

**UNIVERSIDADE DO ESTADO DE MATO GROSSO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO *STRICTO SENSU* EM LINGUÍSTICA
MESTRADO EM LINGUÍSTICA**

KAMYLLA MARTOS VITTORAZZI RIOS

**ESTUDO DO *SUJEITO* PRONOMINAL DE REFERÊNCIA ARBITRÁRIA NA FALA
CARIOCA**

**CÁCERES – MT
2022**

KAMYLLA MARTOS VITTORAZZI RIOS

**ESTUDO DO *SUJEITO* PRONOMINAL DE REFERÊNCIA ARBITRÁRIA NA FALA
CARIOCA**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Linguística da Universidade do Estado de Mato Grosso, como requisito parcial para a obtenção do título de Mestre em Linguística.

Orientadora: Dra. Gislaine Aparecida de Carvalho.

R586e	RIOS, Kamylla Martos Vittorazzi. Estudo do Sujeito Pronominal de Referência Arbitrária na Fala Carioca / Kamylla Martos Vittorazzi Rios - Cáceres, 2022. 78 f.; 30 cm. Trabalho de Conclusão de Curso (Dissertação/Mestrado) - Curso de Pós-graduação Stricto Sensu (Mestrado Acadêmico) Linguística, Faculdade de Educação e Linguagem, Câmpus de Cáceres, Universidade do Estado de Mato Grosso, 2022. Orientador: Gislaine Aparecida de Carvalho 1. Sujeito Nulo. 2. Sujeito Preenchido. 3. Referência Genérica. 4. Ineterminação do Sujeito. I. Kamylla Martos Vittorazzi Rios. II. Estudo do Sujeito Pronominal de Referência Arbitrária na Fala Carioca: . CDU 81'27
-------	--

KAMYLLA MARTOS VITTORAZZI RIOS

**ESTUDO DO SUJEITO PRONOMINAL DE REFERÊNCIA ARBITRÁRIA NA FALA
CARIOCA**

BANCA EXAMINADORA

Prof.^a Dr.^a Gislaine Aparecida de Carvalho
PPGL - UNEMAT
(Orientadora)

Prof. Dr. Albano Dalla Pria
PPGL – UNEMAT
(Avaliador Interno)

Prof.^a Dr.^a Juliana Bertucci Barbosa
UFTM
(Avaliador Externo)

**CÁRECES – MT
2022**

Dedico este trabalho a minha família pelo apoio incondicional, em especial, aos meus filhos Lucas e Arthur, meus maiores incentivos na realização deste trabalho.

AGRADECIMENTOS

Agradeço

À Deus, primeiramente, por me privilegiar com a vida e com a realização deste trabalho.

À **Prof.^a Dr.^a Gislaine Aparecida de Carvalho**, por confiar em mim, pela orientação valiosa, pela paciência, disposição e carinho ao longo do curso.

Aos membros da banca, **Prof. Dr. Albano Dalla Pria** e **Prof.^a Dr.^a Juliana Bertucci Barbosa**, por aceitarem o convite de participar da banca e, sobretudo, pelas valiosas contribuições com a minha pesquisa e à **Prof.^a Dr.^a Dircel Aparecida Kailer**, pelas sugestões durante a qualificação.

Aos meus amigos, **Welliton Bindandi**, **Wellington Oliveira** e **Thainá Aparecida**, pelo incentivo a continuar estudando, pelo apoio imensurável que me concedem, pela injeção de ânimo e encorajamento, pelas conversas, risadas, que sempre me alegram.

Por fim, agradeço imensamente a minha família, que sempre me apoiou em minhas decisões e durante todo o percurso no Curso de Linguística. Agradeço, principalmente, a minha mãe, **Maria**, e meu pai, **Avadir**, que se dispuseram a cuidar dos meus filhos, enquanto estive ausente para me dedicar à pesquisa. Pelo apoio emocional que recebo deles, sem os quais não conseguiria avançar em meus estudos.

Ao meu esposo, **Cleiton**, pelo companheirismo, pela parceria, pela paciência, pelo apoio e torcida. A sua presença me dava forças para continuar.

Ao Programa de Pós-Graduação em Linguística (PPGL), em especial, aos professores, pelo conhecimento compartilhado. À coordenação do curso, por toda atenção dada, durante a qualificação.

À Universidade do Estado de Mato Grosso (UNEMAT), pela formação e qualificação de profissionais com excelência, para trabalhar na área da educação e pesquisa no Estado e no país.

A todos que direta e indiretamente contribuíram com a minha qualificação, neste percurso tão desafiador da pesquisa.

[...] a etiqueta linguística é exatamente o registro da diversidade da linguagem de um povo.

(Jean Louis Calvet).

RESUMO

Esta pesquisa fundamenta-se na Teoria de Variação e Mudança Linguística, nos estudos de Weinreich, Labov e Herzog (2006 [1968]). O principal objetivo do estudo consiste em fazer um levantamento do preenchimento pronominal em concorrência com o *sujeito nulo*, na fala dos cariocas, enfocando a atuação dos pronomes de referência arbitrária e sua correlação com a alteração do Parâmetro do *Sujeito Nulo*. Para analisar a realização do *sujeito* de referência arbitrária na fala carioca, a pesquisa baseou, em parte, do *corpus* utilizado por Duarte e Reis (2018), do **Projeto Concordância**, do site CORPORAPORT, da UFRJ, que disponibiliza materiais diversos para análises linguísticas de pesquisadores. Portanto, o *corpus* é constituído de entrevistas gravadas no período entre 2009 e 2010, que englobam falantes de Copacabana e Nova Iguaçu, Rio de Janeiro (RJ), estratificados em faixa etária, escolaridade e gênero. Deste *corpus*, foram selecionadas somente as entrevistas dos falantes de Copacabana. Os dados obtidos passaram por um tratamento estatístico com o auxílio do programa GoldVarb 2001, que permitiu uma análise tanto quantitativa quanto qualitativa dos resultados, de modo a evidenciar: quais fatores condicionam o preenchimento do *sujeito*, bem como quais contextos ainda privilegiam a gramática [+*Sujeito nulo*]. Os resultados mostraram a preferência dos cariocas pelo *sujeito*-pronome no uso do *sujeito* indeterminado, apresentando evidências que indicam um processo de mudança em direção ao *sujeito* preenchido.

Palavras-chave: *Sujeito nulo. Sujeito preenchido. Referência genérica. Indeterminação do sujeito.*

ABSTRACT

This research is based on the Theory of Linguistic Variation and Change, in the studies of Weinreich, Labov and Herzog (2006 [1968]). The main object is to make a survey of the pronominal filling in competition with the null subject, in the speech of cariocas, focusing on the performance of arbitrary reference pronouns and their correlation with the alteration of the Null Subject Parameter. To analyze the performance of the subject of arbitrary reference in Carioca speech, the research was based, in part, on the corpus used by Duarte and Reis (2018), from the Concordance Project, from the CORPORAPORT website, from UFRJ, which provides various materials for linguistic analyses of researchers. Therefore, the corpus consists of interviews recorded between 2009 and 2010, which include speakers from Copacabana and Nova Iguaçu, Rio de Janeiro (RJ), stratified in age, schooling and gender. From this corpus, only the interviews of the speakers of Copacabana were selected. The data obtained underwent a statistical treatment with the help of the GoldVarb 2001 program, which allowed a quantitative and qualitative analysis of the results, in order to evidence: which factors condition the completion of the subject, as well as which contexts still favor grammar [+Null Subject]. The results showed the preference of cariocas for the subject-pronoun in the use of the indeterminate subject, presenting evidence that indicates a process of change towards the completed subject.

Keywords: Null subject; Subject completed; Generic reference; Indeterminacy of the subject.

LISTA DE QUADROS

Quadro-síntese 1	35
Quadro-síntese 2	35
Quadro-síntese 3	71

LISTA DE TABELAS

Tabela 1- Quantificação e porcentagem dos sujeitos de referência arbitrária na fala carioca.	55
Tabela 2 - Quantificação e porcentagem do cruzamento dos fatores tipo de sujeito e correferencialidade	57
Tabela 3- Quantificação e porcentagem do cruzamento entre os fatores tipo de sujeito x pessoa gramatical	60
Tabela 4 - Quantificação e porcentagem dos fatores tipo de sujeito x escolaridade	62
Tabela 5 - Quantificação e porcentagem do cruzamento dos fatores tipo de sujeito x faixa etária	63
Tabela 6 - Quantificação e porcentagem do cruzamento dos fatores tipo de sujeito x tempo e modo verbal	65
Tabela 7- Quantificação e porcentagem do sujeito de referência arbitrária fator mais animado e menos animado	66
Tabela 8 - Quantificação e porcentagem do cruzamento do fator tipo de sujeito x concordância verbal	68
Tabela 9 - Quantificação e porcentagem do cruzamento dos fatores tipo de sujeito x gênero	68
Tabela 10 - Quantificação e porcentagem do cruzamento dos fatores tipo de sujeito x pessoa gramatical x gênero	70

SUMÁRIO

INTRODUÇÃO	12
SEÇÃO I.....	15
O SUJEITO DE REFERÊNCIA INDETERMINADA	15
1.1 A indeterminação do sujeito para a gramática tradicional	15
1.2 A indeterminação do sujeito e os estudos linguísticos.....	17
1.2.1 Estudos diacrônicos e sincrônicos: a <i>indeterminação do sujeito</i>	17
1.3 A partícula <i>se</i>	20
1.3.1 <i>Se</i> apassivador/indeeterminador	20
1.3.2 A realização do <i>sujeito</i> em sentenças finitas e infinitivas	23
1.4 Finalizando	27
A INDETERMINAÇÃO DO SUJEITO NAS VARIEDADES DO PORTUGUÊS BRASILEIRO.....	29
2.1 O sujeito arbitrário no português brasileiro.....	29
2.2 Finalizando	36
SEÇÃO III	38
ABORDAGEM TEÓRICO-METODOLÓGICA	38
3.1 A Teoria Variacionista	38
3.2 O Parâmetro do Sujeito Nulo	40
3.3 Sobre a pesquisa	41
3.3.1 Objetivo da pesquisa	41
3.3.2 Hipótese.....	41
3.3.3 O <i>corpus</i> da pesquisa.....	42
3.4 Sobre o método.....	42
3.4.1 Suporte quantitativo	42
3.4.2 Grupo de fatores	43
3.5 Fatores linguísticos.....	44
3.5.1 Tipo de ocorrência: <i>sujeito nulo</i> vs <i>sujeito pronominal</i>	44

3.5.2 A pessoa gramatical	45
3.5.3 O fator animacidade	47
3.5.4 A concordância verbal.....	48
3.5.5 Tempo e modo verbal.....	49
3.5.6 A correferencialidade	50
3.6 Fatores sociais	51
3.6.1 Gênero	51
3.6.2 A faixa etária	52
3.6.3 A escolaridade	52
3.7 Finalizando	53
SEÇÃO IV	54
ANÁLISE DO CORPUS	54
4.1 O sujeito de referência arbitrária na fala carioca.....	54
4.1.1 Resultado geral.....	54
4.1.2 Fatores selecionados pelo programa.....	56
4.1.2.1 Cruzamentos e pesos relativos	56
4.1.2.1.1 Sujeito genérico e a correferencialidade	56
4.1.2.1.2 Tipo de sujeito x pessoa gramatical	59
4.1.2.1.3 Tipo de sujeito x escolaridade	62
4.1.2.1.4 Tipo de sujeito x faixa etária	63
4.1.2.1.5 Tipo de sujeito x tempo e modo verbal	64
4.2 Fatores excluídos pelo programa GoldVarb 2001 (peso relativo)	66
4.2.1 A animacidade do sujeito	66
4.2.2 Tipo de sujeito x concordância verbal.....	67
4.2.3 Tipo de sujeito x gênero	68
4.2.4 Tipo de sujeito x pessoa gramatical x gênero.....	69
4.4 Finalizando	70
CONSIDERAÇÕES FINAIS	72

REFERÊNCIAS 74

INTRODUÇÃO

Estudos sobre o Português Brasileiro (PB) têm aumentado consideravelmente, em razão de apresentar peculiaridades que se distinguem do Português Europeu (PE) e das demais línguas românicas. (GALVES (2001); DUARTE (1995; 2003). Um fato interessante diz respeito a algumas mudanças que ocorreram na língua ao longo do tempo e teriam afetado a marcação do parâmetro do *sujeito nulo* no português brasileiro.

Línguas positivamente marcadas pelo Parâmetro de *Sujeito Nulo* apresentam preferencialmente a posição de *sujeito vazia*, tanto para os *sujeitos* definidos quanto arbitrários. Entretanto, no português brasileiro, esse comportamento ocorre de modo diferente, pois há preferência pelo preenchimento da posição do *sujeito* com formas pronominais, como apontam Duarte (1995, 2000), Cyrino (1994), Cyrino, Duarte e Kato (2000), dentre outros.

Tarallo (1983) foi um dos primeiros pesquisadores a evidenciar algumas mudanças sintáticas relevantes que acabaram interferindo no fenômeno de preenchimento do *sujeito*. Segundo o referido autor (1993), “[...] a reorganização do sistema pronominal teve como consequências mais importantes a implementação de objetos nulos no sistema brasileiro de um lado, e *sujeito lexicais* mais frequentes de outro”.

Diversas análises diacrônicas e sincrônicas, conforme Duarte (1995, 2000), Cyrino (1994), Laperuta (2002), Carvalho (2005), confirmam que o português brasileiro já não é uma língua tipicamente [+] *pro-drop*, ou seja, apresenta menos *sujeito nulos*, ao contrário do português europeu que, embora tenha casos de preenchimento do *sujeito pronominal*, ainda é marcado positivamente. (DUARTE, 199).

Duarte (1993) associa a crescente preferência do *sujeito pronominal* à simplificação no paradigma flexional. Para chegar a essa conclusão, foram investigados trechos de peças de teatro, de cunho eminentemente popular, escritas nos séculos XIX e XX, que compreendem o período entre 1845 e 1992. Os resultados da pesquisa mostraram que a curva ascendente no número de *sujeito* preenchidos é perpendicular ao enfraquecimento da morfologia verbal. Em suas palavras: “O português brasileiro convive com um sistema agonizante, em que ainda se refletem as características *pro-drop*, e um sistema em desenvolvimento, em que a riqueza funcional perdida já não permite a identificação de *pro*”. (DUARTE, 1993).

Se a perda da propriedade [+] *pro-drop* do português brasileiro incide, inicialmente, em *sujeito* de referência definida, posteriormente, a preferência pelo preenchimento também

recai sobre o *sujeito* de referência arbitrária. No português brasileiro, as construções pronominais nominativas superaram o uso da categoria vazia como estratégia para a indeterminação do *sujeito* (DUARTE, 2003). Nesse sentido, esses *sujeitos de referência genérica* (indeterminados) são objeto de nossa investigação.

Trazemos, a seguir, dois exemplos ilustrativos:

- a) Quando **eu** estudava em Copacabana... era muito bom... inclusive na época que eu estudava... eh as escolas públicas eram muito boas né... as pessoas que tinham dinheiro colocavam seus filhos em escola pública. (E03-03) – **Sujeito definido**
- b) [...] você tem que buscar o que que há de mais ser humano, dar valores de ser humano pras pessoas, porque realmente elas vão pensar... “se **eu** grito mais do que **eu** devia na rua **eu** tou incomodando outra pessoa que... que recebe um som a mais de carros, de poluição etc e tal... falta isso eu acho... essa criação de valor. (E01-02) – **Sujeito arbitrário**

Neste trabalho, objetivamos fazer um levantamento do preenchimento do *sujeito pronominal* em concorrência com o *sujeito nulo* na fala dos cariocas, enfocando a atuação dos pronomes de referência arbitrária e sua correlação com a alteração do Parâmetro de *Sujeito Nulo*, a fim de verificar o comportamento desse fenômeno nessa variedade do português brasileiro. Além disso, buscamos averiguar quais os fatores condicionam o uso das variáveis *sujeito pronominal* e *sujeito nulo*, se são fatores internos e/ou externos à língua.

Para responder nossa questão de pesquisa levamos em consideração as variáveis linguísticas *animacidade do sujeito*, *pessoa gramatical*, *concordância verbal*, *correferencialidade*, *tempo e modo verbal* e as variáveis extralinguísticas *faixa etária*, *escolaridade* e *gênero*.

Das constatações feitas em estudos anteriores, apontadas por Duarte (1995, 2003) e Carvalho (2005), decorre nossa hipótese de pesquisa: o uso das formas pronominais nominativas será superior à categoria vazia na estratégia de indeterminação do *sujeito*. Com base nos pressupostos teórico-metodológicos da Teoria da Variação e Mudança Linguística, pautados em autores como Weinreich, Labov e Herzog (2006 [1968]), buscamos confirmar ou refutar a hipótese em dados que provêm do *corpus*, que pertence ao **Projeto Concordância**, disponível no site CORPORAPOORT, que registra entrevistas gravadas com falantes cariocas, no período que consiste entre 2009 e 2010.

Cabe ainda dizer que esse *corpus*, que pertence ao **Projeto Concordância**, também foi objeto de análise de Duarte e Reis (2018). Nele, os pesquisadores investigam a realização do *sujeito* de referência definida em entrevistas que englobam falantes de duas localidades de Lisboa (Oeiras e Cáceres) e do Rio de Janeiro (Copacabana e Nova Iguaçu). Conforme mencionado, usamos o mesmo *corpus*, só que nele, interessou-nos o *sujeito de referência arbitrária* e apenas a fala carioca de moradores de Copacabana.

Este trabalho é de grande relevância para os estudos sobre a realização do *sujeito* no português brasileiro, uma vez que amplia a descrição feita por Duarte e Reis (2018), que investigaram os *sujeitos de referência definida*, ao passo que, no nosso estudo, analisamos a *referência arbitrária* na fala carioca.

A trajetória deste estudo está estruturada em quatro seções:

Na primeira seção, *O sujeito de referência indeterminada*, apresentamos o conceito e a classificação dos recursos de indeterminação do *sujeito*, apregoada pelas Gramáticas Tradicionais (GT), bem como alguns estudos empíricos sobre esse fenômeno linguístico.

Na segunda seção, *A indeterminação do sujeito nas variedades do português brasileiro*, expomos os resultados de algumas pesquisas empíricas sobre a realização do *sujeito* de referência arbitrária tanto no português europeu, quanto no português brasileiro. Os resultados dessas pesquisas são comparados com os nossos resultados no capítulo de análise.

Na terceira seção, *Abordagem teórico-metodológica*, discorremos sobre nossas hipóteses, objetivos, bem como sobre todos os aspectos referentes ao *corpus*, tais como os fatores linguísticos e extralinguísticos selecionados para a coleta de dados, além da metodologia e pressupostos teóricos adotados no processo de pesquisa.

Na última seção, *Análise do corpus*, apresentamos os resultados obtidos através do programa GoldVarb 2001, tabelas com quantificação e porcentagens dos dados, pesos relativos indicando a significância dos fatores na influência das variáveis linguísticas e extralinguísticas, bem como realizaremos a análise desses dados, comparando-os com os resultados apresentados nas pesquisas comentadas no segundo capítulo deste trabalho.

Apresentamos, posteriormente, as considerações finais e as referências bibliográficas.

SEÇÃO I

O SUJEITO DE REFERÊNCIA INDETERMINADA

Considerando que o tema central deste trabalho é a realização do *sujeito de referência arbitrária* (indeterminada) no português brasileiro, especificamente, na fala carioca, faz-se necessário compreender como tal conceito é classificado pelos compêndios gramaticais e descrito pelos linguistas. Neste sentido, trazemos, nesta seção, uma abordagem do ponto de vista de gramáticos tradicionais, bem como a perspectiva de alguns linguistas em relação ao fenômeno: a primeira corresponde a abordagem prescritiva/normativa; a segunda apresenta uma visão na perspectiva funcional da língua.

1.1 A indeterminação do sujeito para a gramática tradicional

Embora saibamos do vasto volume de gramáticas normativas do português brasileiro, utilizamos apenas algumas delas, com a finalidade de expor a definição e caracterização dos recursos de indeterminação do *sujeito* concebidos pelos compêndios gramaticais. Assim, examinamos as gramáticas dos seguintes autores: Almeida (1999), Bechara (2003), Cunha e Cintra (2007), Cegalla (2008) e Rocha Lima (2011).

Os conceitos sobre o fenômeno da indeterminação do *sujeito* difundidos pelas Gramáticas Tradicionais são apresentados de modo sucinto, sem muitas especificações. Os recursos expostos pela maioria dos gramáticos são reduzidos, geralmente representados por duas formas: o emprego do verbo na 3ª pessoa do plural sem referência específica e o uso do verbo na 3ª pessoa do singular acompanhado da partícula *se*. Desse modo, veremos, a seguir, a classificação da indeterminação do ponto de vista dos gramáticos dos compêndios mencionados.

Para Almeida (1999), considera-se *sujeito indeterminado* quando for impossível a sua identificação. O que pode ocorrer com verbos que: a) exerçam atividade e, que, acidentalmente indiquem impessoalização, como no exemplo: “Dizem que ele vem” (p. 414); b) acidentalmente impessoalizados na passiva, como em “Precisa-se de um datilógrafo” (p. 414).

Bechara (2003), diferentemente de Almeida (1999), define *sujeito indeterminado* como aquele que não se nomeia, ou por não querer, ou por não saber fazê-lo. Nesse caso, observamos, na classificação desse fenômeno, a intenção do falante, posto que a não

identificação do *sujeito* pode ser proposital, conforme a intenção daquele que fala, e não como um processo acidental, como ressalta Almeida. Entretanto, os recursos de indeterminação são os mesmos reconhecidos pelo autor, ou seja, o uso de verbos na 3ª pessoa do plural com *sujeito nulo*, bem como verbos na 3ª pessoa do singular, acompanhados da partícula *se*.

Do ponto de vista de Cunha e Cintra (2007), a *indeterminação do sujeito* ocorre quando o verbo, em alguns casos, não se refere a uma pessoa determinada, “[...] ou por se desconhecer quem pratica a ação, ou por não haver interesse em conhecê-la”. Os autores asseguram que para efeito de indeterminação vale-se a língua de duas estratégias, são elas: uso de verbo na 3ª pessoa do plural, e construções com a partícula *se*. Tais recursos podem ser evidenciados nos exemplos: “Reputavam-no o maior comilão da cidade” e “Comia-se com a boca, com os olhos, com o nariz”. (CUNHA e CINTRA, 2007).

Cegalla (2008) apresenta um pensamento semelhante a Cunha e Cintra (2007) ao dizer que o *sujeito é indeterminado* quando sua identidade é desconhecida ou escondida. Baseia-se nos mesmos recursos que os demais autores, porém, acrescenta o infinitivo como estratégia de indeterminação, assim como nos exemplos: “Era penoso carregar aqueles fardos enormes”. “É triste assistir a estas cenas repulsivas”. (p. 275).

Rocha Lima (2011) afirma que o *sujeito indeterminado* é aquele que não podemos ou não queremos identificar. O autor apresenta como recursos de indeterminação o emprego do verbo na 3ª pessoa do plural, sem referência anterior a *eles*, e o substantivo no plural; o uso do verbo na 3ª pessoa do singular acompanhado da partícula *se*, tais como: “Falam mal daquela moça”, “Vive-se bem aqui”. (p. 289).

De modo geral, observamos que os gramáticos tradicionais utilizam critérios semânticos para a definição do fenômeno de indeterminação, caracterizando-o como aquele que se desconhece ou não se pode identificar, porém, na classificação dos recursos, baseiam-se em questões sintáticas, pois nas duas principais estratégias (emprego do verbo na 3ª pessoa do plural e, 3ª pessoa do singular acompanhado da partícula *se*) evidencia-se a predicação verbal como fator responsável pela classificação do *sujeito indeterminado*.

Com base no exposto, podemos dizer que a abordagem dos autores supracitados é suficiente para compreender como a *indeterminação do sujeito* é concebida pelos compêndios gramaticais, haja vista que os gramáticos reproduzem os conceitos de seus antecessores, mostrando-se fiéis a tradição normativa. Além disso, essa visão contrasta com o ponto de vista de alguns linguistas que buscam compreender o funcionamento da língua em situações interativas, ampliando o estudo de fenômenos linguísticos como a *indeterminação* que, geralmente, são pouco especificados pelas gramáticas tradicionais.

1.2 A indeterminação do sujeito e os estudos linguísticos

Se, por um lado, a tradição gramatical apregoou o conceito de *sujeito indeterminado* pelo uso do verbo na 3ª pessoa do plural, desacompanhado da partícula *se*, e com verbo intransitivo ou transitivo indireto, na 3ª pessoa do singular acompanhado de *se* (ambas as formas com a posição do *sujeito vazia*), por outro lado, alguns estudos linguísticos, que buscam descrever a língua portuguesa brasileira, reconhecem novas formas de *indeterminação* que vão além dessas apresentadas pelas GTs.

Para comentar sobre os recursos de *indeterminação do sujeito*, na perspectiva linguística, mencionaremos algumas pesquisas que investigaram esse fenômeno, tanto na língua portuguesa brasileira (doravante PB), quanto na língua portuguesa europeia (doravante PE).

1.2.1 Estudos diacrônicos e sincrônicos: a *indeterminação do sujeito*

O português brasileiro, embora seja caracterizado como língua de *sujeitos nulos*, já não atende, tipicamente, a essa propriedade *pro-drop*, confirmam essa afirmação Duarte (1995, 2000), Cyrino (1994), Laperuta (2002), Carvalho (2005). Essa variedade do português apresenta um comportamento peculiar em relação ao *Parâmetro do Sujeito Nulo*, com tendência ao preenchimento da posição do *sujeito*, tanto de referência definida quanto arbitrária.

Tendo em vista a realização do *sujeito de referência arbitrária*, também conhecida como *referência indeterminada*, Duarte, Kato e Barbosa (2000), em um estudo comparativo sobre o PE e PB, analisaram a modalidade escrita de ambas as variedades do português, com o propósito de investigar os recursos de indeterminação utilizados por elas. O *corpus* dessa pesquisa foi constituído de diferentes gêneros textuais (entrevistas, receitas de cozinha e cartas de leitores), transcritos de jornais e revistas portuguesas e brasileiros. Para a análise das estratégias de indeterminação, foram excluídos os verbos na 3ª pessoa do plural, devido sua baixa ocorrência, bem como pela especificidade de seu uso.

Em relação a análise dos dados, Duarte, Kato e Barbosa (*ibidem*) observaram que em sentenças finitas, as construções com *se* foram as formas de indeterminação mais utilizadas pelos portugueses (83%), seguida por *nós* (16%), sempre nulo. Quanto ao português brasileiro, houve a preferência pelo uso de *você* (38%), seguido de *se* (23%) e *a gente* (21%). Diferentemente do PE, o PB apresentou formas predominantemente expressas.

Sobre as estratégias de indeterminação em sentenças infinitivas em PE e PB, em geral, tem essa posição do *sujeito vazia*, porém Duarte, Kato e Barbosa (2000) notaram que cerca de 10% das ocorrências apresentaram um elemento na posição do *sujeito*. Em PE esse preenchimento ocorreu, predominantemente, com o uso do *se* (10%), enquanto em PB usou-se *você* (2%) em substituição do *se* (5%).

Considerando os gêneros textuais entrevistas, receitas, e cartas de leitores, analisados no *corpus* sobre os recursos de indeterminação no PE e PB, Duarte, Kato e Barbosa (2000) perceberam que: a) nas entrevistas, o PE apresentou a preferência pelo uso do *se* + verbo transitivo direto com argumento interno no plural e concordância quase categórica. Já no PB, as ocorrências com argumento interno, no plural, mantiveram o verbo no singular; b) nas receitas de cozinha do PE, as pesquisadoras encontraram 16 ocorrências com *se* com concordância, 36 usam o imperativo e uma usa o infinitivo, em um total de 60 receitas. Das 58 receitas no PB, foram registradas 57 construções com o imperativo para indeterminar o *sujeito* e apenas um caso de indicativo com modal; c) em relação ao gênero carta do leitor, os resultados mostraram que impera as construções com *se* nas duas variedades.

O trabalho de Duarte, Kato e Barbosa (2000) mostra que as estratégias de indeterminação do PE e PB são diferentes, pois, no PE, ainda, prevalece o uso do *se* passivo, independentemente da modalidade e dos gêneros, já no PB as estratégias de indeterminação variam conforme a modalidade e o gênero. A pesquisa evidencia que o PB coloquial tende a preencher a posição do *sujeito* de referência arbitrária com formas pronominais nominativas, evitando construções com o *se*, que, geralmente, são mais encontradas em textos formais que se distanciam da oralidade.

Gravina (2014) fortalece a visão apresentada por Duarte, Kato e Barbosa (2000) em relação à língua portuguesa brasileira, considerando-a como, parcialmente, de *sujeitos nulos*, baseada em diferentes pesquisas que apontam para essa possível mudança em direção a esse parâmetro, dentre elas as de Duarte (1995, 2000). A pesquisadora fez um estudo comparativo diacrônico sobre o uso do *sujeito nulo* no PB e PE. Para isso, utilizou um *corpus* constituído por jornais, que circularam no século XIX e início do século XX na cidade de Minas Gerais, no Brasil, e Distrito Évora, Portugal. A hipótese de Gravina (2014) era de que no PB a tendência seria de preenchimento do *sujeito*, diferentemente do PE, que prevaleceria o uso da categoria *vazia*.

Gravina (2014) norteada pelos resultados obtidos em Gravina (2008, *Apud* 2014) que a levou a perceber que o preenchimento do *sujeito* não ocorre somente pelo uso de formas

pronominais, mas também por *sujeitos* lexicais anafóricos, buscou responder a seguinte questão: “O português brasileiro apresentaria a necessidade de preenchimento no decorrer do tempo ou esse preenchimento seria algo estilístico do gênero jornalístico?”.

Considerando o *sujeito* lexical como forma de preenchimento do *sujeito*, Gravina observou em três jornais analisados no PB, “Recreador Mineiro” (1845-1849), “Jornal Mineiro” (1890-1898) e “Tribuna de Ouro Preto” (1945-1948) o percentual de: 69% - 43% -24% em *sujeito* nulos nos períodos de 1845 a 1948 e, 31% - 57% - 76% de *sujeitos* preenchidos nos respectivos períodos, o que a levou a notar um expressivo aumento no preenchimento do *sujeito* no século XX. Em contrapartida, os resultados apontados para a análise dos jornais do PE

“Ilustração luso-brasileira” (1856-1859), “Manuelinho de Évora” (1888-1900) e “Notícias de Évora” (1945-1948) foram de: 82% - 77% - 74% em *sujeito* nulos nos períodos de 1856 a 1948 e, 18% - 23% -26% em *sujeito* preenchidos no mesmo período. Nesse caso, a pesquisadora constatou com base na estabilidade das ocorrências em PE, que ele ainda tem a preferência pelo uso da categoria vazia em posição de *sujeito*.

Um dos aspectos observados e destacados por Gravina em relação ao texto jornalístico brasileiro que merece menção aqui, corresponde ao frequente uso do *sujeito* nulo de primeira pessoa do plural (nós), cuja forma insere a sociedade como um todo no discurso jornalístico, configurando um *sujeito* indeterminado. Já no PE, os dados quanto ao uso da primeira pessoa não foram tão frequentes como no PB.

Comparando as ocorrências de *sujeito* indeterminados com o uso do pronome *nós* em relação ao uso do *se*, Gravina (2014) chegou aos seguintes resultados: em PB as ocorrências com a primeira pessoa só aumentaram ao longo do tempo (38% -54% - 55%), enquanto o *se* demonstrou uma diminuição no seu uso (62% - 46% 45%); em PE apresentou (42% -40% 41%) para a primeira pessoa e (58% - 60% - 61%) em usos com o pronome *se*, mantendo uma estabilidade para as duas formas.

Os resultados encontrados por Gravina (2014) confirmam a hipótese de Duarte, Kato e Barbosa (2000), evidenciando que o PB tende a preencher a posição do *sujeito*, mesmo em se tratando de textos escritos como os jornais impressos, sugerindo a instauração de uma gramática de *sujeito* nulo parcial para essa variedade do português.

Vale ressaltar que, em ambos os estudos aqui apresentados, fica evidente que as duas variedades do português (PB e PE) se comportam de modo diferente em relação ao Parâmetro de *Sujeito* Nulo. Considerando, especificamente, o *sujeito* de referência arbitrária é notório que o PE ainda privilegia construções nulas de indeterminação, bem como com o uso do *se*, tanto

em textos orais quanto escritos, revelando um comportamento condizente com as prescrições gramaticais. No PB essas ocorrências já não estão entre as preferências dos seus falantes, pois evidencia-se, nessa variedade, o preenchimento da posição do *sujeito*, seja por uma forma pronominal ou por uma forma lexical anafórica (DUARTE (1995), (2000); GRAVINA, (2008), (2014)). Além disso, há uma expressiva distância entre a modalidade escrita e a falada no PB, sendo a segunda o contexto que mais favorece a mudança.

1.3 A partícula *se*

Embora o uso da partícula *se* seja apontado como índice de indeterminação do *sujeito*, não analisamos essas ocorrências. Porém, acreditamos ser relevante conhecer o ponto de vista de alguns linguistas sobre esse recurso, haja vista a Gramática Tradicional apresenta-o como uma das principais estratégias de indeterminação do *sujeito*. Portanto, trazemos, nesta subseção, alguns estudos que se propuseram em analisar essas ocorrências.

1.3.1 *Se* apassivador/indeterminador

Tradicionalmente a partícula *se* tem função apassivadora quando o verbo que a acompanha for transitivo direto, nesse caso, há concordância verbal. Por outro lado, a partícula *se* interpretada como índice de indeterminação ocorre quando o verbo empregado for transitivo indireto ou intransitivo e com verbos de ligação, além disso não há concordância entre o argumento interno. (ALMEIDA, 1999; CUNHA e CINTRA, 2007).

Embora as Gramáticas Tradicionais apresentem funções distintas para o uso do *se* (apassivador/indeterminador), geralmente no PB há a preferência pelo uso do *se* indeterminador como em “Aluga-se casas”, “Vende-se doces”, “Plastifica-se documentos”, etc., ficando o primeiro restrito a escrita formal.

Segundo Nunes (1991), o *se* enquanto índice de indeterminação do *sujeito* foi desencadeado por um processo de reanálise sintática das construções com o *se* apassivador e, atualmente, a forma inovadora tem sido a preferência dos brasileiros, mas esse uso genérico também já encontrou outras formas recorrentes: *você* e *a gente*.

Em seu estudo, Nunes (1991) promove um percurso diacrônico das construções com *se* apassivador/indeterminador no português brasileiro. Para a análise, o pesquisador examinou quatro *corpora*: um *corpus* diacrônico compreendendo o período de 1555 a 1989 constituído

por cartas, diários e documentos, provenientes do Acervo de Linguística Histórica da UNICAMP; um *corpus* sincrônico, composto por 13 entrevistas do Banco de Dados da PUCSP; e dois *corpora*: um corresponde a entrevistas do português europeu e o outro a um conjunto de reportagens da revista “Veja”, publicadas entre 1988 e 1989.

Apresentamos, a seguir, uma síntese dos aspectos observados por Nunes (1991):

- a) O *se* indeterminador surgiu, segundo o pesquisador, a partir de uma reanálise sintática do *se* apassivador, podendo ser incidido por três elementos: I) reanálise do clítico (o *se* deixa de ser identificador da estrutura do predicado e passa a ocupar o papel de participante da relação anafórico-pronominal); II) reanálise do argumento interno que deixa de ser *sujeito* para desempenhar o papel de objeto direto; III) reanálise da categoria vazia na posição do *sujeito*, cuja a posição é reanalisada como um pronome nulo referencial;
- b) Ao longo dos séculos (XVI, XVII, XVIII, XIX e XX) as construções com *se* indeterminador no PB aumentaram significativamente, revelando ser a preferência dos brasileiros;
- c) As construções com *se* indeterminador no PE são categóricas, haja vista que nessa variedade há o predomínio do *se* apassivador;
- d) Novas formas de indeterminação surgiram devido a consolidação da mudança do *se* apassivador substituído pelo *se* indeterminador, dentre essas inovações destaca-se o pronome de terceira pessoa e a supressão do *se* em sentenças infinitivas;
- e) Na avaliação sincrônica das formas inovadoras no PB, baseado em dados da revista Veja, Nunes esperava encontrar poucas ocorrências, pelo fato de se tratar de textos jornalísticos, porém foram registradas 156 ocorrências com discordância verbal. Por outro lado, foram encontradas apenas 3 ocorrências com a elisão do *se*. Dessa forma, o pesquisador chegou à conclusão de que a revista assume as formas inovadoras como adequadas ao discurso jornalístico.

Para Nunes (1991), o uso do *se* no português brasileiro passou por um processo de apassivador para indeterminador e, atualmente, tem sido substituído por formas pronominais que exercem a função de indeterminação do *sujeito*, essas formas inovadoras tornaram-se a

preferência dos falantes brasileiros. Por outro lado, o português europeu demonstrou um comportamento conservador em relação ao *se* passivo, além disso, as formas pronominais de terceira pessoa, bem como o uso do *se* sem concordância são interpretados como determinados nessa variedade. Observou-se uma significativa diferença nas construções de estruturas com *sujeitos indeterminados* no português europeu e português brasileiro.

Bessa (2016) também fez um estudo diacrônico e sincrônico sobre o uso da partícula *se*, no intuito de verificar se a mudança natural a todas as línguas também atingiu esse pronome, quando utilizado como índice de indeterminação do *sujeito* e/ou partícula apassivadora. Em seu estudo diacrônico, Bessa (2016) observou que o pronome *se*, como indeterminador do *sujeito*, não existia no período medieval (NARO, 1976) e que, assim como constatou Nunes (1991), o *se* passou por um estágio até que chegasse à interpretação indeterminada. Primeiramente, o *se-reflexivo* precisou transformar-se em partícula apassivadora para que depois a oração se tornasse passiva se considerasse transposta para a ativa, configurando um *se* indeterminador. (LIMA, 2006. *Apud.* BESSA, 2016). Tais construções sem concordância passaram, enfim, a ser aceitas no PE e PB entre o século XVI e XVII, mesmo contradizendo as regras das gramáticas tradicionais. Porém, cabe ressaltar aqui, que no PE a forma inovadora do *se*, índice de indeterminação do *sujeito*, não é tão frequente quanto no PB. (NUNES, 1991).

Ainda em um estudo diacrônico, Bessa (2016), baseada na concepção de Naro (1976), percebeu que o pronome *se* sem concordância com o argumento interno da frase, pode ser considerado indeterminado, mesmo que o verbo seja transitivo.

Com base nesse pressuposto, a análise sincrônica de Bessa se deu por meio de ocorrências referentes as duas estruturas do *se* (apassivador/indeterminador) no PB, coletadas no período de outubro de 2015 a março de 2016 em diferentes sites. Para a estudiosa, essas estruturas, a que ela chama de estágios, são caracterizadas pela interpretação sociodiscursiva que os falantes fazem no momento da enunciação. Sendo assim, na produção de textos como “Plastifica-se documentos”, tem-se a interpretação de que “Alguém tem documentos para plastificar”, então o que a princípio seria um *sujeito*, transforma-se em objeto (documentos), portanto temos, neste caso, um *se* indeterminador. Enquanto em estruturas como “Alugam-se apartamentos”, segue a interpretação passiva “Apartamentos são alugados”. (BESSA, 2016).

Em relação aos resultados encontrados, Bessa percebeu a preferência pelo uso do *se* sem concordância, considerando, então, que ambas as formas (*se*-apassivador/*se*-indeterminador) são aceitas na língua portuguesa brasileira. Embora sentenças como “Plastifica-se documentos” não sejam aceitas pela tradição gramatical, essas construções são

utilizadas sem nenhum prejuízo na comunicação, dessa forma, a autora sugere a consolidação de uma gramática tipicamente brasileira.

Com base nos estudos de Nunes (1991) e Bessa (2016), notamos que o uso do clítico *se* no PB é utilizado preferencialmente como índice de indeterminação do *sujeito*, tendo em vista que, na maioria dos casos, não há concordância verbal. Posto isto, podemos compreender que o uso do *se*, enquanto partícula apassivadora fica restrito a contextos formais.

1.3.2 A realização do *sujeito* em sentenças finitas e infinitivas

Com base no comportamento do *sujeito* de referência arbitrária no PB, e tendo em vista a tendência ao preenchimento na posição do *sujeito* em sentenças finitas e infinitivas, trazemos, a seguir, algumas pesquisas que tratam, especialmente, dessas ocorrências tanto na modalidade oral quanto escrita, nas variedades da língua portuguesa brasileira e europeia.

Duarte (2008) apresenta uma análise do *sujeito* de referência arbitrária em sentenças infinitivas, baseando-se na fala e na escrita padrão do português brasileiro (PB) e português europeu (PE). Para isso, utiliza como *corpus* amostras de fala culta carioca (amostra NURCRJ) e para o português europeu entrevistas extraídas de Cresti e Moneglia (2005). A análise da escrita se deu através de artigos de opinião e crônicas publicados em jornais portugueses e cariocas. Duarte acreditava que a mudança que atingiam as sentenças finitas no PB, considerando, aqui, o preenchimento da posição de *sujeito* nessa estrutura, atingia igualmente as sentenças infinitivas, isso porque a associação da existência do infinitivo flexionado à orientação para o discurso leva a tal realização lexical do *sujeito*.

Duarte (2008) observou, com base em um breve percurso bibliográfico, sobre *sujeito* referenciais arbitrários em sentenças finitas e infinitivas, que, embora as gramáticas tradicionais prescrevam que a indeterminação do *sujeito* ocorra apenas com a forma nula acompanhada da terceira pessoa do plural e estruturas com *se*, diferentes estudos revelaram a preferência dos brasileiros em preencher com formais pronominais nominativas a posição do *sujeito*, estando no topo o uso de *você* e *a gente*, sendo muito frequentes na fala (KATO; TARALLO (1986); DUARTE (1995); GALVES (1987;1998); CAVALCANTE, (2007)).

Duarte (2008) retoma a pesquisa realizada por ela (1995) sobre a realização do *sujeito* em sentenças finitas analisada em dados de fala. Sobre os resultados, a pesquisadora observou que estratégias como *nós* e *se*, estavam sendo substituídas por *você* e *a gente*, o que, segundo Galves (1987, 1998), decorre do enfraquecimento da concordância no PB.

Em se tratando da indeterminação do *sujeito* em sentenças finitas na modalidade escrita do PB, Duarte expõe os resultados obtidos em sua pesquisa (2007), que revelaram um caráter conservador dessa modalidade, mantendo-se fiel a gramática normativa. Além disso, percebeu que o português brasileiro escrito se aproxima muito da forma lusitana, do português europeu oral e escrito. No PB, a pesquisa em questão levantou cerca de 45% de ocorrências com *sujeito nulo* e 36% de estruturas com *se*, marcadas do século XIX, o que fortalece a perspectiva de que a escrita é ainda conservadora.

Tomando como base esse quadro de realizações do *sujeito* de referência arbitrária em sentenças finitas, Duarte (2008) acreditava que as sentenças infinitivas apresentariam as mesmas possibilidades interpretativas. Neste sentido, os dados em relação às sentenças infinitivas fortaleceram a percepção de que o português brasileiro escrito e o português europeu oral e escrito são muito semelhantes, demonstrando, então, o mesmo comportamento em relação a análise das sentenças finitas: o PE, na escrita, apresentou 94% de ocorrências com *sujeitos nulos* e 7% em estruturas com *se*; o PB obteve em média 90% do índice com *sujeitos nulos* e 6% com o uso do *se*. Entretanto, as diferenças entre as duas variedades do português tornam-se mais evidentes quando analisada a fala: no PE, o índice de *sujeito nulo* fica muito próximo à escrita (91% - 94%), além disso houve uso ponderado do *se*; já no PB, esse comportamento muda, pois há uma redução no quadro de ocorrências de *sujeitos nulos* (78%), o índice de *sujeito* expressos chega a 22%, sendo que a primeira opção dos falantes brasileiros foi o uso de *você* (18,5%), seguido de *a gente* (2%) e *se* (1,5%).

Duarte (2008) percebeu que os resultados com *se* não tinham motivação funcional, portando sua ausência não comprometeria a interpretação arbitrária. Entretanto, observou uma motivação estrutural: nas três estruturas com o uso do *se* em PB e nas quatro das seis ocorrências em PE, houve a presença de preposição; em apenas duas ocorrências, no PE, o *se* apareceu expresso em orações subordinadas não regidas de preposição; quanto ao uso de *a gente*, a pesquisadora notou que o pronome aparece em orações subjetivas, articuladas com predicadores adjetivais.

Quanto ao uso de *você*, a forma preferida dos brasileiros na realização do *sujeito* de referência arbitrária, Duarte observou o destaque de um aspecto extralinguístico e um linguístico: o primeiro diz respeito a faixa etária, na qual o uso do pronome mostrou-se mais recorrente nos grupos mais jovens de falantes; o segundo corresponde ao uso de *você* não regido por preposições, diferentemente das ocorrências com *se*, que essa estrutura foram evidenciadas. Ambos os fatores também são identificados no uso do pronome *a gente*.

Em relação à análise qualitativa dos dados obtidos na escrita, Duarte observou que tanto no PE quanto no PB o uso do *se* previne uma leitura correferencial com um antecedente, garantindo uma interpretação indeterminada do *sujeito*. Além disso, no PE a ausência ou presença de preposição não se mostra tão atuante como no PB.

Posto isto, os resultados obtidos por Duarte (2008) indicam que, se por um lado o PB é fiel às normas prescritas pelos compêndios gramaticais manifestados na escrita, modelo lusitanizado, por outro lado a fala é regida por uma gramática brasileira própria. Além disso, os dados fortalecem a concepção de que o PB apresenta, no seu sistema, diferentes possibilidades de indeterminar o *sujeito*, apontando para uma possível mudança em direção ao Parâmetro do *Sujeito Nulo*.

Corroboram a proposição apresentada por Duarte (2008), os resultados da pesquisa realizada por Cavalcante e Duarte (2009), que expõem uma análise dos *sujeitos* de referência definida e arbitrária em PE e PB em sentenças infinitivas. Para a pesquisa, as autoras basearam-se no mesmo *corpus* utilizados por Duarte (2008)¹. O objetivo das pesquisadoras era tentar responder as seguintes questões: “O preenchimento da posição de *sujeito* está relacionado à mudança na remarcação do Parâmetro do *Sujeito Nulo*? Como uma língua parcialmente orientada para o discurso como o PB permite o infinitivo flexionado?”. Como hipótese, as autoras acreditavam que as duas variedades do português apresentariam motivações diferentes para o preenchimento do *sujeito* de infinitivo. Esperavam que o PE apresentasse uma distribuição complementar, enquanto no PB encontrariam *sujeito* não expressos e expressos.

Cavalcante e Duarte (2009) constataram um comportamento diferente do PB em relação ao PE, considerando a posição do *sujeito* em sentenças infinitivas. Isso ocorreu tanto em contextos de referência definida quanto arbitrária, porém as diferenças entre as duas variedades ficaram mais perceptíveis na análise do *sujeito* de referência arbitrária.

Em relação ao *sujeito* de referência definida, Cavalcante e Duarte (2009) perceberam que no PE e no PB ainda é predominante a categoria vazia na posição de *sujeito* em sentenças infinitivas: na fala, no PE, foram registrados 98% de *sujeitos nulos*, e apenas 0,7% de *sujeito pronominal*, já no PB, foram contabilizados 90% de *sujeitos nulos*, e 6% de *sujeitos pronominais*; Em se tratando da escrita, no PE, o índice de *sujeito nulo* chegou a 90%, e

¹ O *corpus* constituído tanto na pesquisa de Duarte (2008) quanto no estudo de Cavalcante e Duarte (2009) basearam-se em amostras de fala, para o PB, do Projeto NURC-SP (Norma Urbana Culta), gravadas nos anos 90 e, para o PE, foram utilizadas entrevistas extraídas de Cresti e Moneglia (2005). A análise da escrita também se deu através de artigos de opinião e crônicas publicados entre 2003 e 2007 em jornais de Lisboa e do Rio de Janeiro.

nenhuma ocorrência com *sujeito pronominal*, em contrapartida o PB apresentou um total de 92% de *sujeitos nulos* e 2% de *sujeitos pronominais*. Com base nesses resultados, as pesquisadoras notaram a aproximação da escrita do português brasileiro com as formas falada e escrita do português europeu.

Os resultados obtidos por Cavalcante e Duarte (2009), sobre a realização do sujeito de referência arbitrária, indicaram que o português europeu falado e escrito e o português brasileiro escrito estão muito próximos, o *sujeito nulo* obteve em média 90% dos resultados e o clítico *se* ficou entre 6% e 8%. Todavia, o que chamou a atenção das pesquisadoras foi o uso de *você* no PB oral, com 18,5% de preenchimento, o que poderia indicar um encaixamento da mudança.

Vale ressaltar que os resultados alcançados por Cavalcante e Duarte (2009) ampliam a pesquisa de Duarte (2008), pois as pesquisadoras perceberam que a explicação para a diferença entre o uso do infinitivo no PE e PB, já observada pela autora aqui citada (2008), pode estar relacionada ao tipo de concordância de infinitivo nas duas variedades. O PE possui altos índices de *sujeito nulo*, o que pode ser resultado de um infinitivo não-flexionado, por outro lado no PB ocorre o inverso dessa situação, pois essa variedade tem características de flexão finita para sentenças infinitivas, dessa forma, a concordância é capaz de licenciar uma categoria vazia, porém não consegue interpretá-la, sendo assim, o preenchimento da posição do *sujeito* por uma forma pronominal pode interpretá-la corretamente, sugerindo, então, uma mudança em direção a uma língua de proeminência de *sujeito* e tópicos. (CAVALCANTE e DUARTE, 2009).

Ainda sobre a indeterminação do *sujeito* em sentenças infinitivas, porém agora, considerando as não-flexionadas, Colsato (2009) apresenta uma análise do sistema de referencialidade do português brasileiro, baseando-se na descrição de contextos que favorecem o uso do *se* em oposição a categoria vazia. Nesse sentido, a pesquisadora observou o comportamento das estratégias do *sujeito* de referência definida e arbitrária de verbos infinitivos não-flexionados, seguindo a metodologia de testes de produção, o que possibilitou a verificação estatística do desempenho das variantes.

Em relação aos resultados, Colsato (2009) apresentou alguns aspectos interessantes:

- a) Em orações adverbiais, a categoria vazia privilegia a leitura determinada do *sujeito*, enquanto o *se* marca a referência indeterminada;
- b) Em orações completivas de nome e adjetivo, a categoria vazia favorece a interpretação indeterminada, além disso o uso de *se* acompanhado de preposição também marca o *sujeito* indeterminado;

- c) Nas completivas, prevalece a forma nula de *sujeito* de verbos infinitivos para a referência arbitrária do *sujeito* e, ainda, houve casos de pronome tônico exercendo a função indeterminada;
- d) Sobre os verbos inacusativos, a categoria vazia possibilitava a interpretação de um *sujeito* determinado, dessa forma, para a marcação arbitrária os informantes buscavam por um elemento pleno, geralmente o pronome *se*;
- e) No embate com os verbos finitos, em um processo de testes, os infinitivos não-flexionados asseguraram interpretação referencial determinada, enquanto a referência arbitrária mostrou-se marcada pelo uso do pronome *se*, mantendo sua função de indeterminar o *sujeito*.

Na pesquisa de Colsato (2009), os resultados indicam a predominância da categoria vazia e construções com *se* como recursos de indeterminação do *sujeito* em sentenças infinitivas não flexionadas, diferentemente dos resultados apontados por Duarte (2008) e Cavalcante; Duarte (2009) em estudo sobre as sentenças infinitivas flexionadas.

1.4 Finalizando

Nesta seção foram apresentadas duas diferentes perspectivas que buscam conceituar e classificar os recursos de *indeterminação do sujeito*: a primeira concepção impera nos manuais de gramática normativa (geralmente apresentam os recursos de indeterminação como o uso do verbo na 3ª pessoa do plural e construções com a partícula *se*); a segunda, de cunho linguístico, busca descrever a língua tal como ela acontece. Nesta última, a indeterminação do *sujeito* recorre a estratégias não prescritas pelos manuais normativistas. E é a ela que nos filiamos.

Vale ressaltar que vemos a perspectiva prescritiva não como deficiente na abordagem sobre os fenômenos linguísticos, mas restrita a contextos específicos que não contemplam a realidade dos falantes. Nesse sentido, as estratégias de *indeterminação do sujeito*, no PB, apresentadas pelas gramáticas tradicionais imperam na modalidade escrita em contextos formais, pelo fato de ser, predominantemente, conservadora, assim observado em trabalhos como de Duarte, Kato e Barbosa (2000), Gravina (2014). Por outro lado, a fala favorece o uso de inúmeros recursos de indeterminação perfeitamente aceitáveis no processo de interação verbal, desse modo, as formas tradicionais são substituídas por outras inovadoras como *você*, *a*

gente, eu, tu, nós, dentre outras, assim constatado em pesquisas como de Duarte (2008) e Cavalcante e Duarte (2009).

Nos estudos linguísticos aqui apresentados, percebemos a incansável afirmativa de que o português brasileiro caminha para uma mudança em relação ao *Parâmetro do Sujeito Nulo*, podendo ser considerada como uma língua de *sujeitos nulos* parciais. Essa evidência pode ser observada quando comparada com o PE que demonstra, ainda, um comportamento conservador no uso da língua, privilegiando estruturas nulas na realização do *sujeito* tanto em contextos de referência definida quanto arbitrária.

Os trabalhos de cunho descritivo abordados, nesta seção, foram relevantes tanto para emprendermos a acepção dos linguistas acerca do fenômeno da indeterminação, como para conhecermos o atual quadro do português brasileiro em relação ao *Parâmetro do Sujeito Nulo*, com foco no estudo do *sujeito* de referência arbitrária. Sendo assim, na próxima seção, discorreremos sobre algumas pesquisas que tratam do fenômeno em questão especificamente na modalidade falada do PB. Os resultados dessas pesquisas serão retomados na última seção, de análise.

SEÇÃO II

A INDETERMINAÇÃO DO *SUJEITO* NAS VARIEDADES DO PORTUGUÊS BRASILEIRO

Nesta seção apresentamos os resultados de alguns trabalhos empíricos que descrevem a realização do *sujeito* no português brasileiro, com o foco na modalidade oral. A revisão desses estudos possibilitará, mais adiante, fazer uma comparação entre os dados obtidos nesses trabalhos com aqueles resultantes de nossa pesquisa.

Selecionamos pesquisas sobre o *sujeito* arbitrário realizadas em quatro estados brasileiros, sendo eles: Rio de Janeiro, variedade carioca descrita por Duarte (1995); Paraná, variedade do interior paranaense estudado por Godoy (1999); a variedade de falantes araguienses dos estados de Mato Grosso e Goiás, analisada por Carvalho (2005); e o estado da Bahia, variedade de Feira de Santana pesquisada por Assunção (2012). Esses trabalhos mostram o funcionamento do fenômeno da indeterminação do *sujeito* perante ao *Parâmetro do Sujeito Nulo*, sendo assim, observaremos quais as estratégias demonstram mais efetividade na fala dos brasileiros: se aquelas legitimadas pelas gramáticas tradicionais, ou as formas inovadoras.

2.1 O sujeito arbitrário no português brasileiro

Contamos, atualmente, com uma vasta bibliografia sobre a realização do *sujeito* de referência arbitrária no português brasileiro, porém, selecionamos apenas algumas autoras que buscam descrever o fenômeno da indeterminação em algumas regiões do país. Dentre elas destacamos Duarte (1995). A pesquisadora apresenta uma análise sobre a realização do *sujeito* no português brasileiro tanto de referência definida quanto arbitrária. Embora a referência definida tenha sido discutida mais efetivamente por Duarte (1995), traremos os resultados somente sobre o *sujeito* de referência arbitrária, para mantermos a objetividade deste trabalho. Duarte (1995) analisou amostras de fala de informantes cariocas do ensino superior. Foram consideradas as variantes linguísticas *concordância verbal* e *tempo e modo verbal*, em relação aos fatores extralinguísticos a pesquisadora selecionou a faixa etária (25 a 32 anos, de 45 a 53 anos e de 59 a 74 anos) e sexo (feminino e masculino). O objetivo de Duarte (1995) consistia em verificar a implementação da mudança que afeta o *Parâmetro do Sujeito Nulo*, para isso, comparou seu *corpus* com entrevistas de universitários em rádio e TV.

Duarte (1995) constatou que os falantes universitários preferiram utilizar, como estratégias de *indeterminação do sujeito*, as formas pronominais nominativas expressas, estando no topo o uso do pronome “você”, comum em todas as faixas etárias. Em seguida, o pronome “a gente” mostrou-se mais produtivo, seguido de “eles”. Em contrapartida, o uso do pronome “nós” não foi tão atuante como Duarte esperava registrar, além disso, o pronome *se* apresentou baixa ocorrência, resultado do aumento do uso das formas pronominais nominativas.

Comparando os dados obtidos no seu *corpus* com os dados evidenciados em entrevistas de universitários à rádio e à TV, Duarte (1995) atestou a implementação das estratégias de indeterminação “você” e “a gente”. Todavia, nas entrevistas à TV, observou-se a preferência pelo uso de recursos descritos pelas gramáticas tradicionais. Em relação ao fator tempo e modo verbal, a pesquisadora evidenciou que o presente do indicativo é o mais corrente nas construções com *sujeito* indeterminado, porém, afirmou que a referência arbitrária tem se expandido às construções modais e tempo simples com verbos no presente e pretérito perfeito do modo indicativo.

A tendência ao preenchimento do *sujeito* no português brasileiro também foi observada por Godoy (1999), em sua pesquisa sobre a *indeterminação do sujeito* no interior paranaense. A pesquisadora baseou-se em amostras de fala de informantes de Irati, Londrina e Pato Branco. O *corpus* constituído faz parte do projeto VARSUL (Variação Linguística Urbana na Região Sul do país). Para o processo de análise, Godoy (1999) considerou as variáveis linguísticas *tempos* e *modos verbais* e os fatores extralinguísticos como *sexo*, *faixa etária*, *escolaridade* e *localidade*. A estudiosa acreditava que, no interior paranaense, encontraria outras estratégias de indeterminação do *sujeito*, além das legitimadas pelas gramáticas tradicionais e, ainda, tinha a intenção de verificar, em relação aos tempos e modos verbais, se o tempo presente do indicativo seria o mais dominante na *indeterminação do sujeito*.

No que se refere aos resultados, Godoy (1999) atestou que no interior paranaense utilizam-se outras estratégias de indeterminação do *sujeito* que não estão descritas pelos compêndios gramaticais, aliás, os recursos apresentados pelas gramáticas tradicionais mostraram-se menos produtivos do que àqueles considerados informais.

A seguir algumas considerações pertinentes, segundo Godoy (1999, p. 168-170), em relação à realização do *sujeito* na variedade paranaense do português brasileiro:

- a) A variedade paranaense demonstrou tendência em preencher a posição do *sujeito*, o que fortalece a afirmativa de que o português brasileiro caminha para uma mudança em direção ao *sujeito* foneticamente realizado;

- b)** A estratégia “você” mostrou-se mais atuante para a indeterminação do *sujeito*; O uso de “a gente” teve maior representatividade na fala de informantes do sexo feminino;
- c)** O pronome “eu” enquanto recurso de indeterminação teve a mesma frequência tanto na fala de homens quanto de mulheres mais jovens, já em relação a segunda faixa etária (acima de 50 anos), tal estratégia mostrou-se menos ativa na fala de mulheres;
- d)** Em relação ao pronome “tu”, este mostrou-se pouco usado pelas mulheres da segunda faixa etária;
- e)** O uso das estratégias prescritas pelas gramáticas tradicionais mostrou-se mais produtivas na fala de informantes de maior escolaridade;
- f)** Considerando as cidades selecionadas para a pesquisa, constatou-se que na cidade de Irati, utilizam-se com mais frequências os recursos apresentados pelas GTs.

Godoy (1999) confirmou sua hipótese de que no interior paranaense utiliza-se outros tipos de estratégias de indeterminação do *sujeito*, que vão além daquelas prescritas pelas gramáticas normativas tradicionais, evidenciando como principais recursos os pronomes e as formas nominais. Além disso, afirma que uma ou outra preferência decorre tanto de fatores linguísticos quanto extralinguísticos, o que de fato, foi constatado na análise das ocorrências com o pronome *se*, cuja estrutura mostrou-se mais atuante em falantes com maior nível de escolaridade, indicando que a escola, provavelmente, exerce influência sobre a preferência dos falantes.

Ampliando o estudo sobre realização do *sujeito* no português brasileiro, Carvalho (2005) realizou uma análise/descrição da fala do araguaense, tanto de referência definida quanto arbitrária, porém, interessa-nos, aqui, somente a descrição da realização do *sujeito* de referência arbitrária dessa variedade do português.

Carvalho (2005) selecionou, para a constituição do seu *corpus* de análise, duas comunidades de fala, de dois estados brasileiros: Alto Araguaia (Mato Grosso) e Santa Rita do Araguaia (Goiás). A intenção da pesquisadora era verificar se nessas regiões haveria o predomínio do *sujeito* preenchido, da mesma forma que constatado por Duarte (1995) e Godoy (1999). Além disso, buscou identificar quais os contextos que favorecem o *sujeito nulo*, bem como os contextos que contribuem para o preenchimento do *sujeito*. O *corpus* constituído abrangeu entrevistas com 24 informantes. A autora considerou para o processo de análise do

sujeito de referência arbitrária os seguintes fatores extralinguísticos: escolaridade, faixa etária e sexo. Além disso, baseou-se nas variáveis linguísticas: correferencialidade do *sujeito*, a pessoa gramatical, a forma verbal, o tipo de oração e a voz verbal. Carvalho (2005) contemplou todas as ocorrências com: *sujeito nulo*, SN pleno, pronome pessoal, SN modificado e pronome demonstrativo. Vale ressaltar que Carvalho não considerou as ocorrências com a terceira pessoa do singular com *sujeito nulo*, acompanhada ou não da partícula *se*, pois, segundo a estudiosa, tais realizações encontram-se em desuso pelos araguienses.

Carvalho (2005) registrou 2384 ocorrências, dessas 440 eram de referência arbitrária, 66% das ocorrências com *sujeito* genérico tinham o *sujeito* preenchido com pronomes pessoais. Tal resultado, de acordo com a pesquisadora, sugere um processo de mudança rumo ao *sujeito* expreso.

Ao detalharmos os resultados obtidos por Carvalho (2005) em relação aos grupos de fatores e seus cruzamentos veremos que:

- a) A correferencialidade foi apontada pela análise probabilística como a que exerce o mais alto grau de significância na realização do *sujeito*, os resultados mostraram que apesar das estruturas não correferentes favorecerem o uso do *sujeito* preenchido, em estruturas em que o *sujeito* é o mesmo da sentença anterior a casa também se mostrou com alto índice de preenchimento, demonstrando certo enfraquecimento desse fator;
- b) No fator tipo de oração, a referência mostrou interferência tanto em contextos que favorecem quanto nos que retardam à mudança, sendo assim, destacou-se como o uso do pronome-*sujeito* a oração relativa, enquanto as orações independentes mostraram-se mais resistentes, bem como as orações de 2ª coordenadas;
- c) Quanto à forma verbal, o pretérito perfeito do indicativo demonstrou resistência ao preenchimento do *sujeito*, o infinitivo apresentou altos índices de preenchimento, contrariando às normas prescritivas e, ainda, o presente do indicativo e o futuro do subjuntivo mostraram-se mais favoráveis ao preenchimento do *sujeito*;
- d) Em relação à pessoa gramatical, notou-se um avanço na mudança em direção ao preenchimento do *sujeito* nas 1ª e 2ª pessoas, enquanto a 3ª pessoa mostrou-se mais resistente à mudança;
- e) A voz verbal ativa mostrou-se bastante produtiva nas ocorrências com *sujeito* preenchido, além disso a voz passiva teve baixos índices de *sujeito* nulos;

- f) Destaca-se como estratégias de indeterminação do *sujeito* utilizada pelos araguienses as formas “você” e “a gente”;
- g) Os grupos de fatores sexo, escolaridade e faixa etária não foram selecionados pela análise probabilística como significativos na realização do *sujeito* de referência arbitrária.

Tendo em vista a realização do *sujeito* de referência arbitrária no PB, Assunção (2012) promove um estudo dos recursos de indeterminação do *sujeito* na variedade de Feira de Santana, Bahia. Para o trabalho, a pesquisadora utilizou o *corpus* proveniente do projeto intitulado “A língua portuguesa no semiárido baiano”, do Núcleo de Estudos da Língua Portuguesa (NELP). Para a análise, Assunção (2012) considerou os seguintes grupos de fatores: *tempo* e *modo verbal*, *tipo de oração*, *tipo de verbo* e *preenchimento ou não do sujeito*. Como fatores extralinguísticos levou em conta: sexo, faixa etária e escolaridade do falante.

Em seu estudo, Assunção (2012) evidenciou outras estratégias de indeterminação do *sujeito*, além das legitimadas pelas gramáticas tradicionais, dentre elas destacam-se: “eu”, “você”, “a gente”, “nós” e “eles”, sendo que as variantes “a gente” e “você” encontram-se no topo da preferência dos feirenses.

Convém frisarmos aqui, que a variante “tu”, ao contrário do que presumia Assunção (*ibidem*), não teve nenhum registro de ocorrência para efeito de indeterminação, demonstrando mais efetividade na referência definida. Quanto ao uso de “nós”, essa variante não se mostrou tão atuante, sendo, por muitas vezes, substituída pelo uso de “a gente”.

Tomando como base a variante “você”, por ser a principal estratégia de indeterminação utilizada pelos falantes de Feira de Santana, Assunção (2012) mostrou que os mais jovens são os que mais fazem uso desse recurso (59%), enquanto a terceira faixa etária, representada por um grupo de falantes acima de 65 anos, quase não utiliza essa variedade (3%), optando pelo uso de “a gente”. Nesse sentido, é notória a influência da faixa etária no uso de uma ou outra variedade, considerando que os mais jovens usam com mais frequência as formas inovadoras. Além disso, Assunção atestou que o uso da variante “você” não apenas concorre com a estratégia “a gente”, mas a substitui, pois, a pesquisadora percebeu que enquanto uma diminuía a outra aumentava a sua frequência na fala dos feirenses.

Considerando o fator sexo, Assunção (*ibidem*) atestou que os homens utilizam com uma frequência maior as variantes inovadoras. No que se refere à escolaridade, a pesquisadora surpreendeu-se ao evidenciar que as estratégias como o uso de “você” para a indeterminação

do *sujeito* foram mais produtivas na fala dos mais escolarizados. Nesse caso, diferentemente dos resultados de Godoy (1999), que evidencia a influência no uso das estratégias de indeterminação pelos falantes, pelo fato de privilegiar estruturas legitimadas pelas gramáticas tradicionais, e claramente observada na fala dos mais escolarizados, Assunção (2012), por outro lado, constatou um alto índice no uso de estratégias não regulamentadas pelas GTs, utilizadas por falantes com curso superior, indicando que não há restrições no uso das variantes inovadoras.

Assunção (2012), ainda, constatou que o fator linguístico tempo e modo verbal exerce grande influência no uso das estratégias de indeterminação inovadoras, estando entre o mais atuante o presente do indicativo, reforçando os resultados de Duarte (1995), Godoy (1999) e Carvalho (1995) que afirmam que esse tempo e modo verbal privilegiam a indeterminação do *sujeito* com estruturas pronominais, pelo fato de apresentar uma característica atemporal, segundo Menon (1994). Contudo, vale destacar, aqui, que ao analisar os pesos relativos, Assunção demonstrou que o futuro do subjuntivo e o infinitivo condicionam o uso da variante “você”.

Segundo Assunção (2012, p.107) a indeterminação do *sujeito* “não se prende a uma estrutura do *sujeito* lexical não preenchido como exigem as gramáticas tradicionais”, fato justificado devido à indeterminação ser considerada como um fenômeno semântico e discursivo, portanto, o que se indetermina não é o *sujeito*, mas sim sua referência, nesse sentido ela é condicionada ao contexto comunicacional, dando origem a diferentes formas de indeterminar o *sujeito*. Nesse sentido, não há como limitar a indeterminação do *sujeito* somente às estratégias como: uso de verbo na terceira pessoa do plural e estruturas com a partícula “se”.

Com base nos estudos apresentados, aqui, podemos perceber a velocidade com que a mudança em direção ao *sujeito* preenchido está sendo implementada no português brasileiro. Apesar de não ser considerada como mudança concluída, os dados sugerem que esse processo avança à medida que o tempo passa. Se considerássemos isoladamente uma pesquisa, não seria possível fazer essa afirmação, porém, em uma revisão bibliográfica, o extenso volume de trabalhos sobre a realização do *sujeito* de referência arbitrária revela a preferência dos brasileiros em preencher a posição do *sujeito*.

Para ilustrarmos os resultados obtidos nas pesquisas de Duarte (1995), Godoy (1999), Carvalho (2005) e Assunção (2012), apresentamos, a seguir, o **quadro-síntese 1** com algumas informações que consideramos pertinentes em relação às pesquisas. Destacamos o percentual

obtido em cada estudo no que se refere às variantes *sujeito preenchido* x *sujeito nulo*, de modo a compararmos os resultados em relação a cada variedade do português brasileiro.

Quadro-síntese 1

O SUJEITO DE REFERÊNCIA ARBITRÁRIA NAS VARIEDADES DO PORTUGUÊS BRASILEIRO
Pesquisadora: DUARTE (1995) – Variedade pesquisada: A fala carioca <i>Sujeito preenchido</i> : 65% <i>Sujeito nulo</i> : 35%
Pesquisadora: GODOY (1999) – Variedade pesquisada: A fala do interior paranaense <i>Sujeito preenchido</i> : 62% <i>Sujeito nulo</i> : 38%
Pesquisadora: CARVALHO (2005) – Variedade pesquisada: A fala araguiense <i>Sujeito preenchido</i> : 66% <i>Sujeito nulo</i> : 34%
Pesquisadora: ASSUNÇÃO (2012) – Variedade pesquisada: A fala dos feirenses <i>Sujeito preenchido</i> : 87% <i>Sujeito nulo</i> : 13%

Fonte: Quadro elaborado pela autora, a partir dos dados das pesquisas.

Os principais recursos de indeterminação do *sujeito* utilizados nas variedades do português brasileiro, apresentados nos estudos de Duarte (1995), Godoy (1999), Carvalho (2005) e Assunção (2012) foram os pronomes pessoais, sendo assim, as estratégias legitimadas pelas gramáticas tradicionais demonstraram pouca eficácia pelos falantes da variedade carioca, paranaense, araguiense e feirenses.

Posto isto, faz-se necessário compreender quais fatores linguísticos foram mais produtivos no uso dessas estratégias destacadas em cada estudo aqui exposto. A seguir, apresentamos o **quadro-síntese 2**, dos fatores que demonstraram maior efetividade no preenchimento do *sujeito*.

Quadro-síntese 2

FATORES LINGUÍSTICOS MAIS SENSÍVEIS AO PREENCHIMENTO DO SUJEITO DE REFERÊNCIA ARBITRÁRIA
Pesquisadora: DUARTE (1995) – Variedade pesquisada: a fala carioca Favorece o <i>sujeito preenchido</i> : Os pronomes “você” e “a gente”, geralmente de 1ª e 2ª pessoas; o tempo e modo verbal presente do indicativo.
Pesquisadora: GODOY (1999) – Variedade pesquisada: a fala paranaense Favorece o <i>sujeito preenchido</i> : Os pronomes “você”, “eles” e “a gente”; tempo e modo verbal presente do indicativo.

Pesquisadora: CARVALHO (2005) – Variedade pesquisada: a fala araguiense Favorece o <i>sujeito preenchido</i>: Os pronomes “você” e “a gente”, geralmente de 1ª e 2ª pessoas; a voz verbal ativa; as formas verbais presente do indicativo e futuro do subjuntivo; as orações relativas e as independentes de 2ª coordenadas.
Pesquisadora: ASSUNÇÃO (2012) – Variedade pesquisada: a fala dos moradores de Feira de Santana Favorece o <i>sujeito preenchido</i>: Os pronomes “a gente” e “você”; o tempo e modo verbal presente do indicativo.

Fonte: Quadro elaborado pela autora, a partir dos dados das pesquisas.

2.2 Finalizando

Nesta seção discorreremos sobre alguns estudos que se referem à realização do *sujeito* de referência arbitrária em algumas variedades do português brasileiro: a fala carioca (RJ), a fala do interior paranaense (PR), a fala araguiense (MT e GO) e a fala dos feirenses (BA). Tais estudos descrevem o atual quadro de realização do *sujeito* no português brasileiro, de modo a mostrar a evolução do processo de mudança em relação ao *Parâmetro do Sujeito Nulo*, tendo em vista que nessa variedade os falantes optam por outros recursos que não se limitam à omissão do *sujeito* em contextos arbitrários, de modo que as variedades inovadoras sobrepujam à categoria vazia.

As pesquisas de Duarte (1995), Godoy (1999), Carvalho (2005) e Assunção (2012), embora tratem de variedades diferentes do português brasileiro, apresentam resultados semelhantes, revelando diversas estratégias para a indeterminação do *sujeito*. Além disso, constata-se a ineficácia dos recursos de indeterminação prescritos pelas gramáticas tradicionais, de modo a serem reconhecidos apenas em contextos formais e não em situações reais de uso da língua.

O estudo de Duarte (1995), Godoy (1999), Carvalho (2005) e Assunção (2012), sobre a indeterminação do *sujeito* em variedades do português brasileiro, corroboram os resultados de pesquisadores como Tarallo (1972), Duarte (1993) entre outros, que postularam a perda do *status* de língua de *sujeitos nulos* para o PB. Nesse sentido, por meio da nossa pesquisa, buscamos confirmar ou refutar tal afirmativa com base no estudo do *sujeito de referência arbitrária* na variedade carioca. Para isso, selecionamos os fatores linguísticos que, com base nas pesquisas de Duarte (1995), Godoy (1999), Carvalho (2005) e Assunção (2012) afetam negativa e positivamente o preenchimento do *sujeito* de referência arbitrária. Os fatores

linguísticos, extralinguísticos, bem como nossos objetivos e hipóteses são discutidos com mais detalhes na próxima seção.

SEÇÃO III

ABORDAGEM TEÓRICO-METODOLÓGICA

Este trabalho está inscrito nos estudos sociolinguísticos de linha laboviana, nesse sentido pretendemos apresentar, nesta seção, a abordagem que ancora nossa pesquisa, que tem como precursor o pesquisador William Labov. Além disso, comentamos sobre a metodologia utilizada, bem como as hipóteses, os objetivos e os fatores linguísticos e sociais selecionados para o processo de análise.

3.1 A Teoria Variacionista

A Sociolinguística é uma área ou campo da linguística. Neste campo se inter cruzam outros campos como o social e o linguístico, ou seja, a sociolinguística se propõe investigar a relação entre o mundo linguístico e o mundo social. Esse ramo surgiu com o advento da linguística, segundo movimentos teóricos que buscavam mostrar o social como parte do funcionamento da língua.

Os trabalhos na área da sociolinguística partem do pressuposto de que há uma relação entre o corpo da língua e seu exterior. Nessa abordagem, considera-se primeiro a autonomia do sistema linguístico, para depois proporem a inter-relação com o mundo social. Diferentemente de outras abordagens, segundo Pagotto (2006, p.51), “[...] a sociolinguística trabalha com a ideia de que o sistema linguístico tem um funcionamento próprio, independente do mundo social embora submetido a ele”.

Mesmo parecendo um tanto paradoxal considerar o sistema linguístico autônomo e dependente ao mesmo tempo em relação à estrutura social na qual o sistema atua, todo trabalho que utiliza dessa abordagem teórica procura investigar satisfatoriamente questões linguístico-sociais.

Na sociolinguística, buscam-se lugares de intersecção entre o mundo social e a dimensão linguística. E essas intersecções normalmente são definidas por meio do funcionamento linguístico. As linhas de pensamento da sociolinguística se estabelecem do lugar que privilegiam tal intersecção.

Reconhecem-se como grandes áreas da sociolinguística as linhas de pensamento: Teoria da Variação e Mudança Linguística, Etnografia da Fala e Sociologia da Linguagem.

Todas estudam aspectos relevantes do funcionamento da língua na sociedade, porém, nos interessa, aqui, a metodologia que abrange a Teoria da Variação e Mudança Linguística, pelo fato de ser a linha de pensamento na qual ancoramos nossa pesquisa.

A Teoria da Variação e Mudança Linguística, também chamada Sociolinguística Quantitativa, pelo fato de utilizar na sua metodologia formas matemáticas de quantificação de dados, se interessa em discutir a maneira como o sistema linguístico é afetado não apenas pelas relações enunciativas, mas também pelas relações de classe e outros grupos sociais. Nesse sentido, considera-se como um conjunto de fatores definidos, com os quais se supõe que a diversidades linguística esteja relacionada: o gênero, a faixa etária, a escolaridade e a região geográfica onde vive o falante. (TARALLO, 2001, p.8).

Willian Labov foi o pioneiro nos estudos da Variação e Mudança Linguística e juntamente com Weinreich e Herzog formalizaram esse novo modelo teórico. Labov (1963) desenvolveu uma pesquisa sobre a variação entre ditongos no inglês falado em Martha's Vineyard no Estado de Massachussetts. Mais tarde, em 1964, o pesquisador realizou um estudo sobre a estratificação social do inglês em Nova York, dessa forma o autor fixou uma metodologia que descreve e interpreta fenômenos linguísticos, considerando como aspectos essenciais no processo de variação fatores internos e externos à língua. Esse modelo teórico pressupõe que todo sistema linguístico é caracterizado por sua sistematicidade e heterogeneidade, e por isso é passível a variações e mudanças linguísticas ao longo do tempo.

O fato da mudança linguística estar relacionada à sistematicidade da heterogeneidade linguística, implica dizer que a língua apresenta uma heterogeneidade ordenada e não aleatória, portanto, é regulada por um conjunto de regras, por fatores variáveis independentes que podem ser tanto linguísticos quanto sociais. Dessa forma, a sociolinguística variacionista busca identificar e medir cada fator, descrevendo o seu funcionamento no processo de variação e mudança, para isto a teoria dispõe de um método quantitativo que mede o grau de influência desses fatores.

Segundo Labov (1972, [2008]) a variação é regra na língua, é prova de que elas estão vivas. São produtos sociais e como tal estão condicionadas a controles sociais, que “manipulam” as formas linguísticas. Nesse sentido, novas formas surgem, sem que isso signifique que uma se sobreponha à importância da outra, o que acontece, nesse caso, é um processo de encaixamento, que uma nova variante após um período de variação é implementada no sistema linguístico e social.

Vale ressaltar aqui, que, conforme Labov (1972, [2008]), toda mudança surge da variação, entretanto, nem toda variação pressupõe mudança linguística. Nesse caso, objetiva-se, com a metodologia sociolinguística, observar e descrever as variações como um processo inerente a todas as línguas, observando quais os fatores que retardam ou aceleram o processo de mudança.

Tendo em vista todas as discussões expostas aqui, adotamos a abordagem teórico-metodológica da Sociolinguística Variacionista pela sua adequação à nossa pesquisa, cujo objeto de investigação é o *sujeito* de referência arbitrária na fala carioca.

3.2 O Parâmetro do Sujeito Nulo

Podemos definir *sujeito nulo* como uma interpretação de *sujeitos referenciais* (definidos ou indefinidos) em sentenças, em que um pronome, um sintagma nominal ou expressão referencial estejam subentendidos, apresentando a posição do *sujeito* vazia (VERÍSSIMO, 2017).

O Parâmetro *pro-drop* é um conjunto de propriedades correlacionadas, dentre as quais inclui-se o *sujeito* vazio, também denominado *Parâmetro do Sujeito Nulo*. Esse parâmetro enquadra-se na Teoria de Princípios e Parâmetros delineada por Chomsky (1981). Essa teoria, em se tratando da realização do *sujeito*, divide a língua em dois grupos: as que permitem uma categoria vazia na posição do *sujeito* sem que isso comprometa a gramaticalidade da sentença, denominada, nesse caso, *[+] pro-drop*, e as que exigem *sujeito* expresso, caracterizada *[-] pro-drop*. (CARVALHO, 2009).

O português, assim como o espanhol e o italiano, é rotulado como língua *[+] pro-drop*, por apresentar “todas” as propriedades que assim o denomina. Entretanto, estudos empíricos, sobre o português brasileiro, contradizem tal afirmativa, apresentando evidências de que essa variedade do português (PB) não possui todas as propriedades necessárias para que seja caracterizada como língua tipicamente de *sujeitos nulos*, revelando a preferência por formas pronominais expressas, demonstrando um comportamento diferente do português europeu. (DUARTE, 1995; 2000; 2003; GODOY, 1999; CYRINO, 1994; CYRINO, DUARTE e KATO, 2000).

Baseando-nos em estudos anteriores sobre a realização do *sujeito* no português brasileiro, buscamos descrever a realização do *sujeito* de referência arbitrária na fala carioca,

observando, com base em hipóteses linguísticas e extralinguísticas, se essa variedade sustenta o *status* de língua tipicamente de *sujeitos nulos* ou se, como apontado por Duarte (1995) e Duarte e Reis (2018), percorre um caminho de mudança em direção ao *sujeito preenchido*.

3.3 Sobre a pesquisa

3.3.1 Objetivo da pesquisa

Tendo em vista o extenso volume de pesquisas diacrônicas e sincrônicas (DUARTE 1993; 1995; 2000; CYRINO, 1994; LAPERUTA, 2002; CARVALHO, 2005), que confirmam que o português brasileiro já não é uma língua tipicamente [+] *pro-drop*, ou seja, uma língua que apresenta menos *sujeito nulos* – ao contrário do português europeu que, embora tenha casos de preenchimento do *sujeito* pronominal, ainda é marcado positivamente para o *Parâmetro do Sujeito Nulo* – nos propomos a, com esta pesquisa, fazer um levantamento do preenchimento do *sujeito pronominal* em concorrência com o *sujeito nulo* na fala dos cariocas, enfocando a atuação dos pronomes de referência arbitrária e sua correlação com a alteração do *Parâmetro do Sujeito Nulo*, a fim de verificar o comportamento desse fenômeno nessa variedade do português brasileiro.

3.3.2 Hipótese

Se a perda da propriedade [+] *pro-drop* do português brasileiro incide, inicialmente, em *sujeito de referência definida*, posteriormente, a preferência pelo preenchimento também recai sobre o *sujeito de referência arbitrária*. No português brasileiro, as construções pronominais nominativas superaram o uso da categoria vazia como estratégia para a indeterminação do *sujeito* (DUARTE, 2003). E são esses *sujeitos* de referência genérica o objeto de nossa investigação.

Duarte (1993) associa a crescente preferência pelo *sujeito* pronominal à simplificação no paradigma flexional. Para chegar a essa conclusão, investigou trechos de peças de teatro de cunho eminentemente popular, escritas nos séculos XIX e XX, que compreendem o período entre 1845 e 1992. Os resultados da pesquisa mostraram que a curva ascendente no número de *sujeito* preenchidos é perpendicular ao enfraquecimento da morfologia verbal.

Com base nesse pressuposto, acreditamos que um dos fatores mais relevantes no comportamento do preenchimento do *sujeito* pronominal de referência arbitrária seja a redução do paradigma flexional. Segundo Duarte (1993), contamos, atualmente, com um paradigma que apresenta quatro formas, graças à perda da segunda pessoa direta, bem como a introdução da forma inovadora “a gente”, que, por vezes, concorre com o pronome “nós”, que se combina com formas verbais de terceira pessoa do singular. Essa alteração no paradigma flexional proporciona alterações profundas na representação do *sujeito* pronominal, resultando no maior índice de preenchimento, o que indica uma possível mudança linguística.

3.3.3 O *corpus* da pesquisa

Para confirmar ou refutar a hipótese desta pesquisa nos baseamos em dados que provêm do *corpus* que pertence ao **Projeto Concordância**, disponível em (www.concordancia.ufrj.br), com entrevistas gravadas entre 2009 e 2010, estratificados em faixa etária, escolaridade e gênero. Esse *corpus* também foi objeto de análise de Duarte e Reis (2018). Nele, os pesquisadores investigam a realização do *sujeito* de referência definida em entrevistas que englobam falantes de duas localidades de Lisboa (Oeiras e Cácia) e do Rio de Janeiro (Copacabana e Nova Iguaçu). Conforme mencionado, usamos o mesmo *corpus*, só que nele, interessou-nos o *sujeito* de referência arbitrária e apenas a fala carioca (Copacabana). Seleccionamos para nossa análise apenas as entrevistas de Copacabana, pois, essas entrevistas, por si só, serão suficientes para observar o comportamento do *sujeito* de referência arbitrária na variedade carioca.

Analizamos todos os inquéritos de Copacabana, que contemplam os dois gêneros (Masculino e feminino), três faixas etárias (18 a 35 anos/ 36 a 55 anos/ 56 a 75 anos) e duas escolaridades (Ensino Fundamental e Ensino Superior), um total de 12 falantes. Para as análises multivariacionais, utilizamos o programa GoldVarb 2001 (ROBINSON, LAWRENCE & TAGLIAMONTE, 2001), cujos cálculos estatísticos e probabilísticos nos possibilitaram a análise quantitativa do fenômeno da indeterminação do *sujeito*.

3.4 Sobre o método

3.4.1 Suporte quantitativo

Para a obtenção dos resultados quantitativos, utilizamos como suporte o programa GoldVarb 2001 (ROBINSON, LAWRENCE & TAGLIAMONTE, 2001), que faz parte do pacote VarBrul. Esse programa possibilita análises multivariacionais e probabilísticas, realiza cálculos e fornece resultados quantitativos e percentuais, auxiliando o pesquisador na confirmação ou não de sua hipótese através de evidências.

Dentre os procedimentos realizados por esse programa, conforme Oliveira (2006, *In. Assunção, 2012*), está a análise combinatória dos grupos de fatores e variáveis, seguida de número de ocorrências, porcentagens e pesos relativos. Além disso, a influência relativa dos fatores sobre as variáveis é medida, atestando assim a influência de cada fator em relação às variáveis dependentes. De acordo, ainda, com a autora, considera-se significativo o fator ou variável com peso a partir de 0.05.

A fim de esclarecimento, na análise multivariada ou multidimensional, o programa nos apresenta vários níveis, não só com o cálculo do peso relativo dos fatores de cada uma das variáveis analisadas isoladamente em relação ao seu *input*, como também os cálculos de sua verossimilhança máxima (*log likelihood*) e de sua significância (*significance*). O programa expõe por meio do *step-up* os fatores com maior relevância e, pelo *step-down* elimina as variáveis independentes que não foram significativas.

Para a análise e utilização do GoldVarb, são necessários alguns passos como: construir a amostra da pesquisa; definir as variáveis dependentes, bem como os grupos de fatores; selecionar e codificar os dados; eliminar as ocorrências que não se enquadram nos critérios estabelecidos da variável dependente e preparar o arquivo de ocorrências. (GUY; ZILLES, 2007, *In. VITÓRIO, 2012*).

Dessa forma, o GoldVarb se define como um suporte quantitativo de pesquisa variacionista que além de quantificar as ocorrências, indica os possíveis condicionamentos das variáveis dependentes, contribuindo, então, com o trabalho do pesquisador em sua investigação.

3.4.2 Grupo de fatores

A sociolinguística tem como objeto de estudo a diversidade linguística, isto é, a língua falada em contextos sociais, ou seja, em situações reais de uso da língua. Nesse sentido, considera-se como ponto essencial de pesquisa a comunidade linguística, caracterizada como um grupo de falantes que compartilha a mesma variedade linguística.

Em uma pesquisa de cunho variacionista, é preciso levar em conta, numa comunidade linguística, alguns fatores linguísticos e sociais que, segundo Labov (1972, [2008]), são primordiais na observação e descrição de determinada variedade, pois interferem no modo como cada informante se comporta diante da sua língua. Posto isso, destacamos, aqui, os grupos de fatores que fazem parte da nossa pesquisa.

3.5 Fatores linguísticos

3.5.1 Tipo de ocorrência: *sujeito nulo* vs *sujeito pronominal*

Se considerarmos a tradição gramatical normativa, apregoadada para o bom uso do português brasileiro, veremos que esta variedade é marcada positivamente para o *sujeito nulo*, em outras palavras, a referência do *sujeito* nessa variedade encontra-se implícita na sentença, sendo considerada pela Teoria de Princípios e Parâmetros como língua [+] *pro-drop*. Essa teoria desenvolvida por Chomsky (1981, 1982) divide as línguas em dois grupos: as que exigem *sujeito* expreso, denominadas [-] *pro-drop*, e as que permitem uma categoria vazia nessa posição, consideradas como [+] *pro-drop*. (CHOMSKY, 1981. *Apud*. CARVALHO, 2013).

Diferentemente da classificação apresentada pela Teoria de Princípios e Parâmetros, os estudos variacionistas como de Duarte (1995, 2003), Duarte, Kato e Barbosa (2001), Carvalho (2009), Gravina (2014) mostram que o português brasileiro já não se enquadra mais como língua tipicamente [+] *pro-drop*. Esses estudos confirmam a afirmação feita por Duarte (1993) de que “[...] o português brasileiro convive com um sistema agonizante, em que ainda se refletem as características *pro-drop*, e um sistema em desenvolvimento, em que a riqueza funcional perdida já não permite a identificação de *pro*”. Nesse sentido, é válido mencionar que o português brasileiro usa de outras estratégias para preencher/indeterminar o sujeito.

Pautados na afirmação de que o português brasileiro já não é mais uma língua tipicamente [+] *pro-drop*, as nossas variáveis dependentes são:

SUJEITO NULO
SUJEITO PRONOMINAL

A seguir, trazemos exemplos de *sujeito nulo* e *sujeito pronominal*:

(01) **Documentador (Doravante D):** É muito ruim quando a pessoa não valoriza... deixa eu te falar esse bonde aqui tem pouco tempo né?

Informante (Doravante I): tem

D: como é que era assim antes de ter?

I: **A gente** subia muito de escada era de lá do Cantão até aqui em cima e aqui dá pra ver que só tem muita escada só lá na Laranjeiras que tem ladeira mas era muita escada mesmo... **a gente** sofria, Ø subia com bolsa, peso era muito. (E2-07)

3.5.2 A pessoa gramatical

Estudos diacrônicos e sincrônicos revelam a relevância do fator “pessoa gramatical” no preenchimento do *sujeito* (DUARTE, 1993; 1995; 2000; GODOY, 1999). Esses estudos mostram que a 1ª e 2ª pessoas, pelo caráter dêitico que possuem, favorecem o preenchimento do *sujeito*, enquanto a 3ª pessoa, pelo seu caráter anafórico, necessita de um antecedente que a identifique, portanto, maior resistência ao processo de mudança².

O comportamento resistente da 3ª pessoa para o processo de mudança em direção ao preenchimento do *sujeito* é estudado, especialmente, por Duarte (2008) e Duarte e Reis (2018). Nos referidos estudos, os pesquisadores confirmam o que já havia sido observado por Duarte (1993;1995), Godoy (1999), Carvalho (2005) e Assunção (2012): a 3ª pessoa (singular e plural) é a mais afetada pelo *sujeito nulo*.

É válido mencionar que a pesquisa de Duarte e Reis (2018) contempla o mesmo *corpus* que utilizamos em nossa pesquisa, porém, nele analisam somente os *sujeitos* de referência definida, com foco na 3ª pessoa. De acordo com os autores, a resistência da 3ª pessoa não se deve apenas ao seu caráter anafórico, mas também ao traço [+humano]. Em outras palavras, quanto mais humano/mais pessoa, mais preenchido será o *sujeito*, por outro lado, quanto menos humano/menos pessoa, menos preenchimento. E, de fato, isso é comprovado na pesquisa de Duarte e Reis (2018), que obtiveram maior percentual de preenchimento de *sujeito* de 1ª e 2ª pessoas, enquanto os de 3ª pessoa, mesmo tendo demonstrado certo avanço no processo de

² Para mais detalhes consultar seção 2, pois nela apresentamos um quadro-síntese, que ilustramos os fatores que mais influenciaram, bem como os que mais retardaram o preenchimento do *sujeito* de referência arbitrária em estudos de Duarte (1995), Godoy (1999), Carvalho (2005) e Assunção (2012).

mudança em direção ao *sujeito* expresso, ainda apresentaram índices superiores de não preenchimento.

Baseando-nos em estudos anteriores, acreditamos que, com o sujeito de referência arbitrária, a 1ª e 2ª pessoas pelos pronomes **a gente** e **você** exibirão altos índices de sujeitos preenchidos, mais esta que aquela.

Em nosso estudo, consideramos as seguintes pessoas gramaticais:

EU
TU
ELE
NÓS
VÓS ³
ELES
VOCÊ/VOCÊS
A GENTE

Exemplos:

(02): **Informante:** [...] mas não é uma crítica, mas um mau cristão na verdade... que acha que só ele tá certo... na verdade os... os fanáticos são um problema quando se mistura a religião com criação de valores na verdade ... você tem que buscar o que que há de mais ser humano, dar valores de ser humano pras pessoas, porque realmente elas vão pensar... “se **eu** grito mais do que **eu** devia na rua **eu** tou incomodando outra pessoa que... que recebe um som a mais de carros, de poluição etc. e tal... falta isso eu acho... essa criação de valor. (E01-02).

(03) **Informante:** Ah **tu** vai trabalhar mal, então vai pagar indenização. (E7-07)

(04) **D:** Em relação aos salários você acha que são justos? Esses que ganham bem são o quê? Os apadrinhados?

³ Em nosso *corpus* não foram registradas ocorrências com o pronome “vós”. Nesse sentido, não apresentamos nenhum exemplo com o referido pronome.

I: Normalmente são apadrinhados... lá tem muito a a/ além de apadrinhado tem muito o que a gente chama de panelinha não sei o quê então quem tá por cima hoje no poder puxa aquele pessoal todo pra trabalhar junto... e às vezes o cara vai pra uma área que **ele** não tem conhecimento nenhum, mas como ele é apadrinhado de alguém, **ele** passa lá quatro anos sem produzir nada sem fazer nada. (E9-05)

(05) **D:** E falando assim (perguntando) você falou do seu pai do seu avô... que que você pensa das famílias atuais? Configuração das famílias.

I: **Nós** estamos melhores do que os americanos porque a gente ainda tem família (...) **nós** tamos resgatando um pouco da questão da família. (E10-06)

(06) **D:** a senhora acha o povo brasileiro mais... mais aberto a receber... outras pessoas, de outros países? Aqui dentro do Brasil a senhora vê diferença? foi pra... pra Porto Alegre.

I: Ah sim por exemplo... foi Minas... Minas é uma coisa... Porto Alegre é outra né... eu acho que é diferente (...) o problema é receber... nesses lugares **eles** fecham até a porta quando a gente vai. (E6-12).

(07) **D:** e você acha que hoje em dia tem mais divórcios por que é mais comum assim?

I: Tem mais divórcio do que antes porque hoje em dia **você** pode ficar com a pessoa antes (...) quando você não experimenta isso antes a chance de dar errado é muito grande sem que seja culpa de ninguém. (E10-07).

(08) **D:** A senhora sente diferença entre as famílias católicas e os filhos que são criados em colégios católicos... quando a família está inserida na igreja tem diferença?

I: Quando a família é de fato... cristã né... eu acho que há uma diferença grande ... primeiro porque a mulher não se separa... né... quando conhece Deus... quando **a gente** não conhece **a gente** faz todas as coisas erradas. (E6-05).

3.5.3 O fator animacidade

Estudos variacionistas (PAREDES DA SILVA, 1991; DUARTE, 1993; 1995; 2008); CARVALHO, 2005) apontaram a relevância do fator animacidade na realização/omissão do

sujeito gramatical: o traço [+] animado do referente favorece o preenchimento do *sujeito*, enquanto o [-] animado privilegia construções com o *sujeito nulo*.

A escala de referencialidade proposta por Cyrino, Kato e Barbosa (2000) explicita a relevância dos traços: [+/- humano], [+/- animado], quer para o/ preenchimento/ a omissão do sujeito, quer para o apagamento/não-apagamento do objeto. Argumentos mais humanos se encontram no limite mais alto na hierarquia referencial, enquanto não-argumentos se encontram na posição mais baixa. O falante EU e o interlocutor VOCÊ, por serem inerentemente argumentos humanos, estão no ponto mais alto na hierarquia. A terceira pessoa se situa em um ponto mais baixo, devido à interação de traços [+/- humano] e [+/- específico]. "O *sujeito* que se refere a uma proposição (o *sujeito neutro*) está numa posição ainda mais baixa. No ponto mais baixo da hierarquia estão os *sujeitos* não referenciais". (CYRINO, DUARTE e KATO, 2000).

Para esse grupo de fatores, nossas expectativas estão depositadas em um maior número de *sujeitos preenchidos* quando o referente tiver traço [+ animado].

Segue exemplo:

- (09) I: (...) o problema é receber... nesses lugares **eles** fecham até a porta quando a gente vai. (E6-12).

No exemplo (09) temos um sujeito (+) animado, tendo em vista que se trata de uma pessoa, um ser genérico que representa uma ou mais pessoas que fecham suas portas para visitantes de outros lugares, ou seja, não são receptivos.

3.5.4 A concordância verbal

A concordância *sujeito-verbo* no português brasileiro tem demonstrado certo enfraquecimento, atingindo os limites da simplificação (DUARTE, 1993; LOPES; MACHADO, 2005). Nesse sentido, a introdução da expressão **a gente** e **você** no paradigma pronominal, combinada à mesma forma verbal que outras pessoas do discurso (2ª e 3ª pessoa do singular) resultou na perda da riqueza funcional, bem como no enfraquecimento da morfologia verbal e concordância (DUARTE, 1993). Para Galves (1993), a não concordância é aquela que "[...] não contém pessoa, ou contém pessoa como um traço puramente sintático".

Sendo assim, a questão do *sujeito nulo*, no português brasileiro, está mais relacionada a sua identificação, do que a sua legitimação.

Se o português brasileiro já não é uma língua de concordância forte, a nossa hipótese é: o uso de pronomes e/ou formas inovadoras será superior à categoria vazia, quer com marcas de concordância, quer sem elas, mais neste que naquele caso.

Observe os exemplos a seguir:

(10) **D:** Por que que você acha que a cidade tá tão violenta assim?

I: **a gente** sabe que onde tem muito turista eles vêm com dinheiro... que se o ladrão é um pouco esperto ele vem pra cá... mas acho que a violência tá de maneira geral [...] **a gente** vive numa desigualdade muito grande. (E7-02) – Concordância não se aplica verbal

(11) **D:** Que mais que você faz por aqui? Aqui é um bairro violento?

I: Alguns amigos já foram assaltados mas eu nunca, violentos etc e tal, assalto aqui é um dos lugares menos violentos pra isso [...] é preocupante porque **você não sabe** quando vai acontecer né pode acontecer em qualquer lugar. **Você pode** tá na praia e alguém chegar como aconteceu com aquele turista português de alguém chegar e te dar uma facada do nada. (E01-02) – Concordância não se aplica.

3.5.5 Tempo e modo verbal

O preenchimento ou apagamento do *sujeito*-pronome é também afetado pelo tempo e modo verbal. Duarte (1995), Godoy (1999), Carvalho (2005) e Assunção (2005) concluíram, em suas pesquisas, que o presente do indicativo privilegia o preenchimento do *sujeito* com estruturas pronominais, pelo fato de apresentar uma característica atemporal, como no exemplo a seguir:

(12) **D:** E você costuma ir ao cinema, teatro, o que que você gosta de fazer?

I: Eu costumo ir ao cinema e teatro... não aqui em Copacabana quer dizer cinema em Copacabana é meio difícil mas teatro tem alguns. [...] não frequento muito o Leme mas eu prefiro...toda vez que eu ando por aqui eu prefiro aqui do que Copacabana... aqui é mais bonito até... o Leme tem uma coisa ainda provinciana

assim parece que é um **você afasta** um pouco da cidade e é o único lugar também de praia de zona sul. **Você olha** pra algum lugar e ainda **você vê** mato assim mesmo, **você não vê** só prédio. (E01-01)

Duarte (1995) evidenciou que o preenchimento do *sujeito* em contextos indeterminados tem se expandido a outros tempos e modos verbais, como por exemplo o pretérito do indicativo como em:

(13) D: você lembra dos seus sonhos?

I: Sempre

D: Conta algum assim que tenha sido interessante.

I: Uma vez eu sonhei que tava andando na rua... que tinha tido uma explosão assim... em tudo... tinha várias pessoas tinham morrido... pelas ruas **você via** os prédios desmoronados etc e tal (E1-08)

Baseando-nos em pesquisas anteriores (DUARTE, 1995; GODOY, 1999; CARVALHO, 2005, ASSUNÇÃO, 2012) formulamos nossa hipótese: o presente do indicativo concentrará um maior número de sujeitos pronominais.

3.5.6 A correferencialidade

A correferencialidade é considerada como um fator de grande importância em um estudo sobre a realização do *sujeito*, pois, teoricamente, quando o referente é o mesmo, sua posição inibirá um *sujeito preenchido*, entretanto, se for diferente, o *sujeito* aparecerá preenchido (DUARTE, 1995; CARVALHO, 2005).

A seguir, trazemos um exemplo que ilustra a informação acima, evidenciando uma categoria vazia em contexto correferencial.

(14) **D:** É muito ruim quando a pessoa não valoriza... deixa eu te falar esse bonde aqui tem pouco tempo né. Como é que era assim antes de ter?

I: A gente subia muito de escada era de lá do Cantão até aqui em cima e aqui dá pra ver que só tem muita escada só lá na Laranjeiras que tem ladeira mas era muita escada mesmo. **A gente** sofria, **Ø subia** com bolsa peso era muito. (E2-07).

Carvalho (2005), em sua pesquisa, atestou que em relação ao fator correferencialidade, na fala araguiense, o maior número de *sujeitos preenchidos* se deu em contextos não correferentes, enquanto as estruturas correferentes privilegiaram o *sujeito nulo*. Entretanto, a estudiosa observou que os contextos não correferentes, embora mais resistentes, também já se mostram enfraquecidos.

Duarte e Reis (2018), que pesquisaram o mesmo *corpus* que aqui analisamos, também constataram, em sua pesquisa, que os contextos iniciais ainda favorecem o *sujeito* expresso, por outro lado, os contextos correferentes que, teoricamente, são resistentes ao preenchimento apresentaram índices de *sujeito* expresso superiores à categoria vazia. Isso ocorreu principalmente com a 2ª e 1ª pessoas, respectivamente. Entretanto, a 3ª pessoa demonstrou um caráter mais resistente ao preenchimento do *sujeito* tanto em contextos iniciais quanto correferentes.

Nesse sentido, podemos compreender que os contextos correferentes são resistentes ao *sujeito* expresso, enquanto as estruturas não-correferentes favorecem o preenchimento.

3.6 Fatores sociais

3.6.1 Sexo

Labov (1982, p. 78. Apud. LUCCHESI e ARAÚJO, 2004)) afirma que "[...] na maioria das mudanças linguísticas, as mulheres estão à frente dos homens na proporção de uma geração", nesse caso tendem a utilizar com mais frequência as formas inovadoras. Porém, alguns estudiosos como Chambers e Trudgill (1984) indicam um posicionamento inverso em relação ao fator sexo. Para os pesquisadores as mulheres apresentam maior incidência no uso de variantes de prestígio, enquanto os homens privilegiam estruturas inovadoras.

Duarte (1995) declara que são as mulheres que lideram a mudança rumo ao *sujeito* lexical. Já para Godoy (1999), há uma sutil diferença no preenchimento do *sujeito* por homens e mulheres, entretanto o sexo feminino apresentou em sua pesquisa um número um pouco maior de ocorrências com o *sujeito* expresso. Carvalho (2005) também constatou que o sexo feminino privilegia o preenchimento do *sujeito*, enquanto os homens ainda utilizam construções com o *sujeito nulo*. Em contrapartida a esses estudos, Assunção (2012) constatou a predominância do preenchimento *sujeito-pronome* na fala de informantes do sexo masculino.

Tendo em vista os estudos de Duarte (1995), Godoy (1999) e Carvalho (2005), podemos dizer que, em nosso estudo, as mulheres liderarão o quadro de preenchimento do *sujeito* com as formas inovadoras.

3.6.2 A faixa etária

Segundo Chambers e Trudgill (1984. *Apud.* LUCCHESI e ARAÚJO, 2004) os jovens lideram o processo de mudança, demonstrando preferência pelo uso de variantes inovadoras, enquanto os grupos intermediários tendem a utilizar as formas de prestígio.

Reforçam o estudo de Chambers e Trudgill (1984) as pesquisas de Duarte (1995), Godoy (1999), Carvalho (2005) e Assunção (2012), que pontam a preferência dos falantes mais jovens pelo uso das variantes inovadoras. Nesse sentido, nas pesquisas em questão, evidenciaram-se o uso de estratégias como “você” e “a gente” como indeterminação do *sujeito* na fala do grupo da primeira faixa etária, embora não seja expressiva a diferença para as demais faixas etárias.

Duarte e Reis (2018), em pesquisa cujo *corpus* também faz parte de nosso estudo, observaram um maior índice de preenchimento do *sujeito* na fala de informantes da primeira faixa etária, ou seja, os mais jovens, enquanto o grupo mais velho exibiu índices mais baixos, mantendo um equilíbrio entre *sujeitos nulos* e *expressos* de 3ª pessoa.

Considerando os estudos comentados, aqui, partimos do pressuposto de que os jovens exibirão, em nossa pesquisa, uma maior incidência de *sujeitos* preenchidos, pelo fato de ser o grupo que, geralmente, privilegia as construções inovadoras na linguagem.

3.6.3 A escolaridade

É sabido que a tradição escolar privilegia a variante de prestígio. Isso pressupõe que o falante com maior escolaridade, como a do ensino superior, por exemplo, apresente a preferência por formas que a gramática tradicional normativa prescreve, diferentemente do falante de menor escolaridade, que tende a utilizar as formas inovadoras. Para Labov (1982), o processo de mudança é liderado pelos falantes integrados à classe média baixa, pelo fato de ser o grupo que menos tem contato com a variante de prestígio.

Na pesquisa de Godoy (1999), o predomínio das variantes de prestígio, como recursos de indeterminação, ficou evidente na fala de informantes mais escolarizados, enquanto os recursos não legitimados pelas gramáticas tradicionais, tais como as formas pronominais,

abrangeram a escolaridade mais baixa. Em Carvalho (2005), constatou-se, na fala dos mais escolarizados, uma discreta preferência pelo *sujeito nulo*. Em contrapartida, Assunção (2012) atestou uma frequência maior no uso do pronome expresso em contextos indeterminados por falantes mais escolarizados, considerando especificamente, nesse caso, o uso do pronome “você”.

Tendo em vista estudos que indicam que falantes de maior escolaridade privilegiam o *sujeito nulo*, podemos dizer que, em nossa pesquisa, os informantes menos escolarizados apresentarão índices superiores de preenchimento pronominal.

3.7 Finalizando

Nesta seção, expusemos o *corpus*, objetivo, hipóteses, bem como todo o nosso aparato teórico-metodológico. Filiamos essa pesquisa à Teoria de Variação e Mudança, caracterizada pela sistematização da heterogeneidade linguística, que utiliza como metodologia a quantificação de dados, baseada em fatores linguísticos e sociais. Esse método permite evidenciar os fatores que exercem maior ou menor influência em um processo de variação e mudança. Nesse sentido, apresentamos sucintamente os fatores sociais e linguísticos selecionados como principais condicionadores seja para retardar ou acelerar o processo de mudança linguística.

Na próxima seção, apresentamos os resultados obtidos por meio da quantificação dos dados com o auxílio do programa GoldVarb 2001, analisamos e comparamos os resultados com os trabalhos explicitados na seção 2 desta pesquisa.

SEÇÃO IV ANÁLISE DO CORPUS

Apresentamos, nesta seção, a análise dos resultados sobre a realização do sujeito de referência arbitrária na fala carioca obtidos por meio da quantificação dos dados de nosso *corpus* com o auxílio do programa GoldVarb 2001 em relação a realização do sujeito de referência arbitrária na fala carioca. Após a sistematização dos dados, os resultados deste estudo serão comparados com as pesquisas de Duarte (1995), Godoy (1999), Carvalho (2005) e Assunção (2012), que tratam de diferentes variedades do português brasileiro, verificando um comportamento semelhante entre elas: o preenchimento pronominal do *sujeito*. Além disso, buscamos, quando possível, retomar alguns aspectos do estudo de Duarte e Reis (2018), que utilizaram o mesmo *corpus* que o nosso, porém, com foco no *sujeito definido*, em especial, na 3ª pessoa, a fim de, também, comparar resultados, entretanto, no nosso caso, entre *sujeitos definidos* e *arbitrários*.

4.1 O sujeito de referência arbitrária na fala carioca

4.1.1 Resultado geral

Para nosso estudo sobre o comportamento do sujeito de referência arbitrária na fala carioca, selecionamos duas escolaridades, sendo elas: Ensino Fundamental e Ensino Superior. Vale ressaltar, ainda, que todas as ocorrências com *sujeitos-pronomes arbitrários* foram contabilizadas, com exceção de sentenças repetidas (selecionamos apenas uma ocorrência para esses casos), bem como *sujeitos arbitrários* acompanhados de verbos no infinitivo ou imperativo, tendo em vista que esses modos verbais privilegiam a categoria vazia do sujeito, conforme Godoy (1999), embora tenhamos encontrado casos de preenchimento com esses modos verbais.

Seguem exemplos:

- (15) **Informante:** hoje em dia você pode entrar nas Casas Bahia e comprar uma máquina de lavar e pagar doze reais por mês durante muitos meses, mas **você consegue...você consegue** botar uma máquina de lavar dentro de casa (E10-06).

Sentença repetida

(16) **D:** você acha que você fala diferente dependendo da situação que você tá, onde, com quem?

I: Todo mundo ninguém fala a mesma... não tem como **você falar** do mesmo jeito com todo mundo (E1-06). Verbo no infinitivo com preenchimento pronominal

No cômputo geral dos dados desta pesquisa, foram registradas 817 ocorrências de sujeitos de referência arbitrária na fala carioca, sendo 681 (83%) de *sujeitos preenchidos* e 136 (17%) de *sujeitos nulos*.

Observe, na tabela 1, a seguir, a quantificação que ilustra o resultado geral:

Tabela 1- Quantificação e porcentagem dos *sujeitos de referência arbitrária* na fala carioca

Sujeito preenchido	Sujeito nulo	Total
681 (83%)	136 (17%)	817

Fonte: Elaborado pela autora.

Conforme tabela acima, podemos perceber a preferência dos falantes cariocas pelo uso do *sujeito pleno* com alto índice de preenchimento da casa do *sujeito* com pronomes, enquanto a categoria vazia alcançou baixos índices, mesmo estando nos referindo a um contexto arbitrário, que, teoricamente, privilegia o *sujeito nulo*. Tal resultado corrobora a afirmativa de autoras como Duarte (1995), Godoy (1999), Carvalho (2005), Assunção (2012), entre outras (os) que se dedicam a estudar sobre o tema, de que o português brasileiro já não carrega mais o *status* de língua tipicamente [+] *pro-drop*, pela grande incidência de preenchimento pronominal.

Vale destacar, ainda, que na pesquisa de Duarte e Reis (2018) em relação à realização dos *sujeitos de referência definida* na fala carioca, que utilizam o mesmo *corpus* que o nosso, foram computados 2.215 dados de 1ª, 2ª e 3ª pessoa, sendo 90% de *sujeitos preenchidos* em todas as faixas etárias. Assim, os autores concluíram que a mudança em direção ao *sujeito* preenchido de referência definida avança na fala carioca, além disso, afirmam que os resultados apresentados em sua pesquisa mostram que a Teoria de Variação e Mudança associada à Teoria de Princípios e Parâmetros vai além de uma sociolinguística de probabilidades, e servem para incluir o português brasileiro como língua de *sujeitos nulos* parciais.

Nesse sentido, compartilhamos da tese de Duarte e Reis (2018) de que o *sujeito* preenchido avança na fala carioca, em nosso estudo a comprovação se dá em relação a

referência arbitrária, com altos índices de *sujeito preenchido* com formas pronominais expressas, até mesmo em contextos considerados resistentes a esse preenchimento, como por exemplo os correferentes, como veremos mais adiante neste trabalho.

Cabe esclarecer, aqui, que para obtermos os pesos relativos aos fatores selecionados para análise, foi necessário acrescentar seis ocorrências fictícias, em nosso *corpus*, para os fatores que tiveram *knockouts*⁴ a fim de que o programa rodasse. Ressaltamos que esse procedimento ocorreu somente para os pesos relativos, foram adicionadas somente seis ocorrências, uma para cada fator que não obteve variação. Nesse sentido, podemos dizer que tal processo não interferiu na análise probabilística dos fatores, uma vez que tais resultados apresentaram-se absolutos. Desse modo, acrescentar ocorrências falsas ou excluir esses fatores não altera os nossos resultados, essa é apenas uma forma de fazer rodar o programa para verificarmos a significância dos fatores, nesse sentido optamos por não excluí-los. Houve *Knockouts* para o fator “ [-] animado”, para os pronomes “vocês”, “tu” e “vós” com os verbos no futuro do indicativo, passado do subjuntivo e futuro do subjuntivo.

4.1.2 Fatores selecionados pelo programa

4.1.2.1 Cruzamentos e pesos relativos

No processamento da análise da rodada *step up /step down*, o programa GoldVarb 2001 revelou que o melhor *step up* foi o 30 (*Input* 0,935, *Log likelihood* = -219,502, *Significance* = 0,046). Nessa mesma rodada, o programa selecionou os seguintes grupos de fatores em ordem de relevância: (1º) correferencialidade, (2º) pessoa gramatical, (3º) escolaridade, (4º) faixa etária), (5º) tempo e modo verbal.

4.1.2.1.1 Sujeito genérico e a correferencialidade

Na análise probabilística, a correferencialidade foi apontada como o fator de maior significância na influência das variáveis *sujeito preenchido* e *sujeito nulo*. Embora as estruturas não correferentes tenham maior incidência de *sujeitos pronominais* como esperávamos, com

⁴ *Knockout* é o termo usado, no GoldVarb, para indicar que um determinado resultado é absoluto, ou seja, não houve variação. Para que o programa prossiga na execução das rodadas estatísticas, é necessário que se retire da análise o fator com o qual se dá o *knockout* ou que se adicione uma ocorrência para aquele fator que não obteve variação.

94% de ocorrências e peso relativo 0.747, as estruturas correferenciais também apresentaram grandes índices de preenchimento (62%), com peso 0.112.

Ainda que as ocorrências com preenchimento pronominal se sobressaíam em contextos com *sujeitos* diferentes, o que nos chamou a atenção foi o avanço do *sujeito* pleno em estruturas correferentes que, segundo Duarte (1995) e Carvalho (2005), resistem ao preenchimento, reforçando, desse modo, a perspectiva de que o português brasileiro tem a tendência de preencher a casa do *sujeito*. Vejamos a tabela 2:

Tabela 2 - Quantificação e porcentagem do cruzamento dos fatores tipo de *sujeito* e correferencialidade

Correferência	Sujeito preenchido	Sujeito nulo	Total
Mesmo sujeito	172 (62%)	105 (38%)	277
Sujeito diferente	509 (94%)	31 (6%)	540
Total	681	136	817

Fonte: Elaborado pela autora.

Carvalho (2005) também selecionou para análise do seu *corpus* de pesquisa o fator correferencialidade em relação à fala araguiense. Com base na tabela 2, nosso trabalho corrobora a afirmação da pesquisadora de que as estruturas correferentes resistem ao *sujeito lexical*, enquanto a não correferencialidade o favorece.

O preenchimento pronominal em estruturas com o sujeito diferente foi de 94% em nossa pesquisa, mostrando-se favorável ao *sujeito-pronome*. As estruturas correferentes também apresentaram preenchimento do *sujeito* com 62% de ocorrências, um número significativo se levarmos em conta que a correferencialidade privilegia a categoria vazia. Carvalho (2005) também registrou em sua pesquisa um percentual bastante expressivo, 40% de ocorrências com sujeito arbitrário em estruturas com o mesmo *sujeito*.

Vale a pena ressaltar que na pesquisa de Duarte e Reis (2018), que analisaram o mesmo *corpus* desta pesquisa, os autores concluíram que contextos iniciais ainda favorecem o *sujeito expreso*, por outro lado, os contextos correferentes que, teoricamente, são resistentes ao preenchimento apresentaram índices de *sujeito expreso* superiores à categoria vazia. Isso ocorreu, principalmente, com a 2ª e 1ª pessoas, respectivamente. Entretanto, a 3ª pessoa demonstrou um caráter mais resistente ao preenchimento do *sujeito* tanto em contextos iniciais quanto correferentes. Para análise da 3ª pessoa, Duarte e Reis (2018) selecionaram o fator *padrão sentencial*. “Os padrões sentenciais levam em consideração a função do antecedente e

sua posição no mesmo período em que o elemento é retomado ou em um período adjacente”. (DUARTE e REIS, 2018). Tendo em vista que esse fator corresponde às estruturas correferentes, apresentamos, aqui, aqueles que tiveram maior significância na perda do licenciamento e identificação do *sujeito nulo*, na pesquisa dos autores:

a) Padrão E, com o antecedente em uma subordinada anteposta à matriz:

(17) Se o ladrão é um pouco esperto, ele vem pra cá. (H, 1, S)

b) Padrão D, com o antecedente distante:

(18) O meu filho tava chegando em casa -que nós trabalhamos com festa, como eu te falei, né -e ele tinha ido comprar bolas. Aí, não tinha as bolas que nós queríamos. Ele trouxe o dinheiro. Ele era bem mais novinho. Ø Trouxe o dinheiro no bolso. (M, 2, S)

c) Padrão C, com o antecedente em outra função:

(19) O povo aqui de baixo tinha ódio do Brizola porque ele fez isso nas favelas. (H, 2, F).

Duarte e Reis (2018) atestaram que os padrões B e A favorecem o *sujeito nulo*. Vejamos alguns exemplos:

d) No padrão B, o antecedente é sujeito e se encontra em um período adjacente:

(20) Mas, o garoto ficou. Ø tinha saído do berço dele. (H, 3, F)

e) No padrão A, o sujeito da encaixada posposta é correferente com o sujeito da oração matriz:

(21) Então, o que era comédia assim era o próprio amigo deles que faziam isso. Aí, **ele** fingia que Ø matava. (M, 1, F).

Esses pesquisadores (2018) concluíram, que embora o PB ainda exiba *sujeitos nulos* no padrão B, e, também, apresente a categoria vazia no padrão A, que, conforme as autoras, corresponde a contextos para as línguas de *sujeitos nulos* parciais, o *sujeito nulo* encontra-se em processo de extinção e concorrem com o *sujeito expresso*.

Dessa forma, assim como Duarte e Reis (2018) perceberam o enfraquecimento do parâmetro [+] *pro-drop* para o português brasileiro, também observamos na nossa pesquisa. Os

números quanto ao fator correferencialidade sugerem que essa mudança em direção ao *sujeito preenchido* caminha à passos largos, revelando o crescimento do preenchimento pronominal até mesmo em contextos correferentes.

Trazemos alguns exemplos para mostrar as ocorrências:

(22) **I:** Às vezes a gente pega assim por exemplo cinquenta reais e pede pro menino que tem aqui pra ir comprar umas coisas, compra o básico... pra gente poder doar pra essa pessoa pra ter pelo menos o que comer durante uma semana... **você** não consegue comprar o básico, **você** compra ali um quilo de açúcar, pó de café, pão de forma algumas coisa, **você** não pode dar um queijo. (E3-08). (Estrutura correferente com sujeito preenchido).

(23) **D:** E a chuva?

I: Quando chove aqui **a gente** tem medo eu já perdi... eu não gosto nem de falar essa parte mas você já que você perguntou eu perdi um filho subterrado [...] lá no canto lá na frente hoje é até um prédio... agora hoje em dia não tem esse risco mais porque **eles** fizeram né aquela barragem lá de concreto (E5-06). (Estrutura não correferente).

(24) Nada que surpreenda que você fale “a Ø preciso ver o Sérgio Cabral falando ou também não é uma coisa tipo ah se ele tá falando “ah não Ø vou trocar Ø não aguento mais ver Sérgio Cabral falando” (E1-05). (Estrutura correferente com sujeito nulo).

4.1.2.1.2 Tipo de sujeito x pessoa gramatical

O fator pessoa gramatical é apresentado pelo programa GoldVarb 2001 (peso relativo) como o segundo mais relevante na influência das variáveis dependentes “*sujeito preenchido*” e “*sujeito nulo*”. Posto isto, apresentamos, a seguir, na tabela 3, a quantificação e porcentagem do cruzamento dos fatores *tipo de sujeito* e *pessoa gramatical*, de modo a demonstrarmos os pronomes com maior produtividade no preenchimento do *sujeito de referência arbitrária* na fala carioca, bem como os que mais favorecem a categoria vazia. Vale ressaltar que o *sujeito*

preenchido se mostrou mais atuante do que a categoria vazia, confirmando, então, a nossa hipótese de pesquisa de que os *sujeitos pronominais* superariam os *sujeitos nulos*.

Tabela 3- Quantificação e porcentagem do cruzamento entre os fatores tipo de sujeito x pessoa gramatical

Pessoal gramatical	Sujeito preenchido	Sujeito nulo	Total
Eu	36 (77%)	11 (23%)	47
Tu	1 (50%)	1 (50%)	2
Você	375 (91%)	36 (9%)	411
Ele	7 (70%)	3 (30%)	10
Nós	28 (70%)	12 (30%)	40
A gente	202 (91%)	19 (9%)	221
Vocês	2 (100%)	-	2
Eles	30 (36%)	54 (64%)	84
Total	681	136	817

Fonte: Elaborado pela autora

Como esperávamos, o uso dos pronomes “você” e “a gente” foram as estratégias mais utilizadas, as duas tiveram o mesmo percentual de ocorrências (91%), demonstrando a mesma força no preenchimento do *sujeito*. O pronome “você” obteve peso relativo 0,634 de significância, enquanto “a gente” apresentou significância de 0,595. O que nos chamou a atenção nos números dessa tabela, foram, justamente, as ocorrências com o pronome “eu”, com um percentual de 77% e significância de 0,398, uma vez ser o mais alto na hierarquia referencial.

A primeira e a segunda estratégia de indeterminação do sujeito que se encontram no topo das preferências dos cariocas, demonstrado no nosso estudo, foram os termos: “você” e “a gente”, que, igualmente, se sobressaíram nas pesquisas de Duarte (1995), Godoy (1999), Carvalho (2005) e Assunção (2012), confirmando a afirmativa de que o enfraquecimento do parâmetro do *sujeito nulo* atinge, em primeiro lugar, os pronomes de 1ª e 2ª pessoa, enquanto

os pronomes de 3ª pessoa são menos afetados, privilegiando, no último caso, o *sujeito nulo*. Em nossa pesquisa, o pronome de 3ª pessoa do plural, mostrou-se favorável ao *sujeito nulo*.

O caráter resistente ao preenchimento da 3ª pessoa foi estudado por Duarte e Reis (2018), que se propuseram a analisar os *sujeitos de referência definida* na fala carioca. E ao compararmos os resultados apresentados na tabela 2 com aqueles obtidos pelos autores, constatamos que os pronomes afetados, pelo enfraquecimento do parâmetro [+] *pro-drop* na referência definida, são os mesmos que na referência arbitrária: a 2ª pessoa demonstrou ser a mais atingida pelo preenchimento do *sujeito* na pesquisa de Duarte e Reis (90%) e em nosso trabalho com o percentual de 91%, indicando um quadro de mudança quase que totalmente concluído para a 2ª pessoa; a 1ª pessoa obteve um percentual de 84% para os mais jovens e 78% para os falantes de terceira faixa etária na pesquisa de Duarte e Reis (2018), enquanto em nosso *corpus* foram 91% de *sujeitos preenchidos* para a 1ª pessoa (a gente), seguindo a mesma velocidade no processo de mudança que a 2ª pessoa; a 3ª pessoa do plural de referência definida na fala carioca alcançou índices entre 69% para a terceira faixa etária e 75% para a segunda, na pesquisa dos referidos autores, já em nosso trabalho em relação ao *sujeito de referência arbitrária* o resultado despencou para 36% de preenchimento, enquanto a categoria vazia alcançou 64% das ocorrências, mostrando-se mais produtiva do que o *sujeito* pleno em sentenças com a 3ª pessoa em contexto arbitrário.

A seguir apresentamos alguns exemplos ilustrativos dessas ocorrências:

(25) **I:** [...] mas não é uma crítica, mas um mau cristão na verdade... que acha que só ele tá certo... na verdade os... os fanáticos são um problema quando se mistura a religião com criação de valores na verdade ... você tem que buscar o que que há de mais ser humano, dar valores de ser humano pras pessoas, porque realmente elas vão pensar... “se **eu** grito mais do que **eu** devia na rua **eu** tou incomodando outra pessoa que... que recebe um som a mais de carros, de poluição etc e tal... falta isso eu acho... essa criação de valor. (E01-02).

(26) **D:** Que efeito você pode me apontar assim de ter tanta favela ... na zona sul? ... você acha que gera problemas?

I: A desigualdade social ela ... o confronto é mais intenso assim no sentido de que... a convivência... de quem tem muito dinheiro quem tem pouco dinheiro ela... é constante na praia de/ do Arpoador tem ali... todos os moradores do... do Galo do Pavão Pavãozinho... junto com o pessoal que mora em Ipanema que é... né... o

que tem de mais alto assim... então... é o meu namorado estuda na PUC... e ele... eh... já contou várias vezes assim de tá no ônibus... e aí... o pessoal na rua fala “aí playboy... **vocês** fumam maconha aqui do morro né?” assim... tem uma hostilidade... que... que ela acaba sendo... eh explicitada mesmo assim... dá pra trocar... eh acho que... **você** sai da sua casa com suas condições precárias e Ø desce em Ipanema assim deve ser chocante. (E8-03).

(27) **D:** E você gosta de sonhar? uhn uhn... gosto acho que quando **a gente** evolui no pensamento do sonho que a gente tá evoluindo como ser humano. (E1-08)

(28) **D:** em relação ao comércio você acha que é bom... em comparação aos outros bairros?

I: Olha eu acho que ele é muito bom... eu não tenho nem o que reclamar tanto é que Ø **abriram** já três shoppings aqui, só em Botafogo que não existia nenhum ... atualmente tem três shoppings tem muito comércio pelas ruas, lojas novas abrindo [...]. (E09-02).

4.1.2.1.3 Tipo de sujeito x escolaridade

A escolaridade dos informantes da pesquisa foi apontada, pela análise probabilística, como o terceiro fator mais relevante na influência do preenchimento pronominal. O ensino fundamental obteve peso relativo de 0,636, demonstrando maior influência no preenchimento, enquanto o ensino superior apresentou 0,426 de significância. Quanto aos resultados obtidos, por meio dos cruzamentos, sistematizamos os dados apresentados na tabela 4, como segue:

Tabela 4 - Quantificação e porcentagem dos fatores tipo de sujeito x escolaridade

	Sujeito preenchido	Sujeito nulo	Total
Ensino Fundamental	251 (88%)	35 (12%)	286
Ensino Superior	430 (81%)	101 (19%)	531
Total	681	136	817

Fonte: Elaborado pela autora

Como podemos ver na tabela 4, os resultados revelam que os falantes do ensino fundamental utilizaram com maior frequência o *sujeito pronominal* como *sujeito genérico* do

que os mais escolarizados, e esses números confirmam a nossa hipótese. O ensino fundamental apresentou 88% das ocorrências, enquanto o ensino superior teve 81% de casos de *sujeitos* expressos.

Dessa forma, observamos que os falantes cariocas menos escolarizados privilegiaram as novas estratégias de indeterminação. Entretanto, cabe frisar, aqui, que o número de preenchimento com *sujeito-pronome*, no ensino superior, foi bastante expressivo, por se tratar de uma escolaridade que, a princípio, esperava-se que privilegiaria o *sujeito nulo*, mas, que também demonstrou preferência pelo *sujeito expresso*.

Esses resultados vão ao encontro da pesquisa de Godoy (1999), que mostraram o predomínio das variantes de prestígio como recursos de indeterminação na fala de informantes mais escolarizados, enquanto os recursos não legitimados pelas gramáticas tradicionais, tais como as formas pronominais, abrangeram a escolaridade mais baixa. Carvalho (2005) observou, na fala dos mais escolarizados, uma discreta preferência pelo *sujeito nulo*. Em contrapartida, Assunção (2012), atestou uma frequência maior no uso do pronome expresso em contextos indeterminados, por falantes mais escolarizados, considerando, especificamente nesse caso, o uso do pronome “você”.

4.1.2.1.4 Tipo de sujeito x faixa etária

Na análise probabilística, a faixa etária foi apontada como o quarto fator mais significativo na influência quantitativa das variáveis dependentes. Observe a tabela 5 abaixo:

Tabela 5 - Quantificação e porcentagem do cruzamento dos fatores tipo de sujeito x faixa etária

	Sujeito preenchido	Sujeito nulo	Total
1ª faixa etária	261 (84%)	51 (16%)	312
2ª faixa etária	242 (93%)	18 (7%)	260
3ª faixa etária	178 (73%)	67 (27%)	245
Total	681 (83%)	136 (17%)	817

Fonte: Elaborado pela autora.

No cruzamento dos fatores tipo de *sujeito* e faixa etária, podemos perceber que o preenchimento do *sujeito* com pronomes plenos encontra-se no topo das preferências para os

falantes cariocas de segunda faixa etária, com 242 ocorrências (93%) e peso relativo de 0,700. A primeira faixa etária também demonstrou privilegiar as estratégias inovadoras, com 261 ocorrências (84%) e peso relativo de 0.389. Em relação à terceira faixa etária, foram computados 178 casos de preenchimento (73%) e peso de significância de 0,422. Mesmo que a terceira faixa etária tenha apresentado alto índice de preenchimento pronominal, ainda assim é a mais resistente ao preenchimento, com uma porcentagem maior de *sujeitos nulos* (27%).

Os resultados em relação à faixa etária confirmaram que os dois grupos mais jovens utilizam as formas inovadoras com mais frequência do que o grupo de falantes mais velhos. Fato também comprovado nas pesquisas de Duarte (1995), Godoy (1999), Carvalho (2005) e Assunção (2012), que verificaram a predominância de *sujeitos preenchidos* com formas pronominais expressas para a primeira faixa etária, e maior resistência para a terceira faixa etária.

Na pesquisa de Duarte e Reis (2018), em relação ao sujeito de referência definida, ao levar em conta o preenchimento do *sujeito*, a pessoa gramatical e a faixa etária obtiveram os seguintes resultados: a 2ª pessoa apresentou o mesmo percentual em todas as faixas etárias (90%); a 1ª pessoa alcançou índices de 84% para a primeira faixa etária, 77% para a segunda e 78% para a terceira faixa de idade; a 3ª pessoa apresentou 73% de preenchimento para a primeira faixa etária, 75% para a segunda e 69% para a terceira faixa etária. Nesse sentido, a terceira faixa etária foi a mais resistente ao preenchimento do *sujeito*, quando se trata da 3ª pessoa.

4.1.2.1.5 Tipo de sujeito x tempo e modo verbal

O fator tempo e modo verbal foi apontado pela análise probabilística dos fatores como o quinto com maior significância. Nesse sentido, apresentamos, na tabela 6, a quantificação e porcentagem do cruzamento entre os fatores tipo de *sujeito* e tempo e modo verbal. Vale ressaltar que dispomos na tabela somente os resultados em relação aos tempos e modos verbais que se sobressaíram na pesquisa, sendo eles: o presente do indicativo e o passado do indicativo. Os demais tempos verbais obtiveram um número muito reduzido de ocorrências, isso quando não houve *knockout*, portanto, não sendo relevante comentá-los.

Tabela 6 - Quantificação e porcentagem do cruzamento dos fatores tipo de sujeito x tempo e modo verbal

	Sujeito preenchido	Sujeito nulo	Total
Presente do indicativo	542 (87%)	79 (13%)	621
Passado do indicativo	121 (68%)	56 (32%)	177
Total	663	135	798

Fonte: Elaborado pela autora.

Conforme tabela 6, o presente do indicativo mostrou-se mais atuante no preenchimento do *sujeito* em sentenças de referência arbitrária, com 542 ocorrências (87%) e peso relativo de 0,554, enquanto a categoria vazia apresentou 79 casos (13%). Esses resultados corroboram os resultados das pesquisas de Duarte (1995), Godoy (1999), Carvalho (2005) e Assunção (2012), que são unânimes quanto ao favorecimento do presente do indicativo no preenchimento do *sujeito*, pelo fato de apresentar uma característica atemporal, além de ser utilizado para indicar fatos habituais. (MENON, 1994).

Em consonância com Duarte (1995), observamos que o preenchimento *sujeito-pronome* tem se estendido a outros tempos e modos verbais, dentre os quais destaca-se o passado do indicativo. Em nossa pesquisa registramos 121 ocorrências (68%) e peso 0,334 para esse tempo verbal, enquanto a categoria vazia apresentou 56 casos (32%). Esses resultados vão ao encontro daqueles alcançados por Duarte (1995) em sua pesquisa “A perda do princípio ‘Evite Pronome’ no português brasileiro”, na qual também analisou a fala carioca e, Assunção (2012), em seu trabalho sobre a variedade de Feira de Santana.

Os exemplos 29 e 30 mostram essas ocorrências:

(29) D: eh agora tá chegando as UPPs né

I: Mas a UPP:... a UPP... é uma fachada, pura fachada... essa semana mesmo eles agrediram um sobrinho meu [...] **você não vê**, mas a droga continua sendo vendida, o tráfico continua o mesmo, os traficantes continuam lá em cima. (E3-04). (Presente do indicativo).

(30) D: Então... o quê que a senhora acha aqui do bairro... o quê que tem de positivo aqui o quê que é bom?

I: Bem hoje em dia eu não, não acho tanto como eu achava antigamente né [...] até cinema que eu vou eu fico triste porque eu só vejo.... é violentando a mulher... violentando crianças né... coisa que é uma realidade [...] na minha época **a gente não ouvia** nem falar nisso talvez houvesse mas não era com tanta frequência. (E6-02). (Pretérito imperfeito do indicativo)

4.2 Fatores excluídos pelo programa GoldVarb 2001 (peso relativo)

Os fatores linguísticos *animacidade* e *concordância verbal*, bem como o fator extralinguístico *gênero* não foram selecionados pelo programa GoldVarb, por meio do peso relativo, entretanto optamos por comentá-los, tendo em vista as nossas hipóteses de pesquisa.

4.2.1 A animacidade do sujeito

Como já mencionado anteriormente o fator animacidade foi excluído pelo programa GoldVarb 2001, por meio do procedimento peso relativo, em razão de não ter apresentado significância na influência das variáveis dependentes. Contudo, consideramos pertinente traçarmos aqui alguns comentários sobre essa variável. Observe a tabela 7, a seguir:

Tabela 7- Quantificação e porcentagem do sujeito de referência arbitrária fator mais animado e menos animado

Animacidade do sujeito	Sujeito preenchido	Sujeito nulo	Total
Mais animado	681 (83%)	136 (17%)	817
Menos animado	-	-	-
Total	681	136	817

Fonte: Elaborado pela autora.

A tabela 7 ilustra o número de ocorrências e porcentagens da animacidade do *sujeito* que, segundo Duarte e Reis (2018), é relevante para o seu preenchimento. Argumentos mais humanos se encontram no limite mais alto na hierarquia referencial, enquanto não-argumentos se encontram na posição mais baixa. Vale, então, dizer que quanto mais humano, mais específico e mais animado for o argumento, mais susceptível ao preenchimento será o *sujeito*,

o que explica o baixo índice de *sujeitos* preenchidos com pronomes de 3ª pessoa do plural em nosso trabalho, dado o seu caráter [+/-] humano, [+/-] animado.

A animacidade foi considerada como um fator relevante na pesquisa de Duarte e Reis (2018) sobre o *sujeito* de referência definida na fala carioca, e os resultados em relação a 3ª pessoa só confirmam o que observamos em nosso trabalho, “[...] os referentes que apresentam o traço [+hum/+/-esp] favorecem o preenchimento, com pesos de 0,572 e 0,441, respectivamente. Aqueles com o traço [-hum] (aqui incluindo apenas os [-anim]) se situam no extremo oposto, com pesos de 0,282 e 0,114”. (DUARTE e REIS, 2018). Desse modo, percebemos que *sujeitos* [+] animados contribuem para o *sujeito* pleno.

Vejamos alguns exemplos:

- (31) Aí, **a minha tia** também tocava. **Ela** tinha acordeão. (H, 3, F). (Exemplo ilustrativo sobre o *sujeito* definido retirado do trabalho de Duarte e Reis (2018)).

Segue exemplo:

- (32) [...] as causas que são até vinte salários mínimos **você** não precisa de advogado... então às vezes as pessoas entram sozinhas escrevem à mão porque o negócio é muito informal mesmo... **você** pode escrever à mão e dar entrada... e os advogados não falam isso pras pessoas que até vinte salários mínimos... porque eles perdem honorário. (E07-08).

4.2.2 Tipo de *sujeito* x concordância verbal

Embora o fator concordância tenha sido excluída pelo programa GoldVarb 2001 por meio do procedimento peso relativo, consideramos relevante apresentar os resultados obtidos por meio do cruzamento dos fatores tipo de *sujeito* e concordância verbal, tendo em vista que se tratava de uma de nossas hipóteses de pesquisa.

Tabela 8 - Quantificação e porcentagem do cruzamento do fator tipo de *sujeito* x concordância verbal

	Sujeito preenchido	Sujeito nulo	Total
Concordância	59 (46%)	69 (54%)	128
Não se aplica ⁵	622 (90%)	67 (10%)	689
Total	681	136	817

Fonte: Elaborado pela autora.

Ao analisarmos a tabela acima, podemos perceber que os casos em que a regra de concordância não se aplica foram mais correntes no uso de estratégias de indeterminação do *sujeito* com formas pronominais expressas (90%). A ausência de flexão incide favoravelmente ao *sujeito preenchido*, confirmando a tese de Duarte (1995). Outro item interessante na tabela a se observar, quando o *sujeito* é nulo a marca da morfologia é mais expressiva (54%). Tais resultados eram esperados, uma vez que os pronomes “você” e “a gente” foram as estratégias mais utilizadas para a indeterminação do *sujeito*.

4.2.3 Tipo de *sujeito* x sexo

O fator sexo também foi excluído pelo programa GoldVarb (peso relativo), porém, apresentamos, a seguir, o resultado proporcionado pelo cruzamento tipo de *sujeito* x sexo, tabela 9, com o objetivo de verificar o gênero que mais apresentou preenchimento pronominal.

Tabela 9 - Quantificação e porcentagem do cruzamento dos fatores tipo de *sujeito* x sexo

	Sujeito preenchido	Sujeito nulo	Total
Homem	453 (81%)	106 (19%)	559
Mulher	228 (88%)	30 (12%)	258
Total	681	136	817

Fonte: Elaborado pela autora.

⁵ Só consideramos a não concordância nos casos em que há marcas explícitas: “tu”, “nós”, “vós”, “eles”, “elas”, “vocês”.

No cruzamento dos fatores tipo de *sujeito* e sexo, podemos perceber que o preenchimento do *sujeito* com pronomes plenos lidera tanto no sexo feminino, quanto no masculino. As mulheres foram as que mais preencheram a casa do *sujeito* com 88% de ocorrências com *sujeito* pronominal, enquanto para os homens registramos 81%, uma amostra bastante equilibrada entre os sexos, com uma ligeira diferença.

Nossos resultados, em relação ao preenchimento ou não do *sujeito* com o fator sexo, vão ao encontro dos estudos de Duarte (1995), Godoy (1999) e Carvalho (2005), que observaram as mulheres à frente no processo de mudança, em direção ao *sujeito* lexical. No entanto, observamos também, em nosso estudo, que esse quadro atinge quase que igualmente os dois sexos.

4.2.4 Tipo de sujeito x pessoa gramatical x sexo

Sistematizamos, na tabela 10, os dados em relação ao tipo de *sujeito*, pessoa gramatical e sexo.

Conforme mostra a tabela 10, podemos observar que o pronome “você” foi o mais corrente no preenchimento do *sujeito* de referência arbitrária tanto na fala dos homens (253-92%), quanto na fala das mulheres (122-90%), com uma sutil diferença de 2%. Logo após, o pronome “a gente” teve grande produtividade, com 121 ocorrências (90%) registradas para o sexo masculino e 81 (94%) para o sexo feminino, também com uma pequena diferença na porcentagem de ocorrências, demonstrando certo equilíbrio no uso das estratégias de indeterminação entre os gêneros.

O pronome “eles” obteve o maior percentual de *sujeito* nulo para o sexo masculino (48 - 67%), mostrando-se favorável à categoria vazia. Esse recurso é prescrito pelos compêndios gramaticais como uma das principais estratégias de indeterminação do *sujeito*. Apesar disso, também registramos casos de preenchimento com esse pronome. Nesse sentido, percebemos que novas estratégias de indeterminação do *sujeito* estão sendo utilizadas pelos falantes cariocas, além daquelas reconhecidas pelas gramáticas tradicionais.

Tabela 10 - Quantificação e porcentagem do cruzamento dos fatores tipo de sujeito x pessoa gramatical x sexo

	Sujeito preenchido	Sujeito nulo	Sujeito preenchido	Sujeito nulo	Total
	Homem		Mulher		
Eu	23 (74%)	8 (26%)	13 (81%)	3 (19%)	47
Tu	1 (50%)	1 (50%)	-	-	2
Você	253 (92%)	23 (8%)	122 (90%)	13 (10%)	411
Ele/Ela	7 (70%)	3 (30%)	-	-	10
A gente	121 (90%)	14 (10%)	81 (94%)	5 (6%)	221
Nós	23 (72%)	9 (28%)	5 (62%)	3 (38%)	40
Vocês	1 (100%)	-	1 (100%)	-	2
Eles/Elas	24 (33%)	48 (67%)	6 (50%)	6 (50%)	84
Total	453	106	228	30	817

Fonte: Elaborado pela autora.

4.4 Finalizando

Apresentamos, nesta seção, os resultados levantados por meio do programa GoldVarb 2001 sobre a variação das dependentes “sujeito preenchido x sujeito nulo” em contextos arbitrários, na variedade carioca, seguida de grupos de fatores e seus cruzamentos, de modo a verificar a realização do sujeito na fala de moradores de Copacabana (RJ).

Sendo assim, traçamos, a seguir, um quadro-síntese com algumas informações pertinentes sobre os resultados encontrados:

Quadro-síntese 3

SUJEITO DE REFERÊNCIA ARBITRÁRIA NA VARIEDADE CARIOCA
Correferencialidade – As sentenças não correferentes foram favoráveis ao preenchimento do <i>sujeito</i>. Entretanto, as estruturas correferentes (mesmo sujeito) também apresentaram formas pronominais expressas do sujeito, o que mostra certo enfraquecimento do parâmetro do <i>sujeito</i> nulo.
Pessoa gramatical – Os pronomes “você” e “a gente” foram mais produtivos no preenchimento do <i>sujeito</i>, demonstrando a mesma força.
Escolaridade – Os falantes cariocas do ensino fundamental foram os que mais preencheram a casa do <i>sujeito</i>, com 88% de ocorrências, porém, o ensino superior alcançou índices muito próximos, 81%, dessa forma, houve certo equilíbrio no preenchimento pronominal entre as escolaridades.
Faixa etária – Houve a predominância do preenchimento pronominal na fala de informantes da segunda faixa etária, seguido da primeira, com uma pequena diferença, e, por último, a terceira faixa etária. Esta última foi a que mais apresentou <i>sujeitos nulos</i>, demonstrando, ainda, ser a faixa que mais privilegia o <i>sujeito nulo</i>.
Tempo e modo verbal – O tempo e modo verbal presente do indicativo e passado do indicativo demonstraram maior efetividade no preenchimento do sujeito.
Animacidade - Fator excluído pelo programa. No cômputo geral, os sujeitos [+] animados favorecem o <i>sujeito</i> pronominal. Não houve casos de sujeitos [-] animados em nosso <i>corpus</i>.
Concordância verbal – Fator excluído pelo programa. A não marca de flexão incide favoravelmente ao <i>sujeito expresso</i>, enquanto a concordância favorece o <i>sujeito nulo</i>.
Sexo – O fator extralinguístico “sexo” foi excluído pelo programa, peso relativo, por não ter demonstrado influência nas variáveis dependentes. Entretanto, cabe destacar que: os sexos masculino e feminino demonstraram preferência pelo preenchimento pronominal do sujeito, com uma ligeira diferença, as mulheres com 88% e os homens com 81%.

Fonte: Elaborado pela autora.

CONSIDERAÇÕES FINAIS

Pautados em estudos que revelam que o português brasileiro já não é mais considerado como língua tipicamente de *sujeitos nulos*, com a preferência pelo preenchimento com formas pronominais expressas, conforme Duarte (1995; 2000; 2003); Cyrino (1994; Cyrino, Duarte e Kato 2000), investigamos, neste trabalho, a realização do *sujeito de referência arbitrária* na variedade carioca, a fim de verificar se essa variedade ainda sustenta o *status* de língua [+] *pro-drop*, ou seja, língua marcada positivamente para o *sujeito nulo*, como é caracterizado no português brasileiro. Para isso, baseamo-nos em amostras de falas recentes de moradores de Copacabana (RJ), estratificadas em três faixas etárias, duas escolaridades (fundamental e médio) e gênero (masculino e feminino).

Em nosso estudo observamos que a referência arbitrária apresenta índices significativos de preenchimento pronominal, mesmo sendo considerada como um contexto resistente ao *sujeito expreso*, por se tratar de um sujeito indeterminado que, tradicionalmente é prescrito pelos compêndios gramaticais, pelo emprego do verbo na 3ª pessoa do plural, sem referência específica, e o uso do verbo na 3ª pessoa do singular, acompanhado da partícula *se*. (ALMEIDA, 1999; BECHARA, 2003; CUNHA e CINTRA, 2007; CEGALLA, 2008; ROCHA LIMA, 2011).

Assim como Duarte (1995), Godoy (1999), Carvalho (2005) e Assunção (2012), cujas pesquisas apresentamos no capítulo 2, observamos que os falantes brasileiros, no caso de nosso *corpus* - os cariocas, têm a preferência pela forma predominantemente expressa do *sujeito*, confirmando, assim, o que as autoras ressaltam em relação a essa língua que, segundo elas, percorre um caminho de mudança que afeta o parâmetro de *sujeito nulo*, indicando que o português brasileiro já não é mais uma língua tipicamente [+] *pro-drop*, mas sim parcialmente de *sujeitos nulos*, justamente por aceitar em seu sistema o *sujeito-pronome* em contextos que, tradicionalmente, seria marcado por uma categoria vazia.

Registramos em nosso *corpus* 817 ocorrências com sujeitos de referência arbitrária, deste total 83% eram *sujeitos* preenchidos com pronomes e 17% representados pela categoria vazia. Tais porcentagens sugerem que a variedade carioca, assim como as variedades do sul do Brasil estudada por Godoy (1999), do Centro-Oeste, analisada por Carvalho (2005) e do Nordeste, pesquisada por Assunção (2012), privilegiam o *sujeito pronominal* ao invés do *sujeito nulo*. Nossos resultados se aproximam dos apresentados por Assunção (2012), que

também registrou alta porcentagem de *sujeitos preenchidos* (87%) e baixa porcentagem de *sujeitos nulos* (13%).

Ao analisarmos os fatores selecionados para esta pesquisa, confirmamos nossas hipóteses quanto ao favorecimento do preenchimento pronominal do sujeito de referência arbitrária, como seguem:

- a) A não correferencialidade privilegia o *sujeito-pronome*;
- b) Os pronomes “você” e “a gente” demonstraram maior produtividade no preenchimento do *sujeito*;
- c) Os falantes do ensino fundamental utilizam com maior frequência o *sujeito-pronome*, embora o ensino superior tenha apresentado alto índice de preenchimento. Ainda, assim, os menos escolarizados privilegiam as novas estratégias de indeterminação;
- d) As faixas etárias mais jovens privilegiam o uso do *sujeito pronominal*, enquanto a terceira faixa etária o *sujeito nulo*;
- e) O tempo e modo verbal presente do indicativo é o mais corrente no preenchimento do *sujeito*.

Os resultados aqui apresentados mostram que a variedade carioca tem a preferência por formas pronominais expressas, na realização do sujeito de referência arbitrária, porém, o que expomos, neste trabalho, é apenas uma amostra do que está acontecendo em todo o Brasil, como já explicitado nos trabalhos de Godoy (1999), Carvalho (2005) e Assunção (2012), sobre as diferentes variedades do português brasileiro, atestando o uso de novas estratégias de indeterminação do *sujeito*. Dessa forma, aos poucos, o português brasileiro perde suas características que o nomeia língua tipicamente de *sujeitos nulos*, apresentando em seu sistema outras possibilidades.

Falamos aqui do *sujeito de referência arbitrária*, porém, há um vasto volume de estudos em relação ao *sujeito de referência definida*, que apresentam evidências de que o português brasileiro tem a preferência pelo uso do *sujeito* foneticamente realizado. Destacamos a pesquisa de Duarte e Reis (2018), que analisam o mesmo *corpus* desta pesquisa, mas que centram o seu olhar para o caráter resistente da 3ª pessoa para o preenchimento do *sujeito*. Os autores revelam que o preenchimento com a 3ª pessoa avança mais paulatinamente, entretanto, segue o mesmo caminho de mudança que a 1ª e 2ª pessoa.

Esperamos, enfim, ter contribuído com os estudos sobre a realização do *sujeito* no português brasileiro, que tem demonstrado um comportamento diferente daquele esperado para as línguas marcadas positivamente para o *Parâmetro do Sujeito Nulo*, revelando preferência pelo *sujeito expresso*, dentre os quais destacam-se os pronomes.

REFERÊNCIAS

- ALMEIDA, Napoleão Mendes de. **Gramática metódica da língua portuguesa**. 44 edição. Rio de Janeiro: Saraiva, 1999.
- BARBOSA, Pilar; DUARTE, Maria E. Lamoglia; KATO, Mary A. **Sujeito indeterminados em PE e PB**. In: XVI Congresso da Associação Portuguesa de Linguística, Coimbra, p. 1-15, 2000.
- BRANCO, Thaísa Helena Castelo. **Construções de tópico em foco: uma investigação sobre o fenômeno da topicalização na cidade de Governador Nunes Freire/ MA**. Dissertação de mestrado. São Luís, 2019.
- BECHARA, Evanildo. **Moderna gramática portuguesa**. 37 edição. Rio de Janeiro: Lucerna, 2003.
- BESSA, Jaqueline Alves Gonçalves. **Uma interpretação passiva ou impessoal do pronome se?** Uma contribuição para os revisores de textos. Trabalho de Pós-Graduação, Centro Universitário de Brasília, Brasília, 2016.
- CARVALHO, G. A. de. **A realização do sujeito na fala do araguiense**. Dissertação de Mestrado em Linguística e Língua Portuguesa. Araraquara: Universidade Estadual Paulista – UNESP, 2005.
- CAVALCANTE, Silvia R. de Oliveira; DUARTE, Maria. E. Lamoglia. **Sujeito de referência arbitrária em sentenças infinitivas do português e o parâmetro do sujeito nulo**. In: XXIV Encontro Nacional da Associação Portuguesa de Linguística, Lisboa, p. 187-197, 2009.
- CEGALLA, Domingos Paschoal. **Novíssima gramática da língua portuguesa**. 48. ed. São Paulo: Companhia Editora Nacional, 2008.
- COLSATO, Andrea. **Estratégias de indeterminação em sentenças infinitivas não flexionadas do português brasileiro**. Estudos Linguísticos, São Paulo, 38 (1), p. 347-357, 2009.
- CUNHA, Celso; CINTRA, Luís F. Lindley. **A nova gramática do português contemporâneo**. 3ª edição. Rio de Janeiro: Lexikon Informática, 2007.

CYRINO, S. M. L. **O objeto nulo no português do Brasil**: um estudo sintático diacrônico. Tese de Doutorado. Universidade Estadual de Campinas, 1994.

DUARTE, M. E. L.; REIS, E. P. R. Revisitando o *sujeito* pronominal vinte anos depois. **ReVEL**, vol.16, Rio de Janeiro, 2018.

DUARTE, Maria Eugênia Lamoglia. O *sujeito* de referência indeterminada em sentenças infinitivas. **Revista GEL**, S. J. do Rio Preto, v. 5, n. 1, p.9-30, 2008.

DUARTE, Maria Eugênia Lamoglia. A evolução na representação do *sujeito* pronominal em dois tempos. In: PAIVA, M. C.; DUARTE, M. E. L. **Mudança linguística em tempo real**. Rio de Janeiro: Contra Capa, p. 115-128, 2003.

DUARTE, Maria Eugênia Lamoglia. **O sujeito pronominal no português coloquial europeu**. Trabalho inédito, [1995, ?].

GRAVINA, Aline Peixoto. **Diacronia e sujeito nulo no português brasileiro**: um estudo comparativo. *Filol. Linguística Portuguesa*, São Paulo, v. 16, p. 199-231, 2014.

LAPERUTA, M. **A realização do sujeito pronominal**: um estudo sociolinguístico paramétrico para a cidade de Londrina – Norte do Paraná. Dissertação (Mestrado em Linguística e Língua Portuguesa). 2002. 138f. Faculdade de Ciências e Letras da Universidade Estadual Paulista, Araraquara, 2002.

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

LUCCHESI, D; ARAÚJO, S. **A Teoria da Variação e Mudança**. Projeto vertentes: Bahia, 2021.

NUNES, Jairo M. *Se* apassivador e *se* indeterminador: o percurso diacrônico no português brasileiro. **Cad. Est. Linguísticos**, Campinas, 20, p. 33-58, 1991.

PAGOTTO, Emílio Gozze. Sociolinguística. In: PFEIFFER, C.C. & NUNES, J. H. (org.) **Linguagem, História e Conhecimento**. Campinas-SP: Pontes, 2006.

TARALLO, Fernando. **A pesquisa sociolingüística**. São Paulo: Ática, 2001.

VITÓRIO, Elyne Giselle de Santana Lima Aguiar. **Ter/haver existenciais na fala alagoana: variação estável ou mudança em progresso?** Tese apresentada ao Programa de Pós-Graduação em Letras e Linguística da Universidade Federal de Alagoas. Maceió, 2012.

VERÍSSIMO, Victor. A evolução do conceito de parâmetro do sujeito nulo. **Entrepalavras** - Revista de Linguística do Departamento de Letras Vernáculas - Universidade Federal do Ceará, Fortaleza, v. 7, p. 76-90, jan./jun. 2017.