o fator *informação não compartilhada* – inerente à variável tipo de informação – aparenta anular significativamente a realização do artigo definido frente a antropônimos, conforme é possível verificar no quadro a seguir:

**Quadro 11** - Resultado da variável Tipo de informação

<i><b>Tipo de Informação</b></i>	<b>Presença</b>	<b>Ausência</b>	<b>P.R.</b>
Compartilhada/ famoso	60.5% (179/296)	39.5% (117/296)	0.743
Não compartilhada/ não famoso	20.1% (88/438)	79.9% (350/438)	0.328

**Fonte:** Autoria própria

Como é possível examinar no quadro acima, os resultados obtidos sugerem que o fator *informação compartilhada* condiciona significativamente a realização do artigo definido em sintagmas antroponímicos: 60.5% (179/296). Isso se mostra mais evidente quando comparamos com o percentual mais baixo de ausência do artigo para tal fator: 39.5% (117/296). Como já especificamos em seções anteriores, o padrão da comunidade da fala tem uma tendenciosa

---

<sup>80</sup> É importante destacarmos que, nesse contexto específico, pode haver semelhança sonora entre “o” final do verbo “ouvindo” e o “o” artigo. É válido ratificarmos, no entanto, que essa sentença só foi quantificada porque houve uma longa pausa – por parte do informante – antes da pronúncia do sintagma antroponímico “o Belchior”. De tal modo, foi possível confirmar que realmente houve o uso do determinante nesse dado analítico.

preferência pela não marcação do artigo definido diante de antropônimos, e isso é observável quando averiguamos os resultados estatísticos do fator *informação não compartilhada*: 20.1% (88/438) de realização do determinante contra 79.9% (350/438) de não realização.

Todavia, esses resultados parecem apontar para generalizações interessantes acerca da comunidade de fala investigada. Se levarmos em consideração o percentual de realização do artigo definido diante de antropônimos atribuído a referentes famosos – 60.5% (179/296) – podemos admitir que, *grosso modo*, essa variável é bastante importante para desencadear a variação na capital pernambucana. Ou seja, no que se refere à realização marcada de determinantes em sintagmas antroponímicos na cidade do Recife, podemos delegar intenso condicionamento ao fator *informação compartilhada*. Nesses termos, o número pouco expressivo de realização frente a antropônimos de referentes não famosos (39,5% (117/296)) só ratifica essa generalização.

A proeminência da variável *Tipo de informação* pode ser ressaltada, também, mediante avaliação do peso relativo determinado pelos resultados fornecidos pelo *Software*. Conforme verificamos no quadro acima, o fator *Informação compartilhada* obteve um peso relativo de 0.743, ao passo que o fator *Informação não compartilhada* apresentou um P.R. de 0.328.

Em conformidade com a seção metodológica da pesquisa em tela, observamos que o P.R. é concedido, justamente, por meio do cruzamento de todos os fatores disponibilizados pelo programa. Logo, os números apontados por ele apresentam, de maneira mais circunstanciada, a importância da variável e o comportamento de seus respectivos fatores. Sendo assim, os resultados abalizados pelo P.R. também sancionam a influência do fator *informação compartilhada* para a atuação do artigo antepondo nomes próprios de pessoas. Nesse alinhamento argumentativo, podemos depreender que, com base no P.R. obtido, o fator *informação compartilhada* atua de maneira significativa para a realização do artigo definido.

É adequado empregarmos algumas considerações pertinentes ao percentual elevado de ausência do artigo definido atribuído ao fator *informação não compartilhada* e o percentual, igualmente elevado, de presença conferido ao *fator informação compartilhada*. Como a variável em questão faz menção ao nível de conhecimento que os membros da entrevista têm dos referentes (dos antropônimos e/ou das pessoas que eles designam), o *fator informação não compartilhada* faz alusão, justamente, aos antropônimos que – ao serem enunciados no discurso – são conhecidos apenas dos informantes. Na grande maioria desses casos, são pessoas que

fazem parte dos círculos afetivos e de convivência dos informantes e, por conseguinte, não-famosos<sup>81</sup>.

Ao conceber um percentual expressivo de ausência do artigo – 79.9 % (350/438) – nos casos de *informação não compartilhada*, podemos induzir que a comunidade de fala parece anular o artigo definido, especialmente, quando os informantes mencionam pessoas próximas a eles. Isso se mostrou mais notório ao revisitarmos, atentamente, os nossos dados e ratificarmos que os sintagmas antroponímicos habilitados com artigo referem-se, de modo geral, a antropônimos famosos (60.5% (179/296)), o que sanciona a nossa hipótese para esta variável.

Intuímos – com os resultados apresentados e discutidos nesta dissertação – que na comunidade de fala estudada, a seguinte constatação se mostra bastante admissível: em Recife, os informantes tendem a definir, por meio do uso marcado do artigo definido, exatamente aqueles referentes que já são de domínio público, ou seja, famosos (cf. 48, 49, 50 e 51).

Sem embargos, só trabalhos ainda mais aprofundados podem confirmar, ou não, os resultados das generalizações proposta até aqui. Todavia, nessa esteira de pensamento, lançamos mão de algumas hipóteses de intuição linguística<sup>82</sup> na tentativa de explicar (se existem e) quais são os parâmetros que regulamentam o uso acentuado de artigos definidos diante de nomes próprios de famosos:

- (i) Como são referentes famosos que têm forte ligação com as mídias em geral, podemos sugerir que a influência para a marcação do determinante reside, justamente, no fato de termos contato constante – por meio da televisão e/ou mídias sociais – com dialetos que priorizam o uso marcado do artigo.
- (ii) Outra hipótese bastante aceita por nós para explicar o uso acentuado do artigo definido diante de nomes de famosos se baseia, justamente, no alto índice de *ausência* dos determinantes antepondo referentes *não famosos*, isto é, próximos dos informantes. Nessa linha de raciocínio, podemos supor que o artigo definido é utilizado para marcar

---

<sup>81</sup> Conforme destacamos acima, em todos os casos em que a informação era compartilhada entre os membros da entrevista, os referentes eram pessoas de domínio público, isto é, referentes famosos.

<sup>82</sup> É válido destacarmos que aqui damos crédito à intuição linguística da inquiridora responsável pela pesquisa em tela e não para a intuição linguística de cada informante colaborador em particular, dado que eles não foram submetidos a questões desta natureza. Com isso, destacamos, também, a importância de trabalhos linguísticos que foquem na perspectiva de intuição linguística dos informantes, por meio de questionários, afirmações e suposições bem mais específicas sobre as variáveis. A importância complementar de trabalhos desta natureza fundamentaria de maneira bastante robusta as lacunas que, porventura, serão deixadas por esta pesquisa.

que aquele referente – embora conhecido por todos, por ser famoso – carrega a particularidade de ser distante do informante, ou seja, não denota familiaridade, caso contrário, não viria acompanhado pelo artigo. Forte indício para essa hipótese é o percentual altíssimo de ausência para não famosos: 79.9 % (350/438).

- (iii) Por fim, acreditamos que o uso do artigo diante de antropônimos famosos pode ser, também, um artifício utilizado pelos falantes pernambucanos para particularizar aquela figura pública em detrimento de alguém com mesmo nome, mas que não é famoso. Isto é, seria “O Alceu<sup>83</sup>”, cantor e compositor conhecido e bastante famoso no estado, e não outra pessoa;

É válido inferirmos, no entanto, que as afirmações expressas acima ((i); (ii); (iii)) foram elaboradas, sobretudo, na tentativa de explicar tais usos e, por isso, não são afirmações conclusivas sobre a nossa pesquisa. Ao sabermos que tal variável influencia o uso marcado dos determinantes em sintagmas antroponímicos, tais hipóteses não passam de conjecturas desenvolvidas com o intuito de melhor explicar os percentuais de uso e a importância de tal variável para o nosso objeto de pesquisa.

Dito isso, na subseção seguinte iremos apresentar os resultados da variável significativa *Referência dos antropônimos*.

#### 5.2.4 Referência dos antropônimos

Conforme já especificado no capítulo metodológico deste trabalho, essa variável diz respeito à natureza do referente; isto é, se o ser que está sendo referenciado no enunciado é uma pessoa *real* ou um indivíduo *ficcional* (nesse grupo, podemos incluir, de modo mais amplo, personagens de filmes, séries, novelas e até desenhos animados). Conforme observamos anteriormente, os estudos de Pereira (2017) abalizam que o fator *personagem* favorece mais a presença do artigo definido do que a sua ausência na cidade de Caruaru- PE. Oliveira (2018), ao controlar tal variável em dados do Sertão Pernambucano, também observa que o fator *personagem* condiciona exponencialmente o uso de artigos definidos antecedendo

---

<sup>83</sup> Dado retirado do nosso *corpus* de pesquisa.

antropônimos, sendo eleita, pela autora, como uma das principais agentes da variação na cidade de Tabira- PE.

Nessa esteira de pensamento, os nossos resultados percentuais apontam que a capital pernambucana parece permanecer em uniformidade com as demais cidades observadas no Sertão do Pajeú<sup>84</sup>, confirmando, assim, nossa hipótese para essa variável: o fator *personagem* condiciona expressivamente a realização de artigos definidos em sintagmas antroponímicos na cidade do Recife. A seguir, apresentamos alguns exemplos que ilustraram melhor as ocorrências possíveis para os dois fatores dessa variável:

➤ Real

(52) “**Pedro** é muito parecido com meu pai.” (inf. MRMT. F. 2)

(53) “**Luma** é incrível, ela sempre anima o rolê.” (inf. LSS. F. 2)

(54) “**Rejane**, minha professora de física, adorava gatos e eu adorava ela!” (inf. LSS. F. 2)

➤ Personagem

(55) “**O Tyrion** é muito inteligente.” (inf. MRMT. F. 2)

(56) “**A Robin** é incrível em *How met your mother*, ela assume as responsabilidades dela.” (inf. MRMT. F. 2)

(57) “Eu gosto muito da personagem principal, **a Anne**.” (inf. LSS. F. 2)

Conforme é possível verificar nos exemplos expostos acima, podemos inferir que o fator *personagem* condiciona expressivamente a ocorrência do artigo definido diante de antropônimos. Vejamos o quadro, a seguir, que ratifica ainda mais essa explanação:

---

<sup>84</sup> Verificar os resultados dos estudos de Pereira (2017) e Oliveira (2018) discutidos no segundo capítulo desta Dissertação.

**Quadro 12** - Resultado da variável Referência do antropônimo

Referente	Presença	Ausência	P. R
<i>Real</i>	28.9% (187/648)	71.1% (461/648)	0.420
<i>Personagem</i>	93% (80/86)	7% (6/86)	0.920

**Fonte:** Autoria própria

Segundo o quadro acima, os nossos resultados apontam que, quando o antropônimo faz referência a um *personagem*, o percentual expressivo de presença do artigo definido é de 93% (80/86); um número bastante elevado, especialmente quando comparado ao percentual baixíssimo de ausência: 7% (6/86). Essa relevância se mostra ainda mais acentuada quando observamos o número de ocorrências e não apenas os resultados percentuais. Dos 86 sintagmas antropônimos em que o referente era um indivíduo de natureza ficcional (personagens de novelas, desenhos animados, filmes ou séries), 80 deles foram habilitados com a presença do artigo definido.

Já quando os antropônimos designavam indivíduos reais, o percentual de presença do artigo definido caiu significativamente para 28.9% (187/648) e o percentual de ausência se mostrou bem mais elevado com 71.1% (461/648) de não marcação do determinante. Entretanto, é válido observarmos que referentes reais são bem mais comuns em nosso *corpus* de pesquisa e isso fica evidente quando atentamos para o número acentuado de ocorrências desse tipo: 648, ao todo. Mesmo assim, percebemos a relevância do fator *personagem* para a realização do artigo, dado o número elevadíssimo de presença da partícula: (80/86).

A importância da variável *referência do antropônimo* também pode ser avaliada mediante consulta ao do P.R. fornecido pelo *GoldVarbX*. Levando em consideração nossos resultados, verificamos que, quando o antropônimo se refere a personagens o P.R. atingiu alto índice de relevância: 0.920. Entretanto, quando o indivíduo referenciado pelo antropônimo era de um ser de natureza real o P.R. foi bem mais inferior: 0.420.

De tal modo, é adequado concluirmos que o fator *personagem* tem bastante influência para a realização do artigo definido nos dados observados. Como é possível verificar, mesmo tendo um percentual de não realização do artigo maior do que o de realização, os falantes da capital pernambucana tendem a assinalar a presença do artigo nas sentenças, justamente, quando os antropônimos designam personagens fictícios (cf. (55), (56) e (57)); isto é, seres com



os quais não têm relação de convivência. Já, quando os antropônimos são seres reais (cf. (52), (53) e (54)), a comunidade de fala parece invalidar a marcação do artigo.

Dando ênfase aos resultados percentuais, é válido destacarmos que, em termos de marcação do artigo definido em sintagmas antroponímicos, a variável Referência do antropônimo foi a que apresentou dados mais expressivos. Nessa esteira de pensamento, podemos verificar certa ligação entre as variáveis *Tipo de Informação a* e *Natureza do referente*, haja vista que os fatores<sup>85</sup> significativos de ambas se relacionam a antropônimos de domínio público.

Por assim dizer, discorreremos, a seguir, acerca da variável significativa *Tipologia dos antropônimos*.

### 5.2.5 Tipologia dos antropônimos

De todos os empreendimentos analíticos testados neste trabalho, a variável *tipologia do antropônimo* foi pensada, exclusivamente, para os dados da capital pernambucana e pode ser considerada, até certo ponto, como pioneira nos trabalhos sociolinguísticos que envolvem antropônimos até o momento. Isso se deu porque, uma vez que nos debruçarmos sobre as colocações tipológicas de Amaral e Seide (2020), percebemos que pouco se falava – nas pesquisas sociolinguísticas revisadas até aqui – sobre a interferência dos tipos de antropônimos para o fenômeno da variação do artigo definido. O que fizemos nesta dissertação, portanto, foi controlar esse segmento teórico da tipologia<sup>86</sup> dos nomes próprios e associá-lo à Sociolinguística Variacionista, transformando-o, assim, em uma variável linguística independente.

Dessa maneira, os resultados dessa variável parecem apontar para caminhos inovadores enveredados pela pesquisa em tela, haja vista que a tipologia dos antropônimos não apenas se mostrou relevante para a variação, conforme apontam os resultados do programa *GoldVarbX*, como suscitará discussões interessantes quando associada às outras variáveis selecionadas para este empreendimento investigativo. Com isso, queremos dizer que esse tipo de movimento analítico se mostrou relevante para esse trabalho à medida que relacionou, notadamente, a metodologia de base Variacionista a aspectos de ordem linguística que se referem à antroponímia.

---

<sup>85</sup> *Personagem e Informação compartilhada* (famoso).

<sup>86</sup> Tipologia proposta por Amaral e Seide (2020) e discutida no segundo capítulo dessa dissertação.

Conforme foi possível observar em nosso capítulo metodológico, essa variável é a que possui o maior número de fatores: 13 (treze) tipos de antropônimos, segundo a tipologia proposta por Amaral e Seide (2020)<sup>87</sup>. No entanto, após a análise e a codificação dos nossos dados, observamos que apenas 5 (cinco) tipos de antropônimos se mostraram recorrentes em nosso *corpus* de pesquisa: Prenome, Apelido, Nome de urna, Nome de palco e Sobrenome.

É imprescindível destacarmos que todos os 13 tipos de antropônimos foram verificados em nossos dados. Contudo, o número bastante reduzido de algumas ocorrências fez com que excluíssemos esses dados da nossa quantificação. Isto é, quando o programa localiza um número pouco expressivo de ocorrências entre variáveis com muitos fatores – como é o caso da função sintática e da tipologia dos antropônimos – o algoritmo supõe que se trata de um erro de digitação dos códigos e não prossegue com as rodagens de dados. Por essa razão, para evitarmos esse problema logístico, optamos por contabilizar apenas os cinco fatores mais recorrentes dessa variável. Assim sendo, os exemplos a seguir ilustram, respectivamente, algumas dessas ocorrências.

➤ Prenome

(52) Tem um tio meu, **Nilton**, que é caminhoneiro.

(53) A **Luna** é um pouco lerdinha, mas eu gosto dela também.

➤ Apelido

(54) **Nono** é meu tio mais distante.

(55) Tem o meu tio, o **Bim**.

➤ Nome de urna

(56) A **Manuela D'Ávila** é uma mulher na política que eu gosto de acompanhar.

(57) Eu não votaria em **Bolsonaro**.

➤ Nome de Palco

(58) **A Pablo Vittar** acho que ela é um exemplo de superação e talento.

(59) Eu gosto muito **da Glória Groove** que é do vale.

---

<sup>87</sup> A saber: Prenome; Sobrenome; Agnome; Apelido (ou alcunha, ou cognome); Hipocorístico; Pseudônimo; Codinome; Heterônimo; Nome artístico (e nome de palco); Nome religioso; Nome social; Nome de urna e Nome parlamentar.

➤ Sobrenome

(60) **O Mendes** era uma múmia naquele apartamento.

(61) **Nascimento** é aquele tipo de homem *Sugar Daddy*.

Dando seguimento aos nossos resultados, o quadro abaixo apresenta os números percentuais de presença e de ausência para cada fator e, também, o P.R. fornecido pelo programa computacional.

**Quadro 13** - Resultado da variável Tipologia dos antropônimos

<b>Tipologia</b>	<b>Presença</b>	<b>Ausência</b>	<b>P.R.</b>
<i>Prenome</i>	32.8% (170/519)	67.2% (349/519)	0.483
<i>Apelido</i>	28.9% (13/45)	71.1% (32/45)	0.534
<i>Nome de urna</i>	26.9% (21/78)	73.1% (57/78)	0.355
<i>Nome de palco</i>	66.7% (56/84)	33.3% (28/84)	0.662
<i>Sobrenome</i>	87.5% (7/8)	12.5% (1/8)	0.920

**Fonte:** Autoria própria

De acordo como os resultados obtidos pelo programa computacional, o fator que – incorporado à variável tipologia dos antropônimos – mais condicionou a realização marcada do determinante foi o *Sobrenome* com um percentual de 87,5% (7/8) de aplicação, contra apenas 12,5% (1/8) de não realização. O segundo fator mais expressivo para a marcação do artigo definido em sintagmas antroponímicos foi o *Nome de palco* com os seguintes valores: 66,7% (56/84) de realização *versus* 33,3% (28/84) de não realização.

Os fatores *Prenome*, *Apelido* e *Nome de urna*, por sua vez, foram os três componentes da variável que apresentaram os números de não efetivação do artigo maiores do que os números de efetivação. Isto é, o fator *prenome* obteve um percentual de 32.8% (170/519) de presença contra 67.2% (349/519) de ausência. Já o fator *Nome de urna* foi o que menos

condicionou a realização do artigo definido em sintagmas antroponímicos, haja vista que o valor de ausência foi bem mais expressivo do que o de presença: 26.9% (21/78) e 73.1% (57/78), respectivamente. Já o fator *Apelido* também demonstrou não ser um forte condicionador da partícula determinante, com um percentual de aplicação de 28.9% (13/45) em detrimento de 71.1% (32/45) de não aplicação.

Ao nos debruçarmos sobre o P.R. ofertado pelo *GoldVarbX*, observamos que o fator mais relevante para a realização marcada do determinante nos sintagmas antroponímicos controlados por essa variável foi o *Sobrenome* com um P.R. quase categórico de 0.920, justamente pelo número mais reduzido de ocorrências. Em segundo lugar, temos a variável *Nome de palco* com P.R. de 0.662. A variável *Apelido* apresentou o P.R. de 0.534 e as variáveis *Nome de urna* e *Prenome* tiveram um baixo índice de aparição do artigo, conforme podemos verificar, respectivamente, por meio do P.R.: 0.355 e 0.483.

Entretanto, é válido atentarmos para os números de ocorrências de cada um dos fatores, e não apenas para os dados percentuais. Se observarmos com mais precisão os números de aparições dos tipos de antropônimos, percebemos que *Prenomes* e *Nomes de palco* foram os que se mostraram mais recorrentes, isto é, são mais comuns nos enunciados do que os outros. Isso demonstra que a relevância de alguns fatores com poucas ocorrências – como é o caso da variável *Sobrenome* – se dá, justamente, por esse número reduzido de sentenças; especialmente quando comparados àqueles conjuntos tipológicos que são mais recorrentes. Desse modo, só um número mais expressivo de dados proporcionará uma discussão mais aprofundada acerca da tipologia dos antropônimos. Todavia, de modo amplo, podemos assumir aqui que a segunda variável mais significativa, *Nome de palco*, é uma das principais agentes da variação para essa variável, dado o seu número de ocorrências com artigo (56/84).

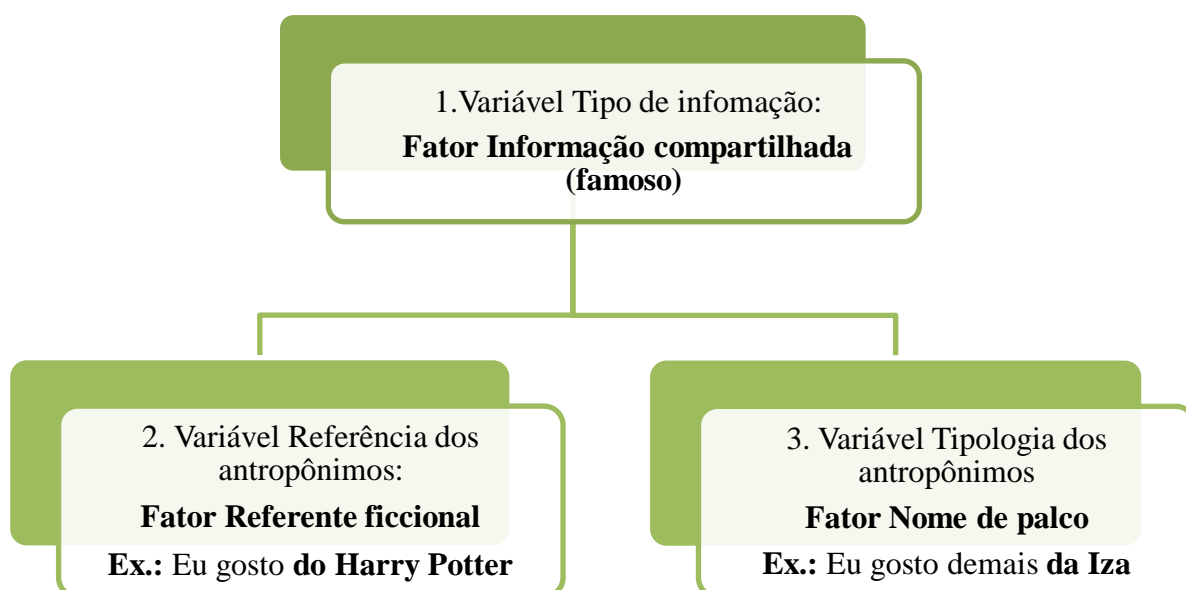
Um fato interessante que merece destaque é a maneira como as três variáveis linguísticas independentes selecionadas pelo programa estão conectadas. Conforme observamos no quadro acima, as variáveis *Tipo de informação*, *Referência do antropônimo* e *Tipologia do antropônimo* parecem condicionar de maneira expressiva o uso marcado do determinante em sintagmas antroponímicos.

Indo mais adiante nessas análises, percebemos que a variável *Tipo de informação* se caracteriza como sendo mais ampla em relação às variáveis *Referência* e *Tipologia* dos antropônimos, isso porque seu fator significativo é capaz de abranger os dois fatores

significativos das outras duas variáveis<sup>88</sup>. Isso pode nos indicar que, mesmo que seja um *personagem* e/ou um *nome de palco*, o antropônimo licenciado com a presença do artigo carrega a particularidade de ser, antes de tudo, um referente famoso: *uma informação compartilhada*.

Em outras palavras, *Personagem* e *Nome de palco* são os fatores significativos das variáveis *Referência* e *Tipologia dos antropônimos*, respectivamente, e ambos estão inclusos no fator relevante da variável *Tipo de informação*, que é *Informação compartilhada*; já que, nos dois casos, designam um referente famoso. A seguir, apresentamos um esquema que ilustra melhor nossa ideia:

**Esquema 4** - Variáveis e fatores favoráveis à marcação do determinante em Recife



**Fonte:** Autoria própria

Conforme exposto no esquema acima, sugerimos que a variável *Tipo de informação* pode ser alocada em posição hierárquica em relação às variáveis *Referência* e *Tipologia dos antropônimos*. O que nos faz lançar mão da seguinte hipótese generalizante para os nossos resultados obtidos até o momento: (i) o *Tipo de informação* parece condicionar significativamente a presença do artigo definido em sintagmas antroponímicos na capital pernambucana. Assim, quanto mais famoso for o referente, mais chances de o antropônimo que o referencia ser licenciado com artigo definido.

<sup>88</sup> *Personagem* e *nome de palco*, respectivamente.

É válido observarmos, que, como já havíamos percebido a influência da variável *Tipo de Informação*, algumas hipóteses sobre tal variável já foram lançadas, com o intuito de explicar sua importância para o uso do artigo definido na capital pernambucana (seção 4.3.2). Em função dessa constatação – de que tal variável está hierarquicamente posicionada acima das outras – realizaremos o cruzamento das variáveis linguísticas (na seção 4.4) sempre em função da variável *Tipo de Informação*.

Isso posto, obtivemos as seguintes constatações acerca das variáveis apresentadas como significativas para a variação do artigo na cidade do Recife:

- (i) Por questões sociais e linguísticas, o fator *masculino* da variável **sexo** favoreceu mais a realização do artigo do que o fator *feminino*;
- (ii) A *primeira faixa etária* beneficiou consideravelmente o uso de sintagmas antroponímicos com artigos definidos;
- (iii) A variável **tipo de informação** foi apontada como sendo a principal variável linguística relevante para este trabalho, uma vez que o fator *informação compartilhada* favoreceu expressivamente a presença do artigo definido. Além disso, levantamos a hipótese de que essa variável liga-se hierarquicamente às outras variáveis linguísticas apontadas como relevantes, mostrando-se uma das principais agentes da variação na fala da comunidade;
- (iv) O fator *personagem* da variável **referência do antropônimo** foi, em termos percentuais, um dos fatores que mais condicionou o uso marcado do artigo definido diante de antropônimos na cidade;
- (v) Os fatores *nome de palco* e *sobrenome* da variável linguística **tipologia dos antropônimos** favoreceram de modo significativo a presença do artigo definido diante de nomes próprios de pessoas na capital pernambucana.

Na próxima seção, iremos apresentar os resultados percentuais das variáveis apontadas como não sendo significativas para a variação do artigo definido na capital do estado de Pernambuco.

### 5.3 VARIÁVEIS APONTADAS COMO NÃO SIGNIFICATIVAS PELO PROGRAMA *GOLDVARBX*

Discutiremos, nesta seção, os resultados percentuais das variáveis apontadas como não significativas pelo programa computacional *GoldVarbX*. É adequado ressaltarmos que, além de demonstrar as variáveis mais expressivas para a variação, o programa também exhibe aquelas que exerceram pouco, ou quase nenhum, alcance para a variação do fenômeno em análise. Desse modo, acreditamos na importância de se discutir todos os dados alcançados, para que assim sejamos capazes de traçar conjecturas que deem conta de elucidar todas as nuances do fenômeno na comunidade de fala investigada.

Tomando por base tais apontamentos, destacamos que todas as variáveis apontadas como não significativas para a variação do artigo definido diante de antropônimos pelo programa computacional na comunidade de fala pesquisada são de base linguística, a saber: *contexto preposicionado*; *função sintática* e *status informacional*. Assim sendo, descreveremos nas próximas subseções cada uma delas, respectivamente.

#### 5.3.1 Contexto preposicionado

Conforme já foi exposto no capítulo metodológico deste trabalho, a variável *contexto preposicionado* leva em consideração se os antropônimos selecionados para investigação são precedidos, ou não, por preposições. É importante destacarmos que, para esta investigação, não levamos em consideração os tipos de preposições – assim como investigado por Pereira (2017). Nesses termos, apresentaremos, no quadro abaixo (14), os resultados percentuais da variável.

**Quadro 14** - Resultado da variável não significativa Contexto preposicionado

<b>Contexto preposicionado</b>	<b>Presença</b>	<b>Ausência</b>	<b>P.R.</b>
<i>Sem preposição</i>	33.5% (197/588)	66.5% (391/588)	0.519
<i>Com preposição</i>	47.9% (70/146)	52.1% (76/146)	0.425

**Fonte:** Autoria própria

De acordo com o quadro acima, podemos perceber que os dois fatores da variável **contexto preposicionado** favorecem mais a ausência do artigo definido diante de antropônimos do que a presença da partícula determinante. Isso se mostra evidente à medida que percebemos que o fator *com preposição* beneficiou em 52.1% (76/146) a ausência dos artigos; enquanto a presença do determinante, mesmo mostrando certa equivalência percentual, foi menos expressiva: 47.9% (70/146). Já o fator sem preposição obteve um percentual pouco significativo de 33.5% (197/588) de presença do determinante contra 66.5% (391/588) de ausência.

Ainda que não tenha se mostrado relevante para variação – considerando que os resultados gerais fornecidos pelo programa computacional *GoldVarbX* descartaram tal variável – é válido observarmos que o fator *com preposição* beneficiou bem mais a realização do artigo definido do que o fator *sem preposição*. Essa constatação, de certa maneira, confirma nossa hipótese para essa variável, haja vista que, conforme pontuam Callou e Silva (1997), também acreditávamos que contextos preposicionados iriam favorecer a aplicação do artigo definido.

Mesmo não sendo o principal elemento linguístico responsável pela variação na comunidade de fala, a presença de preposição parece, ainda que de modo sutil, se destacar na variável. Isto é, quando observamos o percentual de presença do artigo definido nos casos preposicionados (47.9% (70/146)), percebemos que o fator *com preposição* se mostra ligeiramente mais condicionador do que o fator *sem preposição*. Ou seja, os casos em que não há preposição a presença do artigo decai para 33.5% (197/588).

Ao levarmos em consideração o P.R. provido pelo programa computacional, percebemos a ligeira expressividade da variável *sem preposição*, ainda que ambas as variáveis tenham mantido certa equivalência no que concerne ao não condicionamento dos artigos definidos diante de nomes próprios de pessoas: 0.425, *com preposição*; 0.519 *sem preposição*.

Uma hipótese que talvez seja capaz de explicar a baixa relevância da variável *contexto preposicionado* para a variação do artigo definido na capital pernambucana pode ser, justamente, a ocorrência expressiva de antropônimos na posição de sujeito. Se levarmos em consideração que grande parte das ocorrências de sintagmas antroponímicos se deu nesse ambiente sintático, podemos conjecturar que os casos em os antropônimos vinham antepostos por uma preposição foram mais reduzidos; fator que não deu significância à variável e, consequentemente, fez com que ela fosse descartada pelo programa por não ser uma das agentes da variação na cidade do Recife.

De tal maneira, na próxima subseção, falaremos sobre os resultados atribuídos à variável **função sintática**.



### 5.3.2 Função sintática

A Variável **função sintática**, conforme já foi discutido neste trabalho, diz respeito ao posicionamento dos antropônimos e/ou sintagmas antroponímicos nas sentenças estudadas. Ao considerarmos tal variável, acreditávamos que, de modo amplo, ela iria exercer influência em nossos dados analíticos, assim como o fez nos dados de Pereira (2017). Todavia, além de ser apontada como não significativa pelo programa GoldVarbX, apenas um de seus fatores se mostrou condicionador do artigo definido: a posição sintática de *objeto indireto*. De tal modo, indo de encontro a nossa hipótese para tal variável, a posição sintática dos antropônimos não se mostrou muito significativa para o recorte analítico feito nesta pesquisa. A seguir, ilustraremos melhor essas considerações:

**Quadro 15** - Resultado da variável não significativa Função sintática

<b>Função sintática</b>	<b>Presença</b>	<b>Ausência</b>	<b>P.R.</b>
<i>Sujeito</i>	30.7% (106/345)	69.3% (239/345)	0.456
<i>Adjunto</i>	33.3% (2/6)	66.7% (4/6)	0.453
<i>Complemento nominal</i>	37.5% (18/48)	62.5% (30/48)	0.574
<i>Aposto</i>	19.3% (11/57)	80.7% (46/57)	0.455
<i>Tópico</i>	46.2% (6/13)	53.8% (7/13)	0.598
<i>Objeto direto</i>	40.3% (54/134)	59.7% (80/134)	0.555
<i>Objeto indireto</i>	56% (65/116)	44% (51/116)	0.604
<i>Predicativo do sujeito</i>	33.3% (5/15)	66.7% (10/15)	0.488

**Fonte:** Autoria própria

Conforme podemos observar no quadro acima, o fator *objeto indireto* foi a única posição sintática que apresentou um percentual de presença do artigo definido maior do que o número de ausência: 56% (65/116) de realização do artigo contra 44% (51/116) de ausência. A função

de *tópico*, por sua vez, apresentou 46.2% (6/13) de presença do determinante *versus* 53.8% (7/13) de ausência. Seguindo o mesmo padrão de ausência mais elevada, o fator *objeto direto* expôs 40.3% (54/134) de presença e 59.7% (80/134) de ausência.

O fator *adjunto* apresentou um percentual de 33.3% (2/6) contra 66.7% (4/6) de ausência. Já os fatores *sujeito*, *complemento nominal e predicativo do sujeito* expuseram os respectivos valores de presença do artigo 30.7% (106/345), 37.5% (18/48), 33.3% (5/15) contra valores bem mais expressivos de ausência: 69.3% (239/345); 62.5% (30/48); 66.7% (10/15), respectivamente. Por fim, o fator *aposto* foi o que menos condicionou o uso marcado de determinantes em sintagmas antroponímicos na comunidade de fala: 19.3% (11/57) de ausência contra 80.7% (46/57) de presença.

É importante ressaltarmos que em algumas variáveis houve um número reduzido de ocorrências, como podemos verificar nos fatores *tópico* (13), *predicativo do sujeito* (6) e *adjunto* (13). Em outros casos, algumas funções sintáticas foram naturalmente bem mais recorrentes em nosso *corpus* de pesquisa, como é o caso da função de *sujeito*, por exemplo, com 345 ocorrências. Essas constatações são importantes porque, mesmo que de maneira pouco expressiva, todas as funções sintáticas selecionadas para esta análise foram contempladas com ocorrências, assegurando representatividade à variável linguística. É válido observarmos, também, que esse movimento se mostra bastante comum em variáveis com um número mais expressivo de fatores, como é o caso da função sintática. Isto é, à medida que aumentamos os fatores, as ocorrências se redistribuem, fazendo com que alguns fatores sejam mais recorrentes do que outros.

Como já era esperado, apenas o P.R. da função sintática *objeto indireto* se mostrou levemente significativo: 0.604. Seguindo o critério de relevância, tivemos os fatores *tópico*, *complemento nominal* e *objeto direto* com o P.R. baixo, porém bastante semelhante: 0.598, 0.574, 0.555, respectivamente. Seguindo o critério de menos significância, incluímos os fatores: *sujeito*, *adjunto*, *aposto* e *predicativo do sujeito*, com valores também equivalentes de não significância: 0.456; 0.453; 0.455 e 0.488, respectivamente.

Sobre a relevância do fator *objeto indireto*, lançamos mão da hipótese de que a presença da preposição pode ter influenciado o maior número de aparições do artigo nessa posição sintática. Conforme observamos na variável *contexto preposicionado* (4.3.1), mesmo não havendo grande significância para a variável linguística dependente como um todo – visto que a variável *contexto preposicionado* foi descartada pelo programa, o fator *com preposição*

obteve um número mais acentuado de presença do artigo do que o percentual de ausência<sup>89</sup>. Ora, se o objeto indireto carrega a particularidade de ser licenciado, imprescindivelmente, com a presença de preposições, podemos supor que a relevância de tal fator tenha partido dessa ligeira preferência que os informantes parecem ter de usar artigos definidos quando há uma preposição antecedendo o antropônimo.

É válido darmos atenção, também, para o fator *objeto direto* pertencente à variável função sintática. Conforme observamos em nossos resultados percentuais, o fator *objeto direto* teve um percentual de presença do artigo de 40.3% (54/134) contra 59.7% (80/134) de ausência. Nessa esteira de pensamentos, acreditamos que o número levemente elevado de ocorrências do artigo definido conferido ao fator *objeto direto* pode se relacionar, sobretudo, ao número expressivo de sintagmas antroponímicos em contextos existenciais<sup>90</sup> pertencentes a essa função sintática. Isto é, conforme observarmos no segundo capítulo desta dissertação, há contextos existenciais nos quais os antropônimos que são licenciados após o verbo “ter” são considerados objetos direto desse verbo. Os exemplos, retirados de nosso *corpus* de pesquisa, ilustram melhor essas ocorrências:

(62) “Tem o **Selton Melo**, maravilhoso.”

(63) “Tem **Gilberto** que tá fechando.”

(64) “Tem **Raul**, meu namorado.”

Assim como observamos no segundo capítulo da pesquisa em tela, para Müller e Negrão (1998) construções como as expressas em (63) e (64) não parecem ser usais para as autoras. Essa elucidação se dá porque, de acordo com as estudiosas, sentenças existenciais sem a presença do artigo definido antepondo o antropônimo (cf. (63) e (64)) são apontadas como agramaticais no dialeto delas. Todavia, assim como já foi pontuado por Oliveira (2018), esse tipo de construções existenciais – sem os determinastes – são amplamente aceitas no Sertão do Pajeú pernambucano, o que aponta que há em Pernambuco um padrão bem diferente do que propõem Müller e Negrão (1998).

Nessa linha argumentativa, podemos induzir que o mesmo padrão – verificado por Oliveira (2018) – também parece se repetir na capital do estado, haja vista que sentenças como

---

<sup>89</sup> Observar subseção anterior.

<sup>90</sup> Acreditamos que o alto índice de construções existências esteja, intimamente, relacionado ao questionário de pesquisa que, muitas vezes, usa perguntas que direcionam para a utilização desse tipo de ocorrências com o verbo “ter”.

as exibidas em (63) e (64) foram bastante recorrentes em nosso *corpus* de pesquisa. Ainda assim, é oportuno convencionarmos que inúmeras ocorrências em que a posição do determinante estava sim marcada pelo artigo definido no sintagma antroponímico (cf. 62) foram encontradas nesta investigação, pressupondo, assim, que as duas formas são possíveis e existem na cidade do Recife.

É importante observamos, também, que o fator *tópico* apresentou uma quase equivalência entre a presença e a ausência do artigo em nossos dados analíticos, sobretudo quando levamos em consideração o número de ocorrências: 46.2% (6/13) de ausência contra 53.8% (7/13) de presença. Contudo, observamos algumas questões importantes acerca dessa posição sintática que merecem ser discutidas mais profundamente; por essa razão, discorreremos, especificamente, sobre alguns casos de topicalização na seção 5.5 desta pesquisa.

Na subseção seguinte, iremos apresentar os resultados percentuais da última variável descartada pelo programa computacional: a variável *status* informacional.

### 5.3.3 *Status* informacional

Nesta subseção, apresentaremos os resultados percentuais atribuídos à variável *status* informacional. Em conformidade com o que já foi discutido aqui, sabemos que essa variável diz respeito à inserção do antropônimo no contexto comunicativo, isto é, se ele foi citado pela primeira vez (informação nova) e/ou retomado no discurso (informação velha).

De acordo com os estudos de Callou e Silva (1997), o fator *informação nova* favorece a realização de artigos definidos; em face dessa assertiva, controlamos tal hipótese acreditando que o mesmo se daria na cidade do Recife. Todavia, tal proposição não se confirmou, dado que – além de a variável ***status* informacional** não ter sido significativa para a variável depende investigada neste trabalho – o fator *informação nova* anulou mais a presença do artigo definido do que o fator *informação velha*, conforme ilustrado no quadro (16). Assim sendo, nossa hipótese para essa variável também não se confirmou.

**Quadro 16** - Resultado da variável não significativa *Status* informacional

<b>Status Informacional</b>	<b>Presença</b>	<b>Ausência</b>	<b>P.R</b>
<i>Informação nova</i>	35.2% (180/512)	64.8% (332/512)	0.496
<i>Informação velha</i>	39.2% (87/222)	60.8% (135/222)	0.510

**Fonte:** Autoria própria

À medida que observamos os resultados expostos no quadro acima, fica mais evidente que tal variável não se mostra relevante para a marcação do artigo definido diante de nomes próprios de pessoas na cidade do Recife. Conforme podemos verificar, o fator *informação nova* teve o percentual pouco expressivo de 35.2% (180/512) de presença do artigo, ao passo que o percentual de ausência foi bem mais significativo: 64.8% (332/512). Seguindo o mesmo padrão de baixa expressividade, o fator *informação velha* também parece anular a presença do artigo, haja vista os percentuais apresentados: 60.8% (135/222) de ausência contra 39.2% (87/222) de presença.

O P.R. fornecido pelo programa computacional também confirma a hipótese de quase insignificância da variável *status* informacional para a variação na comunidade de fala. O fator *informação velha* apresentou o P.R. levemente acentuado de 0.510, ao passo que o fator *informação nova* exibiu o P.R. menos expressivo de 0.496.

Frente às constatações expostas até aqui, sobre as variáveis apontadas como não significativas pelo programa computacional *GoldVarbX*, apresentamos as seguintes inferências conclusivas: (i) mesmo não exercendo significância para a variação, o fator *com preposição* – da variável **contexto preposicionado** – favoreceu, de modo mais sutil, o uso do artigo do que o fator *sem preposição*; (ii) na variável **função sintática**, apenas o fator *objeto indireto* obteve um número de presença do determinante maior do que o percentual de ausência, por isso, acreditamos na hipótese de que esse resultado advém, justamente, do uso mais expressivo de preposições atribuído à variável contexto preposicionado; (iii) acerca da variável **status informacional**, nenhum dos fatores se mostrou relevante para a variação, haja vista que ambos pareceram anular, de modo equivalente, a presença do artigo definido na comunidade de fala.

À vista disso, acreditamos que as observações acerca das variáveis não significativas proporcionam certa robustez à análise dos dados. Ao levantarmos hipóteses sobre como (e por

que) alguns fatores não exerceram tanta influência para a variação, estamos delineando e, de certo modo, catalogando o comportamento da variável linguística dependente na comunidade de fala estudada. Assim sendo, nos incumbiremos a descrever, na próxima seção, os resultados percentuais do cruzamento de dados realizado entre as variáveis linguísticas apontadas como significativas para esta pesquisa.

#### 5.4 CRUZAMENTO DE ALGUMAS VARIÁVEIS SIGNIFICATIVAS

Nesta seção, elaboraremos o cruzamento de dados entre algumas das variáveis linguísticas apontadas como significativas pelo programa computacional *GoldVarbX*, com vistas a isolar, de modo mais preciso, quais são os fatores mais relevantes para a realização do artigo definido na cidade do Recife. Além disso, realizaremos – de modo adjacente – o cruzamento de dados entre as duas variáveis extralinguísticas apontadas como relevantes para a variação na comunidade de fala, a saber: *sexo e faixa etária dos informantes*.

Em conformidade com as exposições levantadas na seção anterior, presumimos que a variável *tipo de informação* – além de ser designada como significante para a realização do artigo definido em sintagmas antropônimos na comunidade de fala – pode ser alocada em disposição hierárquica em relação às variáveis *referência e tipologia dos antropônimos*. Isto é, ao considerarmos que o fator *informação compartilhada* (referente famoso) pode abranger os fatores relevantes das outras variáveis (personagem e nome de palco), alçamos a hipótese de que tal variável é uma importante agente da variação na cidade do Recife. Em face disso, optamos por realizar o cruzamento das variáveis linguísticas tomando por base a variável *tipo de informação*.

Assim sendo, na próxima subseção apresentaremos o cruzamento de dados edificado entre as variáveis *tipo de informação e referência dos antropônimos*.

##### 5.4.1 Tipo de informação *versus* Referência dos antropônimos

Optamos por realizar o cruzamento entre as variáveis significativas *tipo de informação e referência dos antropônimos* considerando a importância que a primeira variável parece exercer nos dados de língua falada na cidade do Recife. Por essa razão, investigamos de que

maneira essas duas variáveis linguísticas significativas se comportam, sobretudo quando analisadas de maneira interligada.

No quadro 17, a seguir, observaremos os percentuais de presença e ausência dos determinantes de acordo com o cruzamento dos fatores das duas variáveis linguísticas relevantes selecionadas: tipo de informação e referência dos antropônimos.

**Quadro 17** - Cruzamento de dados entre as variáveis Tipo de informação e Referência dos antropônimos

Referência do antropônimo		Tipo de informação			
		<i>Informação não compartilhada (não famoso)</i>		<i>Informação compartilhada (famoso)</i>	
		Frequência	%	Frequência	%
<i>Real</i>	Com artigo	84/433	19%	103/215	48%
	Sem artigo	349/433	81%	112/215	52%
<i>Personagem</i>	Com artigo	4/5	80%	179/296	60%
	Sem artigo	1/5	20%	117/296	40%

**Fonte:** Autoria própria

Conforme exposto no quadro acima, podemos observar que os percentuais de presença e ausência dos determinantes são quase equivalentes quando o antropônimo é um ser real: 48% de presença e 52% de ausência. Além disso, parece prudente inferirmos que, quando se trata de um *personagem famoso*, a possibilidade de o antropônimo ser habilitado com a posição de determinante marcada é bastante expressiva. Em outras palavras, podemos sugerir que, quando o antropônimo carrega essas duas particularidades (ser um personagem e um informação compartilhada), há uma grande possibilidade de o sintagma antroponímico ser licenciado com artigo definido na cidade do Recife.

É oportuno observarmos, também, que o fator *personagem* – da variável referência do antropônimo – condiciona significativamente o uso marcado do determinante nos dois fatores da variável tipo de informação: informação compartilhada (60%) e informação não compartilhada (80%). Essas constatações elucidam, de modo ainda mais preciso, a interferência dos fatores relevantes de cada uma das variáveis significativas e, também, a maneira como elas se correlacionam entre si.

Isso posto, na subseção seguinte discorreremos acerca da associação entre as variáveis *tipo de informação* e *tipologia dos antropônimos*.

### 5.4.2 Tipo de informação *versus* Tipologia dos antropônimos

O cruzamento entre as variáveis *tipo de informação* e *tipologia dos antropônimos* se deve ao fato de acreditarmos que essas duas variáveis têm uma espécie de correlação. De tal modo, é válido investigarmos com se dá essa interseção e seus possíveis desdobramentos para o entendimento do fenômeno variável na comunidade de fala investigada.

A variável *tipo de informação* pode ser considerada binária, já que dispõe de apenas dois fatores: informação compartilhada (famoso) e informação não compartilhada (não famoso). A variável *tipologia dos antropônimos*, por sua vez, compreende um número bem mais elevado de fatores<sup>91</sup>. Contudo, em nossos dados analíticos obtivemos um nível proeminente de recorrência apenas nos seguintes fatores: prenome, apelido, nome de urna, nome de palco e sobrenome. À vista disso, observaremos, no quadro (18) abaixo, de que maneira esses fatores se correlacionam.

**Quadro 18** - Cruzamento de dados entre as variáveis Tipo de informação e Tipologia dos antropônimos

Tipologia dos antropônimos		Tipo de informação			
		Informação não compartilhada		Informação compartilhada	
		Frequência	%	Frequência	%
<i>Prenome</i>	Com artigo	74/391	19%	96/128	75%
	Sem artigo	317/391	81%	32/128	25%
<i>Apelido</i>	Com artigo	4/35	11%	9/10	90%
	Sem artigo	31/35	89%	1/10	10%
<i>Nome de Urna</i>	Com artigo	4/5	80%	17/73	23%
	Sem artigo	1/5	20%	56/73	77%
<i>Nome de palco</i>	Com artigo	-	-	56/82	66%
	Sem artigo	-	-	26/82	34%
<i>Sobrenome</i>	Com artigo	4/5	80%	3/3	100%
	Sem artigo	1/5	20%	-	-

**Fonte:** Autoria própria

Ao analisarmos o quadro acima, percebemos um parâmetro importante entre os índices mais elevados de presença do artigo definido: a relevância da variável *tipo de informação*. Se observarmos os dados percentuais atribuídos à *informação compartilhada*, fica evidente que

<sup>91</sup> Ver o segundo capítulo deste trabalho na seção em que nos debruçamos sobre a tipologia dos nomes próprios.



esse fator condiciona exponencialmente a presença do determinante em quase todos os tipos de antropônimos da variável; com exceção apenas do fator *nome de urna* – que apresenta um número de ausência mais elevado do que o percentual de presença do artigo definido.

A significância da variável *tipo de informação* pode ser observada, também, se tomarmos por base o fator *prenome* da variável *referência dos antropônimos*. Isto é, o fator tipológico com maior número de ocorrências entre todos os outros fatores da variável (447) foi *o prenome*; entretanto, este só se mostrou sensível à presença do artigo definido quando condicionado pelo fator *informação compartilhado*, com um percentual expressivo de 75%. Em contrapartida, quando se tratava do fator *informação não compartilhada*, esse percentual caiu significativamente para 19%, demonstrando a importância de atrelarmos as duas variáveis.

Ao concentrarmos atenção especial aos fatores tipológicos *nome de palco* e *sobrenome*, percebemos, mais uma vez, a atuação que o fator *informação compartilha* exerce para a tipologia dos antropônimos. Nos nomes de palco, por exemplo, não encontramos nenhuma ocorrência em que o antropônimo – que era também uma informação compartilha – estava licenciado com artigo definido. Em se tratando do fator *sobrenome*, percebemos que todas as ocorrências que eram informações compartilhadas ocorreram com presença do determinante (100% (3/3)). Essas constatações se mostram relevantes à medida que somos capazes de perceber, por meio de dados percentuais, a importância do fator *informação compartilhada* no agenciamento da variação na comunidade de fala.

Nesses termos, na próxima subseção, ilustraremos alguns resultados percentuais obtidos por meio do cruzamento das duas variáveis sociais escolhidas para esta pesquisa: *sexo* e *faixa etária* dos informantes.

### 5.4.3 Sexo versus Faixa etária

No quadro 19, a seguir, apresentaremos os resultados percentuais das variáveis extralinguísticas *sexo* e *faixa etária*. Nosso principal objetivo com essa intersecção de dados é, justamente, especificar quais são os perfis dos informantes que mais colaboram para variação na comunidade de fala investigada e quais são aqueles que mais anulam o uso da variável inovadora.

**Quadro 19** - Cruzamento de dados entre as variáveis extralinguísticas Sexo e Faixa etária

Referência do antropônimo		Tipo de informação			
		Feminino		Masculino	
		Frequência	%	Frequência	%
Primeira faixa etária	Com artigo	53/142	37%	173/332	52%
	Sem artigo	89/142	63%	159/332	48%
Segunda faixa etária	Com artigo	40/224	18%	1/36	1%
	Sem artigo	184/224	82%	35/36	99%

**Fonte:** Autoria própria

Mediante análises preliminares, realizadas antes do cruzamento de dados, já suspeitávamos da relação acentuada entre as variáveis *sexo* e *faixa etária*. Após observarmos os resultados percentuais expressos no quadro acima, podemos confirmar que o uso acentuado do artigo definido em sintagmas antroponímicos se dá, justamente, entre os homens pertencentes à primeira faixa etária, com um percentual de 52% (173/332) de presença do determinante. Essa assertiva se mostra ainda mais aceita quando comparamos o número quase nulo de presença de artigos definidos atribuído a segunda faixa etária: 1% (1/36).

O sexo feminino, por sua vez, apresentou um percentual pouco expressivo de presença dos determinantes nas duas faixas etárias selecionadas para esta investigação. Contudo, se levarmos em consideração a *segunda faixa etária*, podemos perceber que as mulheres recifenses mais velhas denotam uma leve tendência a empregar mais o artigo definido em sintagmas antroponímicos do que os homens da mesma faixa etária: 18% (40/224) para o sexo feminino contra apenas 1% (1/36) para o sexo masculino.

Por meio dessas constatações, podemos inferir que o perfil mais acentuado de uso dos determinantes na comunidade de fala – de acordo com o recorte analítico proposto nesta pesquisa – se concentra entre os homens mais jovens. Ou seja, os homens pertencentes à primeira faixa etária são mais sensíveis a forma inovadora na cidade do Recife, que diz respeito ao uso marcado de artigos definidos. Esses resultados, por sua vez, podem estar diretamente relacionados ao alto índice de concordâncias inversas empregadas exatamente pelos homens mais jovens da comunidade, conforme já discutimos na seção (4.2.1) desta pesquisa.

À vista do que foi debatido nos limites desta seção, acreditamos na importância efetiva do cruzamento de dados, sobretudo, porque esse movimento analítico nos permite isolar, especificamente, quais são os principais agentes da variação na comunidade de fala investigada.

Assim sendo, apresentaremos a seguir – na última seção da pesquisa em tela – alguns aspectos estruturais acerca da antroponímia pernambucana.

## 5.5 ALGUNS ASPECTOS ACERCA DA ANTROPONÍMIA NA CAPITAL PERNAMBUCANA

Com base nas discussões propostas no curso desta investigação dissertativa, observaremos, nesta seção, alguns aspectos que julgamos ser extremamente relevantes para o estudo mais amplo da antroponímia na capital pernambucana. É bem verdade que o principal objetivo deste trabalho consistiu em um levantamento quantitativa acerca da variação do artigo definido na cidade do Recife e, de modo adjacente, observamos quais foram as principais variáveis que condicionaram tais usos. Todavia, acreditamos na importância de irmos mais além nos objetivos aqui propostos para isolarmos e analisarmos, também, alguns dados em que a variação – principal objeto investigativo desta pesquisa – não pode ser concebida na comunidade de fala.

Em face dessa problemática, e indo mais adiante nessas investigações, observaremos que há casos em que a variação do artigo definido parece não ser permitida na comunidade de fala; isto é, casos em que há não-opcionalidade para o uso variável do artigo definido. Ao longo das nossas análises, percebemos dois contextos em que a variação parece dar lugar a usos não-opcionais na comunidade de fala: tanto para a marcação obrigatória do artigo definido (com artigo) quando para sua completa ausência em sintagmas antroponímicos (sem artigo).

Constatamos, pelo menos, três casos em que a não-opcionalidade parece ser conferida aos nossos dados. Em apenas um desses casos há a anulação obrigatória do determinante: (i) caso de não-opcionalidade em que o antropônimo é utilizado como etiqueta linguística. Nos outros dois casos de não-opcionalidade, por sua vez, é a marcação do determinante que passa a ser obrigatória: (ii) caso de não-opcionalidade no qual os antropônimos são individualizados por meio de um modificador, (iii) caso de não-opcionalidade em que existe a duplicação do antropônimo descolado à esquerda da sentença. Assim sendo, nas subseções que se seguem mostraremos – a título de exemplo – alguns contextos retirados do nosso *corpus* e as hipóteses, ainda que incipientes, que lançamos para cada um desses contextos.

### 5.5.1 Antropônimos sem artigos: caso de não-opcionalidade em que o antropônimo é utilizado como etiqueta linguística

Com a assistência das informações destacadas por Muller e Negrão (1988) no segundo capítulo deste trabalho, observamos que existem contextos sintático-semânticos nos quais a estrutura [artigo definido + nome próprio] parece não ser permitida no PB falado em Recife, isto é, antropônimos enquanto etiquetas linguísticas e/ou construções denominativas (AMARAL e SEIDE, 2020). Tomando por base essas afirmações teóricas, localizamos – em nossos dados analíticos sobre a capital pernambucana – ocorrências em que o nome próprio de pessoa mencionado no discurso se referia, justamente, à etiqueta linguística dos informantes e não a pessoa em si do discurso. Os exemplos, a seguir, retirados do nosso *corpus* de pesquisa ilustram melhor essas informações:

(65) “Ela, na verdade, se chama **Aparecida**.”

(66) “Eu achava muito bonito o nome **Liana**.”

(67) “Na época, chamava ele de **Bruno**.”

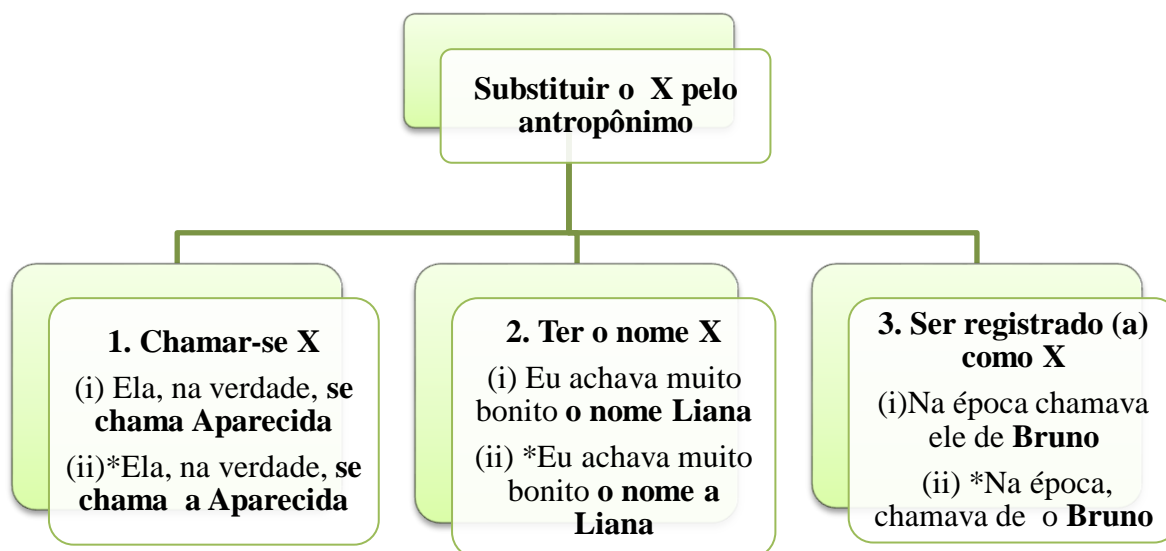
Müller e Negrão (1988) acreditam que em ocorrências como as exemplificadas acima ((65), (66) e (67)) parece não haver opcionalidade para a realização do artigo no PB. Isso porque, a depender da restrição semântica desse antropônimo, ele refere-se ao nome em específico e não a pessoa física que esse nome referencia. Conforme observamos em nosso segundo capítulo, Sedrins (2017) e Oliveira (2018) também fazem coro com as autoras ao reconhecerem que essas ocorrências catalogadas pelas teóricas são, de fato, recorrentes no PB.

No mesmo alinhamento conceitual, Amaral e Seide (2020), como bem sabemos, designam essas ocorrências como sendo construções denominativas. Assim sendo, utilizaremos – no esquema a seguir – as definições dos autores<sup>92</sup> para alocarmos os exemplos ((65), (66) e (67)) provenientes da cidade Recife:

---

<sup>92</sup> Esquema utilizado no segundo capítulo para ilustrar as definições de construções denominativas propostas por Amaral e Seide (2020).

**Esquema 5** - Construções denominativas retiradas do nosso banco de dados



**Fonte:** Autoria própria

Ao nos debruçarmos sobre essas ocorrências, depreendemos que elas elucidam de modo bastante expressivo o que já foi discutido por Müller e Negrão (1988): os antropônimos – na configuração em que estão dispostos nas sentenças – não consentem o licenciamento do artigo definido. Se avaliarmos de maneira semântica as sentenças expressas em (i) no esquema anterior, somos capazes de induzir que os antropônimos não tangenciam as pessoas referenciadas, mas sim os seus nomes – o que optamos por chamar, conforme Amaral e Seide (2020), de **construções denominativas**. Nessa fronteira argumentativa, parece prudente asseverarmos que a colocação do artigo, para esses contextos, é impossível, uma vez que a sua inserção tornaria a sentença, aparentemente, agramatical para o contexto semântico a qual ela se presta<sup>93</sup> (cf. (ii) no esquema acima). Isso posto, na próxima subseção falaremos sobre os antropônimos com uso aparentemente obrigatório de artigos definidos.

<sup>93</sup> É importante ressaltarmos que, para a inserção do \* que representa agramaticalidade no dialeto investigado, estamos considerando a aceitabilidade de um informante recifense; por tanto, sem generalizações que se estendem ao PB.

### 5.5.2 Antropônimos com artigos: nomes próprios de pessoas individualizados por meio de um modificador têm uso obrigatório de artigos definidos?

Em conformidade com o que ressaltamos no segundo capítulo desta pesquisa, acreditamos que, assim como pontua Sedrins (2017), existem casos em que a posição marcada do determinante (artigo definido, nesse caso) parece ser essencial no PB. Para o autor, quando o antropônimo é efetivado com um modificador (adjetivo e/ou sintagma preposicionado) capaz de restringir o referente, a marcação do artigo definido se mostra necessária. Em nossos dados da capital pernambucana, encontramos algumas ocorrências específicas que parecem legitimar o que foi proposto pelo teórico, vejamos:

(68) “**A Monique** da universidade era diferente para mim”

(69) “**O Pedro** bonito é diferente do outro lá...”

Ao nos debruçarmos sobre as ocorrências dispostas acima, presumimos haver mais de um referente com o mesmo nome, por isso a individualização deste antropônimo parece ser necessária. Essa restrição, por sua vez, se dá por meio dos elementos especificadores: [da universidade], sintagma preposicionado e [bonito], adjetivo. Isto é, esses elementos reforçam a ideia de que são esses referentes específicos e não outros (A Monique da universidade e não outra; O Pedro bonito e não outro). Um segundo recurso que parece ser essencial para essa individualização do antropônimo é, justamente, o uso marcado da partícula determinante (cf. (68) e (69)).

Nessa linha argumentativa, acreditamos que esse tipo de individualização – expressa pela presença de um modificador do referente – parece condicionar o uso obrigatório do determinante. Ademais, só um *corpus* de análise mais robusto de dados e, possivelmente, testes de intuição linguística com os informantes da comunidade de fala seriam capazes de confirmar, ou não, essas análises. Assim sendo, na próxima subseção, discorreremos sobre algumas ocorrências de topicalização encontradas em nossos dados.

### 5.5.3 Antropônimos com artigos: nomes próprios de pessoas deslocados à esquerda das sentenças favorecem a aplicação do artigo definido?

Nesta seção, falaremos sobre a terceira hipótese de não-opcionalidade no que tange ao uso variável do artigo definido. Conforme observamos em nossos dados, os casos de topicalização dos antropônimos parecem condicionar, de algum modo, a presença do artigo na comunidade de fala. Essa generalização se mostra mais plausível à medida que observamos que todos os antropônimos com artigos – localizados nesses contextos sintáticos – não faziam parte das outras variáveis já descritas como responsáveis pela variação. Isto é, todos os seis antropônimos licenciados com artigos definidos que ocupavam a posição de tópico das sentenças não eram *referentes famosos*, *personagens*, tampouco, *nomes de palco* e *sobrenomes* (principais fatores condicionadores da variação na comunidade de fala). Tal fator nos faz crer na hipótese de que, de certa maneira, existe uma possível influência da posição sintática de tópico nessas ocorrências.

É importante observarmos que, entre os casos de não-opcionalidade destacados neste trabalho, este episódio de antropônimo deslocado à esquerda das sentenças (topicalização com antropônimos licenciados com a presença de artigos definidos) foi o único que não havia sido previsto em nosso capítulo teórico. Em outras palavras, tanto os casos de construções denominativas, quanto as ocorrências de individualização dos antropônimos por meio de modificadores já haviam sido catalogadas anteriormente por teóricos que elegem os sintagmas antroponímicos como objetos de análise (MÜLLER e NEGRÃO (1988); SEDRINS, (2017) e AMARAL E SEIDE, (2020)). Por essa razão, as observações expostas nesta subseção – sobre os casos de topicalização – serão preliminares e de natureza inteiramente investigativa, haja vista a necessidade de explicarmos melhor essas ocorrências e a maneira como elas se apresentaram em nosso *corpus* de pesquisa.

Em face disso, tentaremos elaborar aqui alguns prognósticos acerca desse tipo específico de construção sintática na qual a posição de determinante se mostrou constantemente assinalada pela presença de artigos definidos. Conforme observaremos nos dados a seguir, as ocorrências em que os sintagmas antropônimos foram descolados para a periferia esquerda das sentenças (para a disposição de tópico), o antropônimo movido estava sempre precedido por um artigo definido:

(70) “**O Daniel**, ele é meio tabacudo”

(71) “**A Mônica**, nossa ela é muito chata.”

(72) “**A Lilian**, ela tem essa mania de mandar em mim”

(73) “**O Sebastião**, ah, Sebastião é meu braço direito aqui em Recife”

Podemos observar aqui, que todos os constituintes antroponímicos estão realizados deslocados à esquerda da disposição sintática. Isto é, parece haver uma relação discursiva de intenção da fala, uma vez que esse antropônimo ganha a posição de ênfase na sentença. É possível observarmos, também, que em três casos acima, o tópico refere-se a uma categoria que está expressa por um pronome pessoal na posição intrassentencial (cf. 70, 71 e 72):

- [O Daniel], ele é meio tabacudo.
- [A Mônica], nossa! ela é muito chata.
- [A Lilian], ela tem essa mania de mandar em mim.

Conforme é possível notarmos, esses deslocamentos à esquerda podem ser conferidos pela retomada, no domínio intrassentencial, de um elemento equivalente aos nomes próprios: o pronome **ela/ ele**, por exemplo. Já no exemplo (73), observamos a ocorrência topicalizada do mesmo antropônimo que se encontra no interior da sentença, isto é, uma repetição do antropônimo. Nesse caso, em particular, percebemos que, ao auferir a posição de tópico sentencial, o antropônimo passa a ser licenciado – assim como nos outros exemplos – com a presença do artigo definido, no entanto, o mesmo parece não se dar quando o antropônimo ocupa a posição intrassentencial de sujeito da sentença:

- [O Sebastião], ah, Sebastião é meu braço direito aqui em Recife.

Como é possível observarmos, esse último exemplo de topicalizações parece retomar um constituinte que se encontra lexicalmente repetido no domínio intrassentencial. Diferente das outras ocorrências, em há que a inserção de um elemento equivalente (o pronome) no interior da sentença. Nesse último caso, em particular, podemos observar que a retomada se dá, justamente, por meio de uma cópia do termo topicalizado: [O Sebastião]; [Sebastião].

À vista do que foi exposto aqui, acreditamos na hipótese de que a posição de tópico das sentenças parece favorecer a realização marcada de artigos definidos diante de antropônimos no português falado na cidade do Recife. Seja para marcar uma ênfase discursiva e/ou destacar o antropônimo deslocado à esquerda, a posição de tópico sentencial parece condicionar a



presença do determinante. Todavia, só uma análise mais acurada e um *corpus* mais robusto de dados linguísticos seriam capazes de conformar e/ou refutar tal hipótese.

Desse modo, no intuito de sistematizarmos nossas observações estruturais, podemos listar as seguintes conclusões acerca desta seção: (i) os dados de Recife – assim como nos estudos de Müller e Negrão (1988) e Amaral e Seide (2020) – parecem sugerir que, quando nos referimos ao nome do indivíduo e não ao indivíduo em si, o uso do artigo é repellido no PB; (ii) conforme aponta Sedrins (2017), nos contextos em que o nome próprio é realizado com um modificador (adjetivo, sintagma preposicionado) a realização do determinante parece ser obrigatória no português falado no Recife; (iii) a posição de tópico das sentenças parece favorecer a presença do artigo definido, dado que os antropônimos que se encontravam na periferia esquerda das sentenças estavam licenciados com o determinante.

Isso posto, apresentaremos, a seguir, nossas conclusões gerais sobre este trabalho de dissertação.

## 6 CONCLUSÕES

A pesquisa em tela foi idealizada, sobretudo, pela necessidade de se atualizar as descrições linguísticas produzidas na cidade do Recife acerca da variação do artigo definido em sintagmas de natureza antroponímica. Isto é, se levarmos em consideração que o último estudo que elaborou esse mapeamento o fez mediante a comparação com outras capitais brasileiras<sup>94</sup>, temos dimensão da importância de se desenvolver uma descrição que – assim como o estudo que fizemos nesta dissertação – priorizou a exposição de vários aspectos linguísticos e sociais alusivos à antroponímia na capital do estado.

A fim de garantir a robustez de nossas análises e uma abrangência da estrutura [artigo definido + antropônimos] no português falado no Brasil, nos baseamos nos estudos de Müller e Negrão (1990); Sedrins (2017), Amaral e Seide (2020), entre outros autores que se dedicam a investigar os aspectos mais estruturais dos sintagmas antroponímicos. Buscamos aparato, também, nas proposições sociolinguísticas desenvolvidas por alguns estudiosos que já observaram o comportamento variável do artigo definido em sintagmas antroponímicos: Callou & Silva (1997), Campos Jr. (2011), Sedrins, Pereira e Siqueira (2015), Pereira (2017) e Oliveira (2018).

De modo sumário, esta pesquisa cooperou para uma descrição mais cuidadosa acerca do perfil sociolinguístico do português falado na capital pernambucana; observando, em primeiro plano, as construções sintáticas que são, ou não, licenciadas com a partícula determinante e, em segunda instância, as principais propriedades que condicionam (ou não) os usos.

Tomando por base os resultados gerais de nossa investigação, constatamos que – assim como sugeria a hipótese geral desta pesquisa – a capital pernambucana tem preferência pela modalidade não marcada do artigo definido em sintagmas antroponímicos. Ou seja, em 63.6% (467/734) dos casos registrados, os informantes não utilizaram o determinante; ao passo que em 36,4% (267/734) das ocorrências, os entrevistados optaram pela presença do artigo.

Os resultados obtidos pelo *GoldVarbX* indicaram que cinco variáveis foram significativas para a variação investigada no Recife: *Sexo*, *Faixa etária*, *Tipo de informação*, *Referência do antropônimo* e *Tipologia do antropônimo*. Sendo as duas primeiras de natureza

---

<sup>94</sup> O principal trabalho elaborado, até então, na capital pernambucana havia sido o mapeamento realizado por Callou & Silva (1997), já descrito no capítulo 2 desta dissertação.

extralinguística e as três últimas de natureza linguística. Assim sendo, observamos que o fator *masculino* (fator significativo dentro da variável sexo) favoreceu em 47.3% (174/368) a presença dos determinantes, ao passo que o percentual de ausência foi de 52.7% (194/368); com o P.R. de 0.587. Acreditamos que os números acentuados de sintagmas antroponímicos licenciados com artigos definidos – atribuídos ao sexo *masculino* – estejam diretamente relacionados com o alto índice de concordâncias inversas de gênero produzidas pelos homens da comunidade de fala.

A variável *faixa etária* dos informantes foi a segunda apontada como significativa para a variação na comunidade de fala. Com um percentual de 47.7% (226/474) de presença do artigo, a primeira faixa etária foi a mais importante para a variação. A segunda faixa etária, por sua vez, obteve apenas 15.8% (41/260) de presença dos determinantes. Os P.R. fornecidos pelo programa também garantem a importância da primeira faixa etária para a variação: 0.635 para a primeira faixa etária e apenas 0.267 para a segunda.

Dando relevância às variáveis linguísticas apontadas como significativas, verificamos que o fator *informação compartilhada*, pertencente a variável *tipo de informação*, obteve um percentual acentuado de presença do artigo: 60.5% (179/296), contra 39,5% (117/296) de ausência. Levando em consideração o P.R., constatamos, ainda mais, a importância deste fator no que concerne ao uso marcado do artigo: 0.743

Sobre a variável *Referência do antropônimo*, abalizamos que o fator *personagem* foi um dos principais agentes para a marcação do artigo na comunidade de fala investigada, com um percentual de aplicação quase categórico de 93% (80/86), contra 7% (06/86) de não aplicação. Se observarmos o P.R. fornecido pelo programa computacional, somos capazes de ratificar a significância expressiva do fator *personagem*: 0.920.

Por fim, ao observarmos a variável tipologia do antropônimo, pensada exclusivamente para esta investigação, percebemos que os fatores *nome de palco* e *sobrenome* foram os que mais condicionaram a realização do artigo nesta variável. O nome de palco obteve um percentual de 66.7% (56/84) de aplicação de determinantes, contra 33.3% (28/84) de ausência e um P.R. expressivo de 0.662. O fator *sobrenome*, por sua vez, favoreceu em 87.5% (7/8) a presença do artigo definido, ao passo que o percentual de ausência foi de 12.5% (1/8). O P.R. obtido foi de 0.920.

Assim sendo, para explicar os dados da pesquisa em tela, fizemos a generalização de que as três variáveis linguísticas apontadas como significativas estão, de certa maneira, conectadas. Isto é, o fator *personagem* e o fator *nome de palco* (pertencentes às variáveis *referência do antropônimo* e *tipologia do antropônimo* – respectivamente) têm a

particularidade de referenciar um indivíduo famoso – informação compartilhada – fator significativo da variável *tipo de informação*.

Ainda de acordo com os nossos resultados gerais, analisamos que as variáveis *função sintática*, *contexto preposicionado* e *status informacional* não foram significativas para a variação do artigo definido na capital pernambucana, haja vista os números pouco expressivos apresentados pelo programa computacional GoldVarbX.

Com base em nosso cruzamento de dados percentuais, obtivemos algumas conclusões importantes: i) os homens mais jovens são os informantes mais sensíveis à variante inovadora na comunidade de fala, isto é, os homens pertencentes a primeira faixa etária usam mais o artigo definido em sintagmas antroponímicos do que as mulheres; ii) antroponimos que referenciam personagens famosos (das variáveis *referência dos antroponimos* e *tipo de informação*) condicionam exponencialmente o uso dos determinantes; iii) *sobrenomes* e *nomes de palco* condicionam significativamente a presença de determinantes, sobretudo, quando os antroponimos desta natureza são, também, informações compartilhadas (referentes famosos).

Conforme sabemos, o objetivo principal desta pesquisa foi apresentar uma análise quantitativa acerca da variação do artigo definido em sintagmas antroponímicos na cidade do Recife. De tal modo, constatamos que (sim) algumas variáveis linguísticas e extralinguísticas estão agindo para o uso variável encontrado na comunidade de fala. Indo mais adiante em nossos objetivos (gerais e específicos), isolamos e discutimos, também, alguns casos nos quais consideramos não haver opcionalidade para o uso variável em recife, a saber: (i) caso de não-opcionalidade em que o antroponimo é utilizado como etiqueta linguística; (ii) caso de não-opcionalidade no qual os antroponimos são individualizados por meio de um modificador, (iii) caso de não-opcionalidade em que existe a duplicação do antroponimo descolado à esquerda da sentença.

É válido ressaltarmos que, por ora, não realizamos o cruzamento de dados entre as variáveis apontadas como não significativas, por acreditarmos que esse movimento analítico não iria influenciar expressivamente para os resultados gerais obtidos por esta pesquisa. Assim sendo, com base nas discussões elaboradas aqui, apresentaremos alguns possíveis apontamentos para pesquisas futuras: i) ampliação do *corpus* de pesquisa; ii) investigação aprofundada sobre os casos de concordância inversa encontrados na comunidade de fala; iii) execução de uma coleta de dados que compreenda, também, a fala de crianças recifenses; iv) efetivação de testes de intuição linguística capazes de isolar os casos de não-opcionalidade para o uso variável dos artigos definidos em sintagmas antroponímicos na cidade do Recife.

## REFERÊNCIAS

- AMARAL, E. T. R.; SEIDE, M. S. **Nomes próprios de pessoa: introdução à antroponímia brasileira**. São Paulo: Blücher, 2020.
- ARAÚJO, I. L. **A natureza do conhecimento após a virada linguístico-pragmática**. Revista de Filosofia, Curitiba, v. 16 n.18, p. 103-137, jan./jun. 2004
- CALLOU, D.; SILVA, G. M. O. O uso do artigo definido em contextos específicos. In: HORA, Demerval da (org.) **Diversidade linguística no Brasil**. João Pessoa: Idéia, 1997.
- CAMPOS JÚNIOR, H. da S. **A variação morfossintática do artigo definido na capital capixaba**. 2011. Dissertação (Mestrado Linguística). Centro de Ciências Humanas e Naturais, Programa de Pós-Graduação em Linguística Universidade Federal do Espírito Santo, Vitória, 2011.
- COELHO, Izete Lehmkuhl. *et al.* **Para conhecer sociolinguística**. São Paulo: Contexto, 2015.
- FREITAG, R. M. K. Sociolinguística no/do Brasil. **Cadernos de Estudos Linguísticos**, v. 58, n. 3, Campinas, SP, p. 445-460, set./dez, 2016.
- \_\_\_\_\_. (org.) **Metodologia de Coleta e Manipulação de Dados em Sociolinguística**. São Paulo: Editora Edgard Blücher Ltda, 2014.
- LABOV, William. **Padrões Sociolinguísticos**. Tradução de Marcos Bagno, São Paulo: Parábola Editorial, 2008.
- LUCCHESI, Dante. **Língua e sociedade partidas: a polarização sociolinguística no Brasil**. São Paulo: Contexto, 2015, 320 p.
- MAGALHÃES, T.M. V. O uso do artigo definido diante de pronome possessivos em textos portugueses do século XVI a XX. In.: SEDRINS, A. P. (org) [et al.], **Por amor à linguística: miscelânea de estudos linguísticos dedicados à Maria Denilda Moura**. Maceió: EDUFAL, 2012.
- MARROQUIM, M. **A língua do Nordeste**. 3ª. ed. Curitiba: HD Livros Editora. 1996.
- MOLLICA, M. C.; BRAGA, M. L. (orgs.). **Introdução à Sociolinguística: o tratamento da variação**. São Paulo: Contexto, 2008.
- MÜLLER, A.; NEGRÃO, E. V. O Uso do artigo definido antes do nome próprio em português: uma análise semântica. In: **XXXVI SEMINÁRIO DO GEL**, 1989, Taubaté, SP.
- NARO, A. Modelos quantitativos e tratamento estatístico. In.: MOLLICA, M. C.; BRAGA, M. L. (orgs.) **Introdução à Sociolinguística: o tratamento da variação**. São Paulo: Contexto, 2015.
- OLIVEIRA, M. A. S. de; SEDRINS, A. P. A variação do artigo definido antepondo nomes próprios de pessoas no sertão do pajeú pernambucano: discutindo a influência de fatores

sociais no português falado na cidade de Tabira, **Rios Eletrônica (Fasete)**, Paulo Afonso, BA, n. 24, p. 75-90, 2020.

OLIVEIRA, M. A. S. de. **Um estudo variacionista sobre o licenciamento do artigo definido diante de antropônimos no português falado da cidade de Tabira – PE**. 2018. Monografia (Licenciatura Plena em Letras) – Universidade Federal Rural de Pernambuco/Unidade Acadêmica de Serra Talhada, Serra Talhada, 2018.

PEREIRA, D. K. F. **A realização do artigo definido no português falado na região do Sertão do Pajeú – PE**. 2017. Dissertação (Mestrado em Letras). Centro de Artes e Comunicação, Departamento de Letras, Programa de Pós-Graduação em Letras, Universidade Federal de Pernambuco, Recife, 2017.

PAIVA, M.C. Transcrição de dados linguísticos. In.: MOLLICA, M. C.; BRAGA, M. L. (orgs.) **Introdução à Sociolinguística: o tratamento da variação**. São Paulo: Contexto, 2015.

PETTER, M. Linguagem, língua e linguística. In: FIORIN, J. L. (org). **Introdução à linguística: objetos teóricos**. São Paulo: Contexto, 2008,

SÁ, E. J. de. **Atlas Linguístico de Pernambuco (ALiPE)**. 2013. Tese (Doutorado em Linguística), Centro de Ciências Humanas, Letras e Artes, Programa de Pós-Graduação em Linguística, Universidade Federal da Paraíba, João Pessoa, 2013.

SANTOS, Renata L. A. **A metodologia da pesquisa em sociolinguística variacionista**. Revista espaço acadêmica. Vol.9, nº 97. 2009

SANTOS, R. L. A.; VITÓRIO, E. G. S. L. A. Uma rodada no Goldvarb X. In: COSTA, J. F. C.; SANTOS, R. L. A.; VITÓRIO, E. G. S. L. A. (orgs.). **Variação e mudança linguística no estado de Alagoas**. Maceió: Edufal, 2011.

SANTOS DE PAULO, A. O trabalho de campo sociolinguístico. Uma rodada no Goldvarb X. In: COSTA, J. F. C.; SANTOS, R. L. A.; VITÓRIO, E. G. S. L. A. (orgs.). **Variação e mudança linguística no estado de Alagoas**. Maceió: Edufal, 2011.

SEDRINS, A. P. Nomes próprios e artigos definidos no português brasileiro, **Revista Letras**, Curitiba, PR, n. 96, p. 239-254, jul./dez. 2017.

SEDRINS, A. P.; PEREIRA, D. SIQUEIRA, A. Variação na realização do artigo definido diante de antropônimos em dados de fala e escrita no sertão de Pernambuco. In: SEDRINS, A. P.; SÁ, E. J. (orgs.). **Aspectos descritivos e sócio-históricos da língua falada em Pernambuco**. Recife: editora UFRPE, 2015.

SEDRINS, A. P. O artigo definido no português brasileiro em contexto de antropônimos e de possessivos pré-nominais. In. MOURA, Maria Denilda; SIBALDO, Marcelo Amorim (Orgs.). **Para a história do português brasileiro**. Maceió: Edufal, 2013.

SILVA, G.M.O. Coleta de dados. In.: MOLLICA, M. C.; BRAGA, M. L. (orgs.) **Introdução à Sociolinguística: o tratamento da variação**. São Paulo: Contexto, 2015.

TARALLO, F. **A pesquisa sociolinguística**. 5. ed. São Paulo: Ática, 2007

## APÊNDICE A – TERMO DE ASSENTIMENTO LIVRE E ESCLARECIDO

UNIVERSIDADE FEDERAL DE PERNAMBUCO  
CENTRO DE ARTES E COMUNICAÇÃO – UFPE  
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS – PPGL  
**TERMO DE ASSENTIMENTO LIVRE E ESCLARECIDO**

(PARA MENORES DE 7 a 18 ANOS)

**OBS:** *Este Termo de Assentimento para o menor de 7 a 18 anos não elimina a necessidade da elaboração de um Termo de Consentimento Livre e Esclarecido que deve ser assinado pelo responsável ou representante legal do menor.*

Convidamos você \_\_\_\_\_, após autorização dos seus pais [ou dos responsáveis legais] para participar como voluntário (a) da pesquisa: Uso variável do artigo definido em sintagmas antroponímicos na capital pernambucana: um estudo Sociolinguístico sobre o português falado em Recife. Esta pesquisa é da responsabilidade da pesquisadora Marília Adrielle Siqueira de Oliveira.

Você será esclarecido (a) sobre qualquer dúvida com o responsável por esta pesquisa. Apenas quando todos os esclarecimentos forem dados e você concorde com a realização do estudo, pedimos que rubrique as folhas e assine ao final deste documento, que está em duas vias. Uma via deste termo lhe será entregue para que seus pais ou responsável possam guarda-la e a outra ficará com o pesquisador responsável.

Você estará livre para decidir participar ou recusar-se. Caso não aceite participar, não haverá nenhum problema, desistir é um direito seu. Para participar deste estudo, um responsável por você deverá autorizar e assinar um Termo de Consentimento, podendo retirar esse consentimento ou interromper a sua participação em qualquer fase da pesquisa, sem nenhum prejuízo.

### INFORMAÇÕES SOBRE A PESQUISA:

**Descrição da pesquisa:** Esta pesquisa tem como objetivo principal apresentar um estudo descritivo e quantitativo sobre a realização do artigo definido diante de antropônimos e possessivos pré-nominais na cidade do Recife, sob a perspectiva teórico-metodológica da Sociolinguística variacionista.

**Esclarecimento do período de participação do voluntário na pesquisa, início, término e número de visitas para a pesquisa:** Espera-se que a pesquisa tenha duração de um (1) dia para cada entrevista com cada um dos informantes (24 informantes por cidade). A gravação da entrevista terá um tempo previsto de 45 minutos por participante. Nesse período, serão coletados todos os materiais necessários para análise dos dados (dados de fala). A sua participação será acompanhada e monitorada pelo seu maior responsável, haja vista que trata-se, apenas, de perguntas e respostas gravadas mediante conversa entre a pesquisadora e o informante.

**Riscos:** Quanto aos riscos e desconfortos, como se tratam de dados de fala que serão transcritos e posteriormente estudados e divulgados, a identidade do participante será preservada, a fim de possibilitar o conforto do informante. Caso você venha a sentir algum desconforto, quanto ao local sugerido para coleta ou em relação às perguntas do questionário, comunique a pesquisadora para que sejam tomadas as devidas providências.

**Benefícios diretos e indiretos** Os benefícios esperados com o resultado desta pesquisa são os seguintes: uma melhor descrição do fenômeno em questão a fim de ajudar na sua compreensão; a criação de um bando de dados proveniente da cidade; será possível a execução de outros projetos de pesquisa que venham a surgir a posteriori e envolvam a mesma linha de pesquisa.

As informações desta pesquisa serão confidenciais e serão divulgadas apenas em eventos ou publicações científicas, não havendo identificação dos voluntários, a não ser entre os responsáveis pelo estudo, sendo assegurado o sigilo sobre a participação do/a voluntário (a). Os dados coletados nesta pesquisa (gravações de áudio) ficarão armazenados em computador pessoal, sob a responsabilidade da pesquisadora Marília Adrielle Siqueira de Oliveira, no endereço acima mencionado, pelo período de mínimo 5 anos.

O (a) senhor (a) não pagará nada e nem receberá nenhum pagamento para ele/ela participar desta pesquisa, pois deve ser de forma voluntária, mas fica também garantida a indenização em casos de danos, comprovadamente decorrentes da participação dele/a na pesquisa, conforme decisão judicial ou extra-judicial. Se houver necessidade, as despesas para a participação serão assumidas pelos pesquisadores (ressarcimento com transporte e alimentação). Em caso de dúvidas relacionadas aos aspectos éticos deste estudo, você poderá consultar o Comitê de Ética em Pesquisa Envolvendo Seres Humanos da UFPE no endereço: (Avenida da Engenharia s/n – Prédio do CCS - 1º Andar, sala 4 - Cidade Universitária, Recife-PE, CEP: 50740-600, Tel.: (81) 2126.8588 – e-mail: cepccs@ufpe.br)

Assinatura do pesquisador (a)

**ASSENTIMENTO DO(DA) MENOR DE IDADE EM PARTICIPAR COMO VOLUNTÁRIO(A)**

Eu, \_\_\_\_\_, portador (a) do documento de Identidade \_\_\_\_\_, abaixo assinado, concordo em participar do estudo “A variação do artigo definido diante de antropônimos e possessivos pré-nominais em duas cidades pernambucanas: Tabira e Recife” como voluntário (a). Fui informado (a) e esclarecido (a) pelo (a) pesquisador (a) sobre a pesquisa, o que vai ser feito, assim como os possíveis riscos e benefícios que podem acontecer com a minha participação. Foi-me garantido que posso desistir de participar a qualquer momento, sem que eu ou meus pais precise pagar nada.

Local e data \_\_\_\_\_

Assinatura do (da) menor : \_\_\_\_\_

Presenciamos a solicitação de assentimento, esclarecimentos sobre a pesquisa e aceite do/a voluntário/a em participar. 02 testemunhas (não ligadas à equipe de pesquisadores):

Nome:	Nome:
Assinatura:	Assinatura:



## APÊNDICE B – TERMOS DE CONSENTIMENTO LIVRE E ESCLARECIDO

UNIVERSIDADE FEDERAL DE PERNAMBUCO  
CENTRO DE ARTES E COMUNICAÇÃO – UFPE  
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS – PPGL  
**TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO**  
(PARA MAIORES DE 18 ANOS OU EMANCIPADOS)

Convidamos o (a) Sr. (a) para participar como voluntário (a) da pesquisa “Uso variável do artigo definido em sintagmas antroponímicos na capital pernambucana: um estudo Sociolinguístico sobre o português falado em Recife” que está sob a responsabilidade da pesquisadora Marília Adrielle Siqueira de Oliveira.

Todas as suas dúvidas podem ser esclarecidas com o responsável por esta pesquisa. Apenas quando todos os esclarecimentos forem dados e você concorde com a realização do estudo, pedimos que rubrique as folhas e assine ao final deste documento, que está em duas vias. Uma via lhe será entregue e a outra ficará com o pesquisador responsável.

Você estará livre para decidir participar ou recusar-se. Caso não aceite participar, não haverá nenhum problema, desistir é um direito seu, bem como será possível retirar o consentimento em qualquer fase da pesquisa, também sem nenhuma penalidade.

### INFORMAÇÕES SOBRE A PESQUISA:

**Descrição da pesquisa:** Esta pesquisa tem como objetivo principal apresentar um estudo descritivo e quantitativo sobre a realização do artigo definido diante de antroponímicos e possessivos pré-nominais na cidade do Recife, sob a perspectiva teórico-metodológica da Sociolinguística variacionista.

**Esclarecimento do período de participação do voluntário na pesquisa, início, término e número de visitas para a pesquisa:** Espera-se que a pesquisa tenha duração de um (1) dia para cada entrevista com cada um dos informantes (24 informantes por cidade). A gravação da entrevista terá um tempo previsto de 45 minutos por participante. Nesse período, serão coletados todos os materiais necessários para análise dos dados (dados de fala). Se preferir, a sua participação poderá ser acompanhada e monitorada por terceiros, haja vista que trata-se, apenas, de perguntas e respostas gravadas mediante conversa entre a pesquisadora e o informante.

**Riscos:** Quanto aos riscos e desconfortos, como se tratam de dados de fala que serão transcritos e posteriormente estudados e divulgados, a identidade do participante será preservada, a fim de possibilitar o conforto do informante. Caso você venha a sentir algum desconforto, quanto ao local sugerido para coleta ou em relação às perguntas do questionário, comunique a pesquisadora para que sejam tomadas as devidas providências.

**Benefícios diretos e indiretos** Os benefícios esperados com o resultado desta pesquisa são os seguintes: uma melhor descrição do fenômeno em questão a fim de ajudar na sua compreensão; a criação de um bando de dados proveniente da cidade; será possível a execução de outros projetos de pesquisa que venham a surgir a posteriori e envolvam a mesma linha de pesquisa.

Todas as informações desta pesquisa serão confidenciais e serão divulgadas apenas em eventos ou publicações científicas, não havendo identificação dos voluntários, a não ser entre os responsáveis pelo estudo, sendo assegurado o sigilo sobre a sua participação. Os dados coletados nesta pesquisa (gravações de áudio) ficarão armazenados em computador pessoal, sob a responsabilidade da pesquisadora Marília Adrielle Siqueira de Oliveira, no endereço acima mencionado, pelo período de mínimo 5 anos.

Nada lhe será pago e nem será cobrado para participar desta pesquisa, pois a aceitação é voluntária, mas fica também garantida a indenização em casos de danos, comprovadamente decorrentes da participação na pesquisa, conforme decisão judicial ou extra-judicial. Se houver necessidade, as despesas para a sua participação serão assumidas pelos pesquisadores (ressarcimento de transporte e alimentação).

Em caso de dúvidas relacionadas aos aspectos éticos deste estudo, você poderá consultar o Comitê de Ética em Pesquisa Envolvendo Seres Humanos da UFPE no endereço: (Avenida da Engenharia s/n – 1º Andar, sala 4 - Cidade Universitária, Recife-PE, CEP: 50740-600, Tel.: (81) 2126.8588 – e-mail: cepccs@ufpe.br).

(assinatura do pesquisador)

### CONSENTIMENTO DA PARTICIPAÇÃO DA PESSOA COMO VOLUNTÁRIO (A)

Eu, \_\_\_\_\_, CPF \_\_\_\_\_, abaixo assinado, após a leitura (ou a escuta da leitura) deste documento e de ter tido a oportunidade de conversar e ter esclarecido as minhas dúvidas com o pesquisador responsável, concordo em participar do estudo “A variação do artigo definido diante de antroponímicos e possessivos pré-nominais em duas cidades pernambucanas: Tabira e Recife”, como voluntário (a). Fui devidamente informado (a) e esclarecido (a) pelo(a) pesquisador (a) sobre a pesquisa, os procedimentos nela envolvidos, assim como os possíveis riscos e benefícios decorrentes de minha participação. Foi-me garantido que posso retirar o meu consentimento a qualquer momento, sem que isto leve a qualquer penalidade.

Local e data \_\_\_\_\_  
Assinatura do participante: \_\_\_\_\_

Impressão digital  
  
(opcional)

**Presenciamos a solicitação de consentimento, esclarecimentos sobre a pesquisa e o aceite do voluntário em participar.** (02 testemunhas não ligadas à equipe de pesquisadores):

Nome:	Nome:
Assinatura:	Assinatura:

UNIVERSIDADE FEDERAL DE PERNAMBUCO  
CENTRO DE ARTES E COMUNICAÇÃO – UFPE  
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS – PPGL  
**TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO**  
**(PARA RESPONSÁVEL LEGAL PELO MENOR DE 18 ANOS)**

Solicitamos a sua autorização para convidar o (a) seu/sua filho (a) \_\_\_\_\_ para participar, como voluntário (a), da pesquisa Uso variável do artigo definido em sintagmas antropônimos na capital pernambucana: um estudo Sociolinguístico sobre o português falado em Recife

Esta pesquisa é da responsabilidade da pesquisadora Marília Adrielle Siqueira de Oliveira (contatos: inserir contatos; endereço: inserir endereço e está sob a orientação do professor Marcelo Amorim Sibaldo.

O/a Senhor/a será esclarecido (a) sobre qualquer dúvida a respeito da participação dele/a na pesquisa. Apenas quando todos os esclarecimentos forem dados e o/a Senhor/a concordar que o (a) menor faça parte do estudo, pedimos que rubrique as folhas e assine ao final deste documento, que está em duas vias.

Uma via deste termo de consentimento lhe será entregue e a outra ficará com o pesquisador responsável. O/a Senhor/a estará livre para decidir que ele/a participe ou não desta pesquisa. Caso não aceite que ele/a participe, não haverá nenhum problema, pois desistir que seu filho/a participe é um direito seu. Caso não concorde, não haverá penalização para ele/a, bem como será possível retirar o consentimento em qualquer fase da pesquisa, também sem nenhuma penalidade.

### INFORMAÇÕES SOBRE A PESQUISA

**Descrição da pesquisa:** Esta pesquisa tem como objetivo principal apresentar um estudo descritivo e quantitativo sobre a realização do artigo definido diante de antropônimos e possessivos pré-nominais na cidades do Recife, sob a perspectiva teórico-metodológica da Sociolinguística variacionista.

**Esclarecimento do período de participação do voluntário na pesquisa, início, término e número de visitas para a pesquisa:** Espera-se que a pesquisa tenha duração de um (1) dia para cada entrevista com cada um dos informantes (24 informantes por cidade). A gravação da entrevista terá um tempo previsto de 45 minutos por participante. Nesse período, serão coletados todos os materiais necessários para análise dos dados (dados de fala). A participação do menor será acompanhada e monitorada pelo maior responsável, haja vista que trata-se, apenas, de perguntas e respostas gravadas mediante conversa entre a pesquisadora e o informante.

**Riscos:** Quanto aos riscos e desconfortos, como se tratam de dados de fala que serão transcritos e posteriormente estudados e divulgados, a identidade do participante será preservada, a fim de possibilitar o conforto do informante. Caso você venha a sentir algum desconforto, quanto ao local sugerido para coleta ou em relação às perguntas do questionário, comunique a pesquisadora para que sejam tomadas as devidas providências.

**Benefícios diretos e indiretos** Os benefícios esperados com o resultado desta pesquisa são os seguintes: uma melhor descrição do fenômeno em questão a fim de ajudar na sua compreensão; a criação de um bando de dados proveniente da cidade; será possível a execução de outros projetos de pesquisa que venham a surgir a posteriori e envolvam a mesma linha de pesquisa.

As informações desta pesquisa serão confidenciais e serão divulgadas apenas em eventos ou publicações científicas, não havendo identificação dos voluntários, a não ser entre os responsáveis pelo estudo, sendo assegurado o sigilo sobre a participação do/a voluntário (a). Os dados coletados nesta pesquisa (gravações de áudio) ficarão armazenados em computador pessoal, sob a responsabilidade da pesquisadora Marília Adrielle Siqueira de Oliveira, no endereço acima mencionado, pelo período de mínimo 5 anos.

O (a) senhor (a) não pagará nada e nem receberá nenhum pagamento para ele/ela participar desta pesquisa, pois deve ser de forma voluntária, mas fica também garantida a indenização em casos de danos, comprovadamente decorrentes da participação dele/a na pesquisa, conforme decisão judicial ou extra-judicial. Se houver necessidade, as despesas para a participação serão assumidas pelos pesquisadores (ressarcimento com transporte e alimentação).

Em caso de dúvidas relacionadas aos aspectos éticos deste estudo, você poderá consultar o Comitê de Ética em Pesquisa Envolvendo Seres Humanos da UFPE no endereço: (Avenida da Engenharia s/n – Prédio do CCS - 1º Andar, sala 4 - Cidade Universitária, Recife-PE, CEP: 50740-600, Tel.: (81) 2126.8588 – e-mail: cepccs@ufpe.br)

Assinatura do pesquisador (a)

### CONSENTIMENTO DO RESPONSÁVEL PARA A PARTICIPAÇÃO DO/A VOLUNTÁRIO

Eu, \_\_\_\_\_, CPF \_\_\_\_\_, abaixo assinado, responsável por \_\_\_\_\_, autorizo a sua participação no estudo “A variação do artigo definido diante de antropônimos e possessivos pré-nominais em duas cidades pernambucanas: Tabira e Recife”, como voluntário(a). Fui devidamente informado (a) e esclarecido (a) pelo (a) pesquisador (a) sobre a pesquisa, os

procedimentos nela envolvidos, assim como os possíveis riscos e benefícios decorrentes da participação dele (a). Foi-me garantido que posso retirar o meu consentimento a qualquer momento, sem que isto leve a qualquer penalidade para mim ou para o (a) menor em questão.

Local e data \_\_\_\_\_

Assinatura do (da) responsável: \_\_\_\_\_

**Presenciamos a solicitação de consentimento, esclarecimentos sobre a pesquisa e aceite do voluntário em participar.** 02 testemunhas (não ligadas à equipe de pesquisadores):

Nome:	Nome:
Assinatura:	Assinatura:

## APÊNDICE C – TERMO DE COMPROMISSO E CONFIDENCIALIDADE

**Título do projeto** Uso variável do artigo definido em sintagmas antroponímicos na capital pernambucana: um estudo Sociolinguístico sobre o português falado em Recife

**Pesquisador responsável:** Marília Adrielle Siqueira de Oliveira

**Instituição/Departamento de origem do pesquisador:** Centro de Artes de Comunicação – Programa de Pós-Graduação em Letras

**Telefone para contato:**

**E-mail:**

O pesquisador do projeto acima identificado assume o compromisso de:

- Garantir que a pesquisa só será iniciada após a avaliação e aprovação do Comitê de Ética e Pesquisa Envolvendo Seres Humanos da Universidade Federal de Pernambuco – CEP/UFPE e que os dados coletados serão armazenados pelo período mínimo de 5 anos após o término da pesquisa;
- Preservar o sigilo e a privacidade dos voluntários cujos dados serão estudados e divulgados apenas em eventos ou publicações científicas, de forma anônima, não sendo usadas iniciais ou quaisquer outras indicações que possam identificá-los;
- Garantir o sigilo relativo às propriedades intelectuais e patentes industriais, além do devido respeito à dignidade humana;
- Garantir que os benefícios resultantes do projeto retornem aos participantes da pesquisa, seja em termos de retorno social, acesso aos procedimentos, produtos ou agentes da pesquisa;
- Assegurar que os resultados da pesquisa serão anexados na Plataforma Brasil, sob a forma de Relatório Final da pesquisa;

Recife, ..... de ..... de 20..... .

---

**Assinatura Pesquisador Responsável**

## APÊNDICE D – FICHA SOCIAL

### UNIVERSIDADE FEDERAL DE PERNAMBUCO CENTRO DE ARTES E COMUNICAÇÃO – UFPE PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS – PPGL FICHA SOCIAL

Uso variável do artigo definido em sintagmas antroponímicos na capital pernambucana: um estudo Sociolinguístico sobre o português falado em Recife

**Pesquisadora:** Marília Adrielle Siqueira de Oliveira

**Orientador:** Marcelo Amorim Sibaldo

<b>FICHA DO INFORMANTE</b>		
<b>DADOS PESSOAIS DO INFORMANTE</b>		
NOME:		APELIDO:
DATA DE NASCIMENTO:	IDADE:	SEXO:
ENDEREÇO:		
ESTADO CIVIL:	NÚMERO DE FILHOS: HOMENS: MULHERES:	
NATURALIDADE:	COM QUE IDADE CHEGOU A ESTA COMUNIDADE? (CASO NÃO SEJA NATURAL DA LOCALIDADE)	
DOMICÍLIOS E TEMPO DE PERMANÊNCIA FORA DA LOCALIDADE:		
NÍVEL DE ESCOLARIDADE:		ANOS DE ESCOLARIZAÇÃO:
<b>DADOS PROFISSIONAIS DO INFORMANTE</b>		
QUAL PROFISSÃO EXERCE?		
HÁ QUANTO TEMPO TRABALHA NESTA PROFISSÃO?	FEZ CURSOS ESPECÍFICOS PARA A PROFISSÃO?	
<b>REDES DE RELACIONAMENTO</b>		
( ) LOCAL ( Mesma comunidade) ( ) DISPERSA( Outras comunidade)	SITES DE RELACIONAMENTO: ( ) Facebook ( ) E-mail ( ) WhatsApp ( ) Outros	
<b>CONTATO COM MEIOS DE COMUNICAÇÃO</b>		
ASSISTE TV?	PROGRAMAS FAVORITOS: ( ) Novelas ( ) Prog. Religioso ( ) Esportes ( ) Filmes ( ) Noticiário ( ) Prog. Auditório ( ) Outros	

OUVE RÁDIO?	PROGRAMAS PREFERIDOS: ( )Noticiário geral ( ) Esportes ( ) Prog. Religioso ( ) Música ( )Noticiário policial ( ) Prog. c/ partic. do ouvinte ( ) )Outros	
LÊ JORNAL?	SEÇÕES QUE GOSTA DE LER NO JORNAL:	
LÊ REVISTA?		
QUE RELIGIÃO PRÁTICA?		
PARTICIPAÇÃO EM DIVERSÕES: ( ) CINEMA ( ) TEATRO ( ) SHOWS ( ) MAN. FOLCLÓRICAS ( ) FUTEBOL ( ) OUTROS ESPORTES ( ) OUTROS _____		
<b>PARA PREENCHIMENTO APÓS A ENTREVISTA</b>		
ESPONTANEIDADE DA ELOCUÇÃO:		
GRAU DE CONHECIMENTO ENTRE INFORMANTE E INQUIRIDOR:		
INTERFERÊNCIA OCASIONAL DE CIRCUNSTANTES:		
CARACTERIZAÇÃO SUMÁRIA DO(S) CIRCUNSTANTE(S):		
AMBIENTE DO INQUÉRITO:		
OBSERVAÇÕES:		
INQUIRIDOR:	CIDADE:	DATA:
DURAÇÃO:		

## **APÊNDICE E – ROTEIRO DE PERGUNTAS**

### **ROTEIRO DE PERGUNTAS 1**

- 1- Você acha que essa é uma boa cidade para se morar, por quê?
- 2- No que você trabalha? Você gosta do que faz?
- 3- Qual sua relação com seus colegas de trabalho? Existe alguém com quem você se dá melhor. Por quê?
- 4- O que gosta de fazer durante o tempo livre?
- 5- Você tem uma família grande?
- 6- Você mora com familiares? Quais?
- 7- Como é a sua relação com as pessoas de sua família?
- 8- Você tem parente que moram fora? quais?
- 9 – Lembrando da sua infância, você poderia relatar atividades que realizava junto com membros da sua família e que são boas lembranças?
- 10– Numa família geralmente há uma tendência em cada um ter uma característica marcante (um é mais temperamental, outro é mais calmo, outro é mais ativo, mais preguiçoso, mais fofoqueiro). Pensando em seus irmãos, primos, tios, como você vê aqueles com características mais marcantes?
- 11– No seu tempo de escola, como era seu relacionamento com os colegas de classe? Como você poderia se identificar como aluno (o mais inteligente, o mais comportado). E os seus colegas de escola mais marcantes, quais foram e por quê?
- 12- Você teve bons professores durante a escola? Fale sobre alguns deles
- 13- Você viaja com frequência? Quando viaja, costuma ir com alguém?
- 14 - Você assiste? (novelas, programas de televisão, filmes ou séries)
- 15- Quais os seus preferidos? Por quê?
- 16- Quais os atores/personalidades, seja do cinema, da tv, da política que você admira e por quê? E aqueles que você desaprova o comportamento.
- 17- Você gosta de ouvir música? Quais cantores \ estilo você ouve?
- 18 – Como você avalia a atual situação política brasileira?
- 19- Você poderia falar alguma coisa sobre algum político em que você votaria e algum em que você não votaria?
- 20 – Você já passou por alguma situação de perigo junto com alguém?



**APÊNDICE F – CARTA DE ANUÊNCIA**

**UNIVERSIDADE FEDERAL DE PERNAMBUCO  
CENTRO DE ARTES E COMUNUCAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS**

**REFERENTE À CARTA DE ANUÊNCIA**

Eu, Marília Adrielle Siqueira de Oliveira, portadora do RG: e CPF: , venho, através deste, solicitar dispensa da apresentação do documento “carta de Anuência” na Plataforma Brasil, haja vista que, para a realização da minha pesquisa de mestrado, não serão coletados dados em instituições de ensino ou órgãos públicos. A pesquisa intitulada “Uso variável do artigo definido em sintagmas antroponímicos na capital pernambucana: um estudo Sociolinguístico sobre o português falado em Recife se propõe a realizar coletas com indivíduos escolhidos aleatoriamente nos municípios pernambucanos selecionados para análise. Assim sendo, não terá um local específico para a coleta e, por esse motivo, não será possível a apresentação do referido documento.

Recife, 28 de maio de 2019