

UNIVERSIDADE FEDERAL DE ALAGOAS
FACULDADE DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA E LITERATURA

EMANUELLE CAMILA MORAES DE MELO ALBUQUERQUE LIMA

A SEGUNDA PESSOA DO SINGULAR NA COMUNICAÇÃO DIGITAL DE
EMPRESAS ALAGOANAS: A VARIAÇÃO ENTRE PREENCHIMENTO
PRONOMINAL E SUJEITO NULO

MACEIÓ- AL
2023

UNIVERSIDADE FEDERAL DE ALAGOAS
FACULDADE DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA E LITERATURA

EMANUELLE CAMILA MORAES DE MELO ALBUQUERQUE LIMA

A SEGUNDA PESSOA DO SINGULAR NA COMUNICAÇÃO DIGITAL DE
EMPRESAS ALAGOANAS: A VARIAÇÃO ENTRE PREENCHIMENTO
PRONOMINAL E SUJEITO NULO

Tese apresentada ao Programa de Pós-Graduação em Linguística e Literatura como requisito parcial para obtenção do título de Doutor em Linguística, sob orientação da professora Dra. Telma Moreira Viana Magalhães.

MACEIÓ- AL
2023

**Catálogo na fonte Universidade
Federal de Alagoas Biblioteca Central
Divisão de Tratamento Técnico**
Bibliotecário: Marcelino de Carvalho Freitas Neto – CRB-4
– 1767

L732s Lima, Emanuelle Camila Moraes de Melo Albuquerque.

A segunda pessoa do singular na comunicação digital de empresas alagoanas : a variação entre preenchimento pronominal e sujeito nulo / Emanuelle Camila Moraes de Melo Albuquerque Lima. – 2023.

134 f. : il.

Orientadora: Telma Moreira Viana Magalhães.

Tese (doutorado em Linguística) – Universidade Federal de Alagoas.

Faculdade de Letras. Programa de Pós-Graduação em

*Aos meus filhos, João Vitor e João Vinícius,
minhas razões para sempre seguir em frente.*

AGRADECIMENTOS

Enfrentar os desafios de cursar doutorado, sendo professora, gestora universitária e mãe de duas crianças, em um cenário de pandemia, foi, sem dúvidas, um dos momentos mais árduos de minha vida! Assim, reconheço e sou grata a todas as forças que me auxiliaram durante essa caminhada.

Primeiramente, agradeço a Deus por todo sustento que foi dado nesse percurso. A fé renovou-me em todos os momentos que eu precisei!

Agradeço ao PPGLL por toda compreensão e auxílio diante das dificuldades enfrentadas e por todo conhecimento compartilhado junto aos docentes e colegas discentes.

De maneira especial, agradeço à professora, orientadora e amiga querida Telma Magalhães, por toda força, confiança e carinho depositadas em mim desde o mestrado até aqui. Admiro sua força, inteligência e seu coração sempre tão generoso. Sem seu acolhimento, não teria conseguido!

Meu agradecimento especial ao meu companheiro, Bruno Pessoa, por compreender minhas ausências, por ser ânimo, força, cuidado, carinho, torcida constante e amor sem medida. Apenas ele sabe o quão desafiador foram esses anos e quantos esforços foram necessários para concluir essa etapa. Obrigada, meu amor!

De forma especial, também agradeço a minha mãe que é uma verdadeira força da natureza que me guia e orienta. Meu maior exemplo de perseverança, amor e resiliência.

Agradeço aos meus familiares pelo carinho e torcida que sempre demonstraram em tudo que empreendi na vida. Eles são meu esteio e o afeto que fizeram a diferença nesse percurso.

Agradeço aos meus amigos que tanto aliviaram as dificuldades não só desse processo, mas de todos os demais momentos de minha vida. Especificamente, me refiro aqui aos grandes amigos que ganhei de presente da UFAL e que carregarei pra sempre. Aos meus “Letrados e Letradinhos” e “FIG liberado”, meu muitíssimo obrigada!

Aos amigos Rafael, Adeilson, Eudes e Pedro, agradeço por serem mais do que amigos, mas verdadeiras inspirações de profissionais e de pessoas que eu tenho a sorte de cultivar na vida.

À Universidade Federal do Agreste de Pernambuco, agradeço por ser casa, acolhimento, lugar de crescimento e de construção de laços tão significativos. Em especial, minha gratidão por toda compreensão e confiança do professor Airon Melo, reitor da UFAPE; À minha equipe querida e comprometida da Pró-Reitoria de Ensino e Graduação da UFAPE pelo suporte diário diante de todas as demandas; Aos meus colegas docentes com quem aprendo todos os dias e, de forma muito especial, aos meus alunos e minhas alunas que tanto me fizeram e fazem crescer durante meu fazer docente.

À Selma Cruz, por todo auxílio e paciência para me auxiliar a desvendar o universo do software R.

Agradeço muitíssimo, ainda, aos professores Marcello Marcelino, Renata Araújo, Jair Barbosa e Adeilson Sendrins, pelo aceite para participação da banca de defesa e pelas valiosas contribuições que aprimorarão a versão final do trabalho

E, por fim, mas não menos importante, aos meus meninos João Vitor e João Vinícius que foram e sempre serão o combustível diário para meu crescimento. A razão para ultrapassar qualquer desafio imposto e o amor incondicional que alimenta meus dias.

“Se você gritasse,
se você gemesse,
se você tocasse a valsa vienense,
se você dormisse,
se você cansasse,
se você morresse...
Mas você não morre,
você é duro, José!”

Carlos Drummond de Andrade

RESUMO

Esta tese analisa o preenchimento versus não preenchimento da segunda pessoa do singular na posição de sujeito na comunicação digital, especificamente, na rede social Instagram. Estudos como os de Duarte (1993, 1995, 2003), Duarte *et al.*, (2012) e Santos, (2014), entre outros, apontam para o processo de variação linguística entre preenchimento pronominal e o sujeito nulo no português brasileiro, nesse sentido, nossa investigação busca verificar se o fenômeno também é recorrente na comunicação digital, especificamente em relação à segunda pessoa do singular, e quais fatores são determinantes para o preenchimento ou não dessa posição de sujeito. Avaliando as características específicas dessa modalidade de uso da língua, também pretendemos apresentar as similitudes ou distinções com os dados de fala do português brasileiro. Para tal, amparamos nossa investigação à luz do modelo teórico-metodológico da Sociolinguística Variacionista, proposto por Labov (1972), que analisa as relações entre a língua e o contexto social da comunidade linguística. Através de postagens de quatro empresas naturais do estado de Alagoas, buscamos verificar como é realizada a segunda pessoa do singular na posição de sujeito nesse contexto de comunicação digital no qual o consumidor é o interlocutor direto com o qual cada empresa busca efetivar a comunicação. O *corpus* da pesquisa foi constituído através de dados retirados de postagens públicas na página oficial do Instagram de cada uma das empresas selecionadas, no período de fevereiro de 2022 a maio de 2023, buscando traçar uma análise longitudinal. As variáveis independentes elencadas para nossa pesquisa foram o tipo de sentença (interrogativa e não interrogativa), tempo verbal (passado e não passado), paralelismo e modo verbal (indicativo, subjuntivo, imperativo e infinitivo pessoal). Após a coleta, os dados foram tratados estatisticamente através do software R. Os dados apontam para a predominância do sujeito nulo e para a exclusividade do uso do pronome inovador “você” nos casos da expressão de sujeito pleno. As variáveis analisadas mostraram-se influentes no que tange à escolha entre as variantes analisadas e a variável paralelismo foi selecionada pelo software R como a mais relevante estatisticamente. Os resultados revelam que, na modalidade de comunicação digital, existe um processo de variação estável, no qual as realizações de uso do preenchimento do sujeito diferem das encontradas na fala e na escrita convencional do português.

PALAVRAS-CHAVE: Segunda Pessoa do Singular. Preenchimento do Sujeito. Não Preenchimento do Sujeito. Variação Linguística. Comunicação Digital.

ABSTRACT

This thesis analyzes the filling versus non-filling of the second person singular when in subject position in digital communication, specifically, on the social network Instagram. Studies such as Duarte (1993, 1995, 2003), Duarte et al., (2012) and Santos, (2014), among others, point to the process of linguistic variation between pronominal filling and the null subject in Brazilian Portuguese, in that regard, our investigation seeks to verify whether the phenomenon is also recurrent in digital communication, specifically in relation to the second person singular, and which factors are determining whether or not this subject position is fulfilled. Evaluating the specific characteristics of this form of language use, we also intend to present the similarities or distinctions with the forms of spoken language in Brazilian Portuguese. For this purpose, we support our investigation considering the theoretical-methodological model of Variationist Sociolinguistics, proposed by Labov (1972), which analyzes the relationships between language and the social context of the linguistic community. Through posts of four companies from the state of Alagoas, we sought to verify how the second person singular is performed in the subject position in the context of digital communication in which the consumer is the direct interlocutor with which each company seeks to carry out communication. The corpus was constituted through data taken from public posts of the official Instagram page of each of the selected companies, from February 2022 to May 2023, seeking to draw up a longitudinal analysis. The independent variables listed for our research were the type of sentence (interrogative and non-interrogative), verbal tense (past and non-past), parallelism and verbal mood (indicative, subjunctive, imperative and personal infinitive). After the collection, the data were statistically processed using the R software. The data points to the predominance of the null subject and the exclusive use of the innovative pronoun “você” in cases of full subject expression. The variables analyzed proved to be influential in terms of the choice between the analyzed variants and the parallelism variable was selected by the R software as the most statistically relevant. The results reveal that, in the digital communication modality, there is a process of stable variation, in which the realizations of using subject completion differ from those found in conventional Portuguese speech and writing.

KEYWORDS: Second Person Singular. Filling in the Subject. No fulfillment of the Subject. Linguistic Variation. Digital communication.

RESUMEN

Esta tesis analiza el relleno versus no relleno de la segunda persona del singular en la posición de sujeto en la comunicación digital, específicamente, en la red social Instagram. Estudios como los de Duarte (1993, 1995, 2003), Duarte et al., (2012) y Santos, (2014), entre otros, señalan el proceso de variación lingüística entre el relleno pronominal y el sujeto nulo en el portugués brasileño. En este sentido, nuestra investigación busca comprobar si el fenómeno también es recurrente en la comunicación digital, específicamente en relación con la segunda persona del singular, y qué factores están determinando que se cumpla o no esta posición de sujeto. Al evaluar las características específicas de esta forma de uso de la lengua, pretendemos también presentar las similitudes o distinciones con las formas de lengua hablada en el portugués brasileño. Para ello, apoyamos nuestra investigación a la luz del modelo teórico-metodológico de la Sociolingüística Variacionista, propuesto por Labov (1972), que analiza las relaciones entre la lengua y el contexto social de la comunidad lingüística. A través de publicaciones de cuatro empresas del estado de Alagoas, buscamos verificar cómo se desempeña la segunda persona del singular en la posición de sujeto en este contexto de comunicación digital en el que el consumidor es el interlocutor directo con el que cada empresa busca realizar comunicación. El corpus de investigación se constituyó a través de datos tomados de publicaciones públicas en la página oficial de Instagram de cada una de las empresas seleccionadas, desde febrero de 2022 hasta mayo de 2023, buscando elaborar un análisis longitudinal. Las variables independientes enumeradas para nuestra investigación fueron el tipo de oración (interrogativa y no interrogativa), tiempo verbal (pasado y no pasado), paralelismo y modo verbal (indicativo, subjuntivo, imperativo e infinitivo personal). Después de la recopilación, los datos fueron procesados estadísticamente utilizando el software R. Los datos apuntan al predominio del sujeto nulo y al uso exclusivo del pronombre innovador “tú” en los casos de expresión completa del sujeto. Las variables analizadas demostraron ser influyentes en términos de la elección entre las variantes analizadas y la variable de paralelismo fue seleccionada por el software R como la más relevante estadísticamente. Los resultados revelan que, en la modalidad de comunicación digital, hay un proceso de variación estable, en el que las realizaciones del uso de la finalización de materias difieren de las encontradas en el habla y la escritura portuguesa convencional.

PALABRAS-LLAVE: Segunda Persona del Singular. Relleno del Sujeto. No Relleno del Sujeto. Variación Lingüística. Comunicación Digital.

LISTA DE TABELAS

Tabela 1. Pronomes pessoais na GT.....	27
Tabela 2. Distribuição dos pronomes de segunda pessoa.....	49
Tabela 3. Quantitativo geral dos dados.....	97
Tabela 4. Codificação das variáveis.....	98
Tabela 5. Percentuais de realização das variantes por tipo de sentença.....	109
Tabela 6. Percentuais de realização das variantes na variável paralelismo....	121
Tabela 7. Significância das variáveis excluídas	123
Tabela 8. Significância da variável paralelismo.....	123
Tabela 9. Valores de significância e pesos da variável paralelismo.....	124

LISTA DE GRÁFICOS

Gráfico 1. Total das ocorrências.....	102
Gráfico 2. Distribuição das variantes por empresa.....	106
Gráfico 3. Distribuição total das variantes por tipo de sentença.....	107
Gráfico 4. Distribuição das variantes por tipo de sentença.....	109
Gráfico 5. Distribuição das variantes por tempo verbal.....	111
Gráfico 6. Distribuição das variantes por paralelismos.....	118

LISTA DE IMAGENS

Imagem 1. Perfil oficial Instagram- Empresa Carajás.....	90
Imagem 2. Perfil oficial Instagram- Empresa Sococo.....	91
Imagem 3. Perfil oficial Instagram- Grupo Coringa.....	93
Imagem 4. Perfil oficial Instagram- Grupo Coagro.....	94
Imagem 5. Postagem Empresa Coagro.....	95
Imagem 6. Postagem Empresa Carajás.....	95
Imagem 7. Postagem Empresa Coringa.....	96
Imagem 8. Sujeito pronominal de segunda pessoa.....	103
Imagem 9. Ocorrência de sujeito nulo.....	104
Imagem 10. Variante Você em sentença interrogativa.....	108
Imagem 11. Variante Sujeito Nulo em sentença interrogativa.....	108
Imagem 12. Variante Você no tempo verbal passado.....	111
Imagem 13. Variante Sujeito Nulo no tempo verbal não passado.....	112
Imagem 14. Variante Sujeito Nulo no modo verbal imperativo.....	115
Imagem 15. Variante Você no modo verbal subjuntivo.....	115
Imagem 16. Formas não paralelas na variável paralelismo.....	119
Imagem 17. Forma única/isolada na variável paralelismo.....	120

SUMÁRIO

1. INTRODUÇÃO.....	16
2. A REALIZAÇÃO DA SEGUNDA PESSOA NO PORTUGUÊS BRASILEIRO.....	25
2.1. O SISTEMA PRONOMINAL DO PORTUGUÊS BRASILEIRO E A REPRESENTAÇÃO DA SEGUNDA PESSOA	25
2.2. O SUJEITO NULO NO PORTUGUÊS BRASILEIRO	31
2.3. ESTUDOS ANTERIORES: A REALIZAÇÃO DA SEGUNDA PESSOA DO SINGULAR NOS DADOS DO PB.....	35
2.3.1. A variação nos dados do PB: o preenchimento pronominal do sujeito de segunda pessoa do singular.....	35
2.3.1.1. As pesquisas na região Sul: Lorengian-Penkall (2005) e Franceschini (2010).....	35
2.3.1.2. As pesquisas na região Sudeste: Mota (2008), Silva e Gonçalves (2016) e Modesto (2006).....	37
2.3.1.3. As pesquisas na região Centro-Oeste: Scherre et al (2011) e Andrade (2010).....	39
2.3.1.4. As pesquisas na região Norte: Costa (2013) e Martins (2010).....	40
2.3.1.5. As pesquisas na região Nordeste: Alves (2010), Soares (1980), Nascimento (2017).....	41
2.3.1.5.1. As pesquisas em Alagoas: Vitória (2018), Silva e Vitória (2019) e Silva (2019).....	44
2.3.2. Uma visão geral: a proposta de subsistemas de Scherre <i>et al</i> (2015).....	47
2.4. ESTUDOS ANTERIORES: O SUJEITO NULO NO PORTUGUÊS BRASILEIRO.....	50
2.4.1. Duarte (1995).....	51
2.4.2. Farias e Martins (2021).....	52
2.4.3. Silva (2021).....	54
2.4.4. Nunes (2000).....	55

2.4.5. Santos (2014).....	56
3. A COMUNICAÇÃO DIGITAL: QUE MODALIDADE É ESSA?.....	59
3.1.FENÔMENOS LINGUÍSTICOS NA COMUNICAÇÃO DIGITAL.....	64
3.1.1.Othero et al (2018).....	66
3.1.2. Silva e Pinheiro (2020).....	67
3.1.3. Pinheiro (2021).....	69
4. A SOCIOLINGUÍSTICA VARIACIONISTA.....	72
4.1. O ESCOPO TEÓRICO DA SOCIOLINGUÍSTICA LABOVIANA.....	72
4.2. A METODOLOGIA SOCIOLINGUÍSTICA LABOVIANA.....	78
4.2.1. Os procedimentos metodológicos adotados para o estudo.....	79
4.2.1. 1. As variáveis linguísticas elencadas.....	79
4.2.1.1.1. Variável tipo de sentença.....	80
4.2.1.1.2. A variável tempo verbal.....	82
4.2.1.1.3. A variável modo verbal	83
4.2.1.1.4. A variável paralelismo.....	83
4.3. O LÓCUS DA PESQUISA: A REDE SOCIAL INSTAGRAM.....	84
4.4. A COLETA DOS DADOS E A CONSTITUIÇÃO DO CORPUS.....	86
4.5.SOBRE AS ESCOLHAS:O ESTADO E AS EMPRESAS ALAGOANAS.....	89
4.6. A AMOSTRA.....	94
4.7. O TRATAMENTO ESTATÍSTICO DOS DADOS: SOFTWARE R.....	98
5. APRESENTAÇÃO E ANÁLISE DOS DADOS.....	101
5.1. AS VARIÁVEIS DE SEGUNDA PESSOA DO SINGULAR NA POSIÇÃO DE SUJEITO.....	101
5.2. O PAPEL DAS VARIÁVEIS INDEPENDENTES.....	107
5.2.1. Variável tipo de sentença.....	107

5.2.2. Variável tempo verbal.....	110
5.2.3. Variável modo verbal.....	113
5.2.4. Variável paralelismo.....	117
5.3. AVALIANDO OS RESULTADOS.....	122
6. CONSIDERAÇÕES FINAIS.....	127
REFERÊNCIAS.....	130

1. INTRODUÇÃO

O sistema pronominal do Português Brasileiro (PB) apresenta-se como um terreno fértil de pesquisa para diversas correntes teóricas. Sendo objeto de trabalhos que buscam dar conta da complexidade inerente a esse sistema, bem como as especificidades encontradas ao longo do país no que tange ao seu uso.

Na literatura dos estudos linguísticos, observamos uma série de pesquisas que buscam dar conta de fenômenos que nos são dados pelas diversas formas pronominais, dentre eles, destacamos Scherre (2023), Scherre *et al.* (2015), Duarte (2003), Dias (2007), Menom (2000), entre muitos outros.

No âmbito da Gramática Tradicional (GT), percebemos que ainda há a sustentação da presença exclusiva do “tu” em seus quadros pronominais, e que ela não apresenta o “você” como uma opção para uso na posição de sujeito da segunda pessoa do singular (Bechara, 2004; Cunha *et al.*, 1985). Porém, os dados de diversas pesquisas revelam que o português brasileiro prevê o uso de pelo menos seis formas para a realização dessa função e todas são permitidas e legitimadas pelo sistema linguístico.

Vejamos as sentenças que seguem abaixo:

1. a) *Tu vais/vai¹ para a festa?*
- b) *Você vai para a festa?*
- c) *Cê vai para a festa?*
- d) *Ocê vai para a festa?*
- e) *Ø Vai/vais para a festa?*
- f) *O senhor/ a senhora vai para festa?*

¹Em nossa pesquisa, levaremos em consideração, enquanto dado, sentenças que utilizam a segunda pessoa “tu” com a concordância verbal de segunda e/ou terceira pessoa, sem distinção, tendo em vista as evidências científicas do maior uso da forma “tu” sem concordância no PB e o foco principal na forma de realização explícita ou não da segunda pessoa pronominal em si. Ver Scherre *et al.* (2021)

2. a) *Tu perdeu o horário de novo.*

b) *Você perdeu o horário de novo.*

c) *Cê perdeu o horário de novo.*

d) *Ocê perdeu o horário de novo.*

e) \emptyset *Perdeu/Perdesse o horário de novo.*

f) *O senhor/ a senhora perdeu o horário de novo.*

Observando os exemplos acima, podemos perceber que os seis tipos de sentenças são perfeitamente gramaticais no português. Verificamos que todas as formas pronominais elencadas acima podem preencher a posição de sujeito de segunda pessoa do singular. Nesse sentido, Scherre *et al* (2015) pontuam que :

[...] as formas pronominais de segunda pessoa “você”, “cê”, “ocê” e “tu” estão todas vivas no português brasileiro, com matizes próprios e diversos, a depender da região e/ou da comunidade [...] A ideia mais difundida é a de que o pronome “você” pleno e explícito (sempre com concordância zero) tem uso sistematicamente mais generalizado do que o pronome “tu” (com concordância explícita variável, mas usualmente sem concordância, ou seja, o “você” teria um uso essencialmente não marcado (Scherre *et al.*, 2015, p. 169-170).

Outra importante observação ainda é que, como visto nas sentenças em 1 (e) e 2 (e), a função de sujeito da segunda pessoa do singular pode ser realizada através do “sujeito nulo”, ou seja, o não preenchimento da segunda pessoa pronominal também é licenciado no português como uma possibilidade de referência à segunda pessoa. De acordo Castilho (2010), o sujeito no PB pode ser preenchido por diferentes classes gramaticais e pode ser elidido também, através da anáfora- zero \emptyset .

Sendo assim, as diferenças em relação ao uso de cada uma das variantes parecem incidir apenas nas motivações condicionadas por variáveis linguísticas e extralinguísticas e não em relação às possibilidades reais de uso.

O conceito de sujeito, que nos é importante para análise aqui

empreendida, não é pacífico nas gramáticas e sua classificação é, muitas vezes baseada em critérios de natureza distintas, conforme aponta Sedrins (2021). Contudo, essa função sintática é reveladora de muitas possibilidades de uso da língua e por isso merece nossa atenção.

Para exemplificar tal problemática e justificar a escolha dessa categoria sintática para nossa análise, observemos que de acordo com o que é postulado na gramática de Cegalla, o sujeito “é o ser do qual se diz alguma coisa”. O autor ainda afirma que o sujeito e o predicado são os dois termos essenciais da oração (Cegalla, 1998, p. 295). Já em Bechara (2006, p.409) “chama-se sujeito à unidade ou sintagma nominal que estabelece uma relação predicativa com o núcleo verbal para construir uma oração” (Bechara, 2006, p. 409).

Como visto, a proposta conceitual do segundo gramático já a distingue da do primeiro, pois o critério de Bechara privilegia a forma, a estrutura sintática, enquanto a de Cegalla traz a noção semântica para sua definição, ou seja, a abordagem prescritiva gramatical não possui um consenso quanto a esse termo e a análise de outras gramáticas também corroboram com essa inconsistência conceitual e classificatória.

Focando especificamente no sujeito não expresso, denominado como oculto, as gramáticas, de um modo geral, convergem para sua classificação postulando que é aquele identificável através da desinência verbal, pois não está explícito na oração.

Partindo para o campo de estudos da Linguística, temos uma classificação mais interessante, proposta para Duarte (2007) e que assumimos ao longo desta tese, na qual o sujeito deve ser definido: a) quanto à forma (estrutura) – o sujeito pode vir expresso ou não; b) quanto à referência (seu conteúdo, valor semântico) – o sujeito pode ter referência definida, indefinida ou não ter qualquer referência. Ainda de acordo com Duarte (2007), sempre que buscamos identificar os constituintes de uma oração estamos observando a estrutura argumental que é projetada pelo predicador, enquanto responsável pela estrutura da sentença. Assim, estamos procurando compreender qual a seleção semântica que esse predicador faz. O sujeito, nesse caso, é um dos termos constituintes da oração, enquanto um argumento externo, à esquerda, requerido pelo predicador, podendo ser expresso ou não.

Voltando nossa atenção para as possibilidades de preenchimento ou não dessa posição de sujeito da segunda pessoa do singular e tendo em vista a complexidade envolvida nesse processo de variação no PB, encontramos diversos trabalhos que discutem essa questão no país. Nos dialetos do Sul e do Sudeste do Brasil, temos trabalhos como os de Sousa (2016), Scherre *et al.* (2015), Monteiro (1997), Negrão e Muller (1996), Menon (1996) e Neves (1992). Por outro lado, trabalhos, como os de Alves (2010) e Nascimento (2017), focaram na realização da segunda pessoa do singular em estados nordestinos e, em Alagoas, especificamente, temos as contribuições das pesquisas de Silva (2019) e Silva e Vitória (2017).

Especificamente sobre as formas de preenchimento da segunda pessoa, contudo, de acordo com Scherre, Andrade e Catão (2020):

para que tenhamos um mapa ainda mais próximo da realidade, são necessárias e urgentes mais pesquisas no vasto território brasileiro, com o controle de, pelo menos, cinco possibilidades disponíveis no português brasileiro: você, ocê, cê, tu com concordância e tu sem concordância, com o controle rigoroso dos contextos sintáticos e das nuances interacionais. O mapa é dinâmico e o desafio de seu redesenho está sempre lançado. (Scherre, Andrade e Catão, 2020, p. 274)

Logo, além da necessidade de mais pesquisas que foquem na complexidade impressa nessas formas pronominais de segunda pessoa, acrescentamos que é imperativo observar a função de sujeito da segunda pessoa do singular, também, quando ela não está explícita por elemento pronominal, quando há o que a literatura comumente denomina de sujeito nulo. Assim, assumimos que temos a possibilidade de realização de sujeito tanto explícito por um elemento pronominal como pelo não preenchimento.

Nesse sentido, Duarte (1995) buscou analisar a preferência pelo sujeito pronominal explícito e sua relação com a redução nos paradigmas flexionais verbais, analisando textos escritos em épocas distintas no português brasileiro. Assim, a autora conclui:

que o português do Brasil perdeu a propriedade do Princípio 'Evite Pronome' levando-se em conta o enfraquecimento da flexão, responsável pela identificação da categoria vazia sujeito.

Como consequência dessa perda, a língua caminha na direção das línguas que têm a posição sujeito preenchida. (Duarte, 1995: 30 e 134)

Sendo assim, pretendemos averiguar essa afirmação, partindo dos trabalhos já apontados sobre a realização do preenchimento do sujeito explícito e da necessidade de avaliar sua relação com o não preenchimento dessa posição. Somado a isso, reconhecendo que a língua portuguesa é a quinta língua mais utilizada na internet, segundo as mais recentes estimativas, buscamos ampliar o horizonte de descrição e análise desse processo variacionista para outro ambiente de produção linguística ou de modalidade de uso da língua, o da comunicação digital.

Baron (2013, p. 126) pontua que “poucas tentativas empíricas tem sido feitas para contrastar os dados da comunicação mediada por computadores com os *corpora* comparáveis falados ou escritos”. Assim, a proposta aqui empreendida busca verificar se nosso fenômeno de variação é realizado nessa nova modalidade de uso da língua: a comunicação mediada por computadores, smartphones, tablets e demais dispositivos de comunicação conectados à internet. Pinheiro (2021), destaca que:

a sintaxe dos nossos usos linguísticos feitos em ambiente digital pode (e deve) ser descrita cientificamente, a fim de evitarmos generalizações preconceituosas acerca das suas semelhanças com a fala ou a escrita convencionais. Afinal, os gêneros discursivos digitais se mostram muito presentes nas nossas interações cotidianas – e tendem a fazê-lo cada vez mais. Além disso, devido sobretudo à multiplicidade de plataformas nas quais se realizam e à variedade de esferas comunicativas que integram, tais gêneros parecem por vezes sobressair aos gêneros discursivos convencionais. (Pinheiro, 2021, p.68, grifo nosso)

Assumindo que a comunicação digital possui propriedades específicas que a distinguem da comunicação via escrita e/ou oralidade, almejamos descrever e analisar como se dá o preenchimento ou não da segunda pessoa do singular na posição de sujeito no português na comunicação digital. Para tal, definimos que a rede social Instagram seria o ambiente no qual os dados seriam coletados, por ser uma rede de expressiva adesão no país.

Especificamente, buscamos contextos de uso da língua no qual a interlocução direta com o leitor pudesse acontecer e, assim, selecionamos as postagens realizadas por empresas que se comunicam diretamente com seu público em seus perfis públicos e oficiais. As postagens das empresas foram escolhidas por representarem uma comunicação direta com o interlocutor, no caso, o cliente/consumidor. Nelas, cada empresa fala diretamente com seu cliente, de maneira informal, buscando evidenciar uma relação simétrica, demonstrando proximidade com seu cliente.

Para fins de delimitação do campo de coleta de dados, selecionamos as postagens de quatro grandes empresas do estado de Alagoas no intuito de comparar os resultados encontrados nesse ambiente de comunicação com os encontrados nas pesquisas sobre esse processo variacionista na fala e na escrita.

Destacamos que o ambiente de comunicação digital é reconhecido como um solo fecundo para realização de estudos linguísticos, tendo em vista que grande parte da população se comunica diariamente por meio da tecnologia em seu cotidiano, no entanto, ainda são escassas as investigações que buscam dar conta dessa demanda, especialmente, os estudos que visam o enfoque nas estruturas linguísticas e em seus processos de variação. De acordo com Schlobinski (2012):

No espaço digital encontramos formas de comunicação e mundos textuais muito diversos, e, **do ponto de vista da linguagem, o mundo digital é tão colorido e multifacetado como o real.** Vale a pena observá-lo atentamente, no mínimo para não nos deixarmos levar por preconceitos e para podermos contestar generalizações com uma opinião fundamentada. (Schlobinski, 2012,p. 151, grifo nosso)

Logo, a proposta aqui empreendida parece-nos inovadora e justifica-se como pertinente para maior compreensão dos fenômenos linguísticos nessa modalidade do português brasileiro.

Para tal análise, lançamos mão do arcabouço teórico-metodológico da Sociolinguística Variacionista que trouxe a discussão entre a língua e a sociedade, a partir da segunda metade do século XX, para o cerne dos estudos

linguísticos. De acordo com esse modelo teórico, desenvolvido por William Labov na década de 1960, o meio em que um indivíduo vive, seu contexto social, influencia diretamente no uso linguístico, fazendo com que nossa língua se mantenha viva e em constantes modificações. Logo, a língua passa a ser vista como um sistema heterogêneo.

Segundo Coan e Freitag (2010):

A Sociolinguística laboviana não é uma teoria da fala, nem o estudo do uso da língua com o propósito exclusivo de descrevê-la, mas o estudo do uso da língua no sentido de verificar o que ela revela sobre a estrutura linguística (*langue*). Quando Labov fala em heterogeneidade, refere-se à variação, mas está interessado na variação que pode ser sistematicamente explicada. (Coan; Freitag, 2010, p.173).

De acordo com Camacho (2013, p. 19), a Sociolinguística “incrementou, nas últimas três décadas, uma nova compreensão da natureza ao mesmo tempo variável e mutável da linguagem”. E ainda acrescenta que: “ativou também o reconhecimento do caráter regular e sistemático da heterogeneidade mediante um conjunto de estudos empíricos, de natureza quantitativa com foco na língua em uso no contexto social”.

Para Labov, o desenvolvimento da língua enquanto instrumento de comunicação que se manteve ao longo da história, apesar de todos os processos de variação e mudança a ela concernentes, revela a sua propriedade de estabilidade. Nessa perspectiva, considera-se que a língua deve ser encarada como um instrumento de comunicação usado por uma comunidade de fala e, ainda, que a variação é essencial à própria natureza da linguagem humana (Labov, 1972).

Seguindo a metodologia laboviana, que assume o “caos” linguístico como objeto de estudo (Tarallo, 1994), analisamos a realização da segunda pessoa do singular através de pronomes plenos ou do sujeito nulo em contextos de comunicação digital nos quais o diálogo é estabelecido diretamente entre a empresa – locutor- e o consumidor – interlocutor. O *corpus* da pesquisa foi constituído a partir da coleta de postagens retiradas dos perfis públicos e oficiais,

publicadas no período de fevereiro de 2022 a maio de 2023, na rede social Instagram, de quatro empresas alagoanas que possuem seu público consumidor majoritariamente no estado.

Para a análise, elegemos como nossa variável dependente o preenchimento ou não do sujeito de segunda pessoa do singular, a saber, a forma pronominal canônica “tu”, a forma pronominal inovadora “você e o sujeito nulo; como variáveis independentes, elencamos os seguintes grupos de fatores linguísticos: paralelismo, tempo verbal; modo verbal e tipo de sentença.

A presente tese está dividida em seis seções, incluindo esta introdução, conforme pode ser visto a seguir:

A próxima seção versa sobre o conceito de pronome e a noção de pessoa do discurso a partir dos estudos linguísticos e do que Gramática Tradicional postula, além da discussão sobre o processo de gramaticalização da forma “você”. Também apresentamos os estudos que analisaram o fenômeno da variação entre as formas “tu” e “você” no Brasil, com vistas a observar como se dá o preenchimento pronominal através dessas duas formas. Além disso, trazemos a discussão sobre o não preenchimento do sujeito pronominal, apresentando os estudos já realizados e o que tem se diagnosticado acerca dessa realização no PB. Dessa forma, buscamos apresentar um panorama dos estudos ao longo do país, bem como, de forma mais específica, apontamos algumas pesquisas realizadas no Nordeste e as que foram feitas, especificamente, no estado de Alagoas, dando visibilidade para esse processo e apresentando os principais resultados para fins de comparação, na seção de análise, com os dados obtidos em nossa pesquisa.

A seção de número três apresenta os conceitos basilares acerca da comunicação digital, as características específicas dessa modalidade de uso da língua, alguns estudos linguísticos que tratam da comunicação mediada por dispositivos conectados à Internet, bem como ratificamos a importância do estudo linguístico através da interação realizada no meio digital.

A quarta seção traz o suporte teórico-metodológico utilizado para nossa análise. Nela, apresentamos os fundamentos da Sociolinguística Variacionista a partir de Labov (2008), bem como todo aparato metodológico utilizado para coleta, definição e apresentação dos dados. Demonstramos os parâmetros para definição das variáveis dependentes para o estudo, constituição do *corpus* e procedimentos para a análise dos dados através do tratamento estatístico no software R², desenvolvido para análise de dados e modelos estatísticos, desenvolvido por Rossa Ihaka e Robert Gentleman em 1996.

A seção cinco traz a discussão acerca dos dados encontrados na pesquisa, apresentando as reflexões teóricas e análise dos fatores linguísticos que estão envolvidos no processo de variação entre o preenchimento pronominal e a realização do sujeito nulo na segunda pessoa do singular a partir do *corpus* constituído para esta tese, bem como faz uma reflexão acerca das similaridades e distinções que emergem na comparação com os dados arrolados em pesquisas realizadas nas modalidades escrita e oral da língua portuguesa no país.

A sexta seção fecha a tese, retomando o que foi discutido ao longo do trabalho e apresentando nossas considerações finais após a análise realizada.

Com o presente estudo, pretendemos atingir os objetivos aqui expostos, bem como contribuir com as pesquisas que analisam os fenômenos acerca do preenchimento ou não do sujeito de segunda pessoa do singular, mais especificamente os que investigam esse processo à luz da Teoria da Variação, além de colaborar com o empreendimento dos estudos linguísticos que visam lançar o olhar sobre uso da língua através da comunicação digital.

² Disponível em: <https://www.R-project.org/>

2. A REALIZAÇÃO DA SEGUNDA PESSOA NO PORTUGUÊS BRASILEIRO

Nesta seção, objetivando discutir sinteticamente sobre os aspectos relacionados à realização da segunda pessoa do singular no português, buscaremos apresentar brevemente o conceito de pronome, a entrada da forma “você”, através do processo de gramaticalização, no sistema pronominal do português. Além disso, apresentaremos alguns trabalhos que foram desenvolvidos acerca da variação entre as formas pronominais de segunda pessoa do singular no português do Brasil, através de pesquisas realizadas ao longo país, e os trabalhos que analisam as questões concernentes ao sujeito nulo, com o intuito de estabelecer um delineamento do fenômeno na fala e na escrita e, conseqüentemente, dimensionar a realização do processo de variação no preenchimento ou não da segunda pessoa e subsidiar nossas considerações sobre os dados encontrados na modalidade de comunicação digital.

2.1. O SISTEMA PRONOMINAL DO PORTUGUÊS BRASILEIRO E A REPRESENTAÇÃO DA SEGUNDA PESSOA

Desde o surgimento da gramática grega, que é considerada o primeiro estudo formal da língua, tinha-se um termo específico para designar a classe de palavras que tinham por função serem substitutas do nome³. No latim, o termo grego foi traduzido por *pronomem* e entendido como a palavra que substitui o substantivo, ou que o acompanha para dar-lhe ênfase ou para tornar mais claro o seu significado.

No que tange aos pronomes classificados como pessoais, sua característica básica é que eles são, de fato, aqueles que designam as três pessoas gramaticais no processo de comunicação. Estão intimamente ligados à oralidade e, portanto, são basicamente de 1ª. e 2ª. pessoas (“eu” e “tu”), pois são essas pessoas gramaticais que constituem o eixo do diálogo. A 3ª pessoa e as demais se acrescentam por analogia e podem remeter-se àquelas, “eu” e “tu”.

³1 Os pronomes pessoais são os verdadeiros pronomes, pois vêm no lugar do nome. Sua origem mostra que não são anafóricos. As 3ªs pessoas (seja no singular ou plural) estão fora do eixo básico do diálogo; são dêiticos. (Ver: H. Murachco, “Língua Grega” pp. 163-165 vol. I Ed. Vozes/Discurso, São Paulo 2001).

Dessa forma, os papéis do discurso seriam: a 1a. pessoa (“eu”) é a que fala; a 2a. pessoa (“tu”) é aquela a quem se fala e a 3a. pessoa (“ele”) é a de quem se fala.

De acordo com Jordan (1972) distinguem-se três categorias gramaticais específicas para o pronome pessoal latino: pessoa, caso e número. Segundo o autor, na primeira categoria, cada morfema de pessoa determina um certo modo de organização dentro das demais categorias de morfemas. Em relação ao caso, a língua latina resguardava certa abundância, mantendo seis casos e suas flexões. E na categoria número, parece-nos que não é correto afirmar que os pronomes “nós” e “vós” são, simplesmente, os plurais de “eu” e “tu”. Já que o “nós” indica tanto o conjunto das pessoas que fazem parte do diálogo quanto o conjunto das pessoas incluídas pelo falante. Deste modo, compõe este “nós” o “eu e tu”; o “eu e aquele/outro” e “eu e aquele(s)”. O “vós” indica o conjunto dos co-locutores ou outro conjunto também específico. Assim, compõe este “vós” o “tu e tu (outro)” e “tu e ele(s)”.

Corroborando com as definições acima, Neves (2000, p.457) define que: a) primeira pessoa é aquela de quem parte o discurso, e que só aparece no enunciado quando o locutor faz referência a si mesmo; b) segunda pessoa é aquela a quem se dirige o discurso, e que só aparece no enunciado quando o locutor se dirige a ela; e c) terceira pessoa é aquela sobre a qual é o discurso.

Ao observarmos o que é definido na Gramática Tradicional (GT) do português, os pronomes pessoais ainda resguardam essas categorias, como pode ser visto em Bechara (1997, p. 94), que define que pronome é a expressão que designa os seres sem dar-lhes nome nem qualidade, indicando-os apenas como pessoa do discurso. Para o gramático, os pronomes pessoais designam as três pessoas do discurso e assim são representadas:

TABELA 01: PRONOMES PESSOAIS NA GT

SINGULAR	1 ^a	EU
	2 ^a	TU
	3 ^a	ELE/ELA
PLURAL	1 ^a	NÓS
	2 ^a	VÓS
	3 ^a	ELES/ELAS

Fonte: Adaptado de Bechara (1997, p.95)

Na gramática de Napoleão Mendes de Almeida (1999), a categoria pronome é definida como a palavra que ou substitui ou pode substituir um substantivo. Para o autor, todas as pessoas gramaticais são representadas, taxonomicamente, por pronomes pessoais

Analisando o que é posto nas GTs, percebemos que a segunda pessoa é definida e resguardada apenas através da forma pronominal “tu”, não sendo mencionado, portanto, outro elemento de cunho equivalente no quadro dos pronomes ditos retos, diferentemente do que já é observado nas pesquisas empreendidas na Linguística, que aponta para o uso da forma “você” com estatuto de pronome pessoal de segunda pessoa. De forma inusitada e conservadora, a GT ainda não contempla as transformações linguísticas já largamente apresentadas cientificamente nos estudos.

É notável que o sistema pronominal do português enriqueceu com a entrada de elementos gramaticalizados como “a gente” e “você”, por exemplo. Nesse sentido, percebemos que as diversas pressões que atuam sobre a língua trazem inovações e reorganizações necessárias ao uso linguístico, sem perder a unidade e os mecanismos reguladores que ela possui, demonstrando, portanto, a mutabilidade e adaptabilidade inerentes ao funcionamento linguístico.

De acordo com Said Ali (1950), a forma pronominal “você” é fruto de uma evolução de raízes latinas, iniciadas com a introdução dos pronomes tu/vós no português, usados como tratamento direto a quem se dirigia a palavra. Para distinção na hierarquia social, o “tu” era utilizado para o contexto de intimidade e o “vós” para o tratamento mais formal ou cerimonioso. Cunha e Cintra (2008) apontam que:

No português europeu normal, o pronome *tu* é empregado como forma própria de intimidade [...] No português do Brasil, o uso do *tu* restringe-se ao extremo Sul do país e a alguns pontos da região Norte, ainda não suficientemente delimitados. Em quase todo território brasileiro, ele foi substituído pelo *você* como forma de intimidade. *Você* também se emprega, fora do campo de intimidade, como tratamento de igual para igual ou de superior para inferior [...]. (CUNHA; CINTRA, 2008, p. 305)

Nesse sentido, segundo Castilho (2010):

O derivado *você* passou a ser no PB um tratamento de igual para igual. Para o tratamento cerimonioso, inventou-se o *senhor*. Em regiões brasileiras em que o tratamento *tu* continua vigente, o uso de *você* traz de volta o antigo distanciamento. Onde o *tu* não é mais encontrado, ele e seus derivados são utilizados para expressar distanciamento, como no exemplo: Olha aí o que teu filho aprontou! (CASTILHO, 2010, p.479)

No que tange à inserção do “você”, até então pronome de tratamento exclusivamente, no sistema dos pronomes pessoais do português, tem-se a consolidação de um processo de gramaticalização. Através desse processo, um item lexical passa a assumir uma função gramatical ou um item gramatical passa a assumir funções mais gramaticais ainda.

Sendo considerado um clássico exemplo de gramaticalização, o “você” chegou ao português através do processo de gramaticalização da forma “Vossa Mercê”. Contudo, segundo Lopes (2003), não houve necessariamente a perda completa dos traços categoriais originais nem a assunção definitiva de propriedades de pronome de segunda pessoa. Ainda de acordo com a autora:

Com a inserção de você no quadro pronominal do português, percebe-se a persistência da especificação original de 3ª pessoa, embora a interpretação semântico-discursiva passe a ser de 2ª pessoa [-EU]. [...] Como os processos de mudança não afetariam o sistema linguístico em sua totalidade, a implementação de você no sistema não ocorreu da mesma forma em todas as subcategorias pronominais e criou-se um paradigma pronominal que reflete um sincretismo entre a segunda e a terceira pessoa do singular. Você e tu coexistem no singular e vocês é praticamente categórico no plural na posição de sujeito, [...] (LOPES, 2003, p. 3)

Como podemos perceber, o processo de gramaticalização traz à língua uma nova função para o elemento linguístico, mas não de maneira estática ou exclusiva. As inovações que surgem são constantes, cada vez mais adaptadas às necessidades do sistema linguístico e do uso pelos falantes em qualquer modalidade da língua. Observando as modificações que aconteceram no PB em relação às formas “tu” e “você”, Sousa (2012) define três fases: Fase 1, que vai de 1870 a 1899, tendo a maior frequência do uso do “tu”; Fase 2, de 1900 até 1929, onde se encontra a plena variação entre as duas formas; e Fase 3, no período de 1930 a 1989, com a maior ocorrência do “você” em relação ao “tu”.

Especificamente no português do Brasil, o “você” passou a ser usado de maneira geral pelos falantes, diferentemente do que aconteceu no português europeu, onde esse elemento pronominal ainda é utilizado de maneira restrita, em situações bastante marcadas, segundo Lopes (2003). Outra constatação importante, conforme pontua Bagno (2012) é que:

No tratamento familiar, até meio século atrás, os filhos se dirigiam aos pais como o senhor e a senhora, enquanto estes se dirigiam aos filhos como você. Mais recentemente, com uma maior democratização e informalização dessas relações, se generalizou o emprego de você também dos filhos para os pais (sobretudo nas regiões urbanas). (Bagno, 2012, p.748)

Já no século XV, o uso generalizado da forma “vossa mercê” e suas variantes eram utilizadas pela população não aristocrática em Portugal e no início do século XVI, a forma de tratamento ‘vós’ já se encontrava em desuso,

surgindo, então, o processo de simplificação da forma “vossa mercê”. Dessa forma, temos que o português que chegou ao Brasil já possuía variantes para “vossa mercê”, enquanto forma de tratamento.

Com a inserção do “você” no sistema pronominal do português brasileiro, que, segundo trabalhos de Duarte (1993), Machado (2006) e Souza (2011), aconteceu por volta de 1930, ele já assume a posição de pronome de segunda pessoa do singular e, desde então, inicia-se um processo de variação com a forma “tu”. Sobre o uso do você, Moura Neves (2000) afirma que :

O emprego de você é muito mais difundido do que o emprego de tu, para referência ao interlocutor. Além disso, ocorre frequentemente (embora mais especialmente na língua falada), que se usem formas de 2ª pessoa em enunciados em que se emprega o tratamento você, de tal modo que se misturam formas de referência pessoal de 2ª pessoa e de 3ª pessoa. [...] Esse uso ocorre especialmente na conversação espontânea, e são abundantes os exemplos nos diálogos de peças teatrais. (Moura Neves, 2000, p. 458)

De acordo com a autora, em face ao processo de variação entre as formas “tu” e “você”, o segundo tem seu uso muito mais difundido e ainda relaciona-se com a referência de 3ª pessoa no uso cotidiano do português. Ainda, como afirmam Lopes e Duarte (2004, p. 61): “No português do Brasil [...], você já está perfeitamente integrado ao sistema de pronomes pessoais, substituindo tu em grande parte do território nacional ou convivendo com tu sem que o verbo traga a marca distintiva da chamada segunda pessoa direta”.

Buscando sintetizar os usos da forma “você”, Scherre *et al* explicam que, em termos de uso, eles podem ter:

1. uso marcado, alternando com o pronome —cê” (a exemplo da região Sudeste, em Vitória/ES e em São Paulo/SP) e/ou com o pronome —tu” (a exemplo do Rio de Janeiro/RJ; e da variedade brasiliense, em processo de focalização dialetal, na região Centro-Oeste);
2. uso mais marcado, alternando com os pronomes —cê” ou —ocê” (a exemplo de algumas áreas do interior de Minas Gerais e Goiás ou com o pronome —tu” (a exemplo da região Nordeste);
3. uso como forma de contato, logo substituído pelo pronome —tu” com concordância baixa (a exemplo do interior do Amazonas, na cidade de Tefé, região Norte; e também no Rio de Janeiro);
4. uso estranho à comunidade local,

em que predomina o uso natural do pronome —tu”, na maior parte das vezes sem concordância, que, por sua vez, pode alternar com o pronome —tu” com concordância em situação de mais formalidade, a exemplo do Rio Grande do Sul, na região Sul); 5. uso também estranho à comunidade local, em áreas rurais, com baixa frequência, em que predomina o uso do pronome “cê” (Vale do Gurutuba, no centro-norte de Minas Gerais; e Pombal, no centro de Goiás) ou dos pronomes “cê” e “ocê” (São Francisco, no norte de Minas Gerais; e Arcos, no Centro-oeste de Minas Gerais). (Scherre *et al.*, 2015, p. 170).

Outra característica interessante revelada pelo sistema pronominal do português, para além das formas pronominais como referentes do discurso, é a possibilidade da pessoa do discurso não ser explícita, dando lugar a um “pronome zero” ou sujeito nulo. Em virtude disso, o paradigma do sujeito referencial passa a ser acrescido de mais uma possibilidade de realização: o apagamento do sujeito pronominal. Conseqüentemente, no âmbito da segunda pessoa, essa possibilidade também encontra seu lugar e discorreremos sobre ela na subseção a seguir.

2.2. O SUJEITO NULO NO PORTUGUÊS BRASILEIRO

No âmbito das gramáticas tradicionais, pouco se fala sobre a noção de sujeito nulo, elas o definem como oculto, elíptico ou subentendido. Apresentam classificações de sujeito a partir das noções de sujeito simples, composto, indeterminado, oculto e ainda oração sem sujeito. Em nossa pesquisa, focamos no que a GT denomina como oculto, sendo aquele sujeito que pode ser identificado na sentença através da desinência verbal.

A Gramática de Celso Cunha (2001) define que esse sujeito seria aquele que não está explicitamente expresso na oração, mas que pode ser identificado através das marcas de desinência do verbo ou através da presença do sujeito em outra oração de um mesmo período. Já em Bechara (2004) não se observa menção a esse tipo de sujeito, sendo citado, apenas, no âmbito das figuras sintáticas, quando trata da elipse. Com isso, temos que as GTs defendem que apenas a morfologia verbal seria suficiente para indicar o sujeito, pois a pessoa do discurso já estaria contida na informação dada pelo verbo no português.

Dessa forma, a pessoa explícita através do pronome deveria ser utilizada apenas com o objetivo de dar ênfase a ela.

Para Perini (1998), em sua gramática descritiva, não seria necessário postular sujeito nulo, tendo que se é sujeito, algumas noções atraídas a essa função precisariam ser revistas, como a de que ele seria “o termo que está em relação de concordância com o verbo”. Para o autor:

Se apenas considerássemos os interesses da comunicação, diríamos que uma das fontes é dispensável; mas, é claro, a língua tem outras exigências, de modo que o sujeito explícito pode ser dispensado, mas o sufixo verbal não. [...] não é redundante porque, apesar de ter sujeito, este não aparece explicitamente; em outras palavras, temos um sujeito cujo efeito semântico é suprimido em certos casos. A representação da redundância é indireta, e na verdade não terá nada a ver com a estrutura da sentença, mas com um fato à parte, o caráter “oculto” desse sujeito. (Perini. 1998, p. 369)

O português brasileiro, através dos resultados das pesquisas variacionistas e gerativistas, aparentemente, perdeu sua propriedade que o enquadrava no conjunto de línguas de sujeito nulo, enquanto língua pro-drop. Assume-se que isso ocorreu devido à perda de morfologia verbal, deixando o paradigma de flexão enfraquecido.

Nesse sentido, conforme Duarte (1993, 1995), com o passar do tempo, de forma específica, na segunda metade do século XX, observou-se uma série de mudanças nos processos de preenchimento da pessoa pronominal na posição de sujeito.

Ao buscar apresentar a trajetória do sujeito no português, o trabalho de Duarte (1992), realizado a partir de textos de peças teatrais populares dos séculos XIX e XX, no qual analisou a realização de sujeitos plenos, pronominais e nulos, demonstrou que ao levar em consideração as pessoas do discurso, os dados revelaram a predominância do sujeito nulo em relação ao uso de “tu”, “vós” ou “você” e “senhor”. A autora destaca que a partir de 1918 inicia-se um

processo de declínio desse uso preferencial e em 1937 os dados revelam a perda do uso direto da segunda pessoa pronominal nos textos analisados.

Em específico, no que tange à predominância do uso do sujeito nulo na segunda pessoa, Duarte (1993) atribui à relação com a variação também verificada entre as formas pronominais “tu” e “vós”/ “você”/“vocês”, através da análise de peças teatrais de 1845, 1882 e 1918. Assim, a partir do momento em que se inicia a queda das formas “tu” e “vós”, em virtude do lugar dado às formas “você” e “vocês”, há uma queda considerável no número de ocorrências do sujeito nulo de segunda pessoa, especificamente nos dados das peças analisadas. Passando de 69% nas peças de 1918 a 25% nas do ano de 1937.

Para Duarte (1993), esse processo de mudança em curso que acarretou uma simplificação no paradigma flexional verbal, atrelando verbos de terceira pessoa às formas pronominais “você” e “vocês”. Duarte ainda destaca que no contexto de escrita formal mantém-se um paradigma funcionalmente rico, mas que o mesmo não ocorre no âmbito da fala espontânea.

Em seguida, em outro estudo, Duarte (1995) afirma que, no Brasil, após os processos de variação e mudança, o português teria requerido o preenchimento da posição de sujeito diante das pessoas verbais de primeira e segunda pessoa. A partir disso, temos que a relação entre a entrada de novos elementos pronominais ao sistema do português, bem como a perda de flexões verbais, revela que o uso do sujeito nulo vem sendo evitado pelos falantes que acabam por preferirem expressar sintaticamente, tanto na escrita quanto na oralidade, o elemento referencial pronominal.

Outra visão surge a partir do trabalho de Kato e Duarte (2014), estabelecendo o português brasileiro como uma língua que possui sujeito nulo parcial, regido por algumas restrições: a acessibilidade do referente e a hierarquia de referencialidade. A partir disso, temos que as maiores chances de ocorrer o sujeito pleno estão intimamente ligadas ao nível de referencialidade e de especificidade do sujeito. Com isso, temos que as formas que são referências por excelências são mais propícias ao uso do sujeito pleno, a saber, a primeira e a segunda pessoa.

A discussão acerca do preenchimento ou não do sujeito pronominal também foi tratada no trabalho de Botassini (2002) no qual a autora afirma que:

Atribui-se ao português a característica de língua que prescinde do uso do sujeito pronominal porque as desinências verbais são suficientemente marcadas para indicar o pronome sem que haja a necessidade de explicitá-lo. Entretanto, estudiosos voltados à análise desse tema têm demonstrado uma situação diversa em trabalhos, isto é, têm demonstrado que o português contemporâneo do Brasil tem preenchido cada vez mais a casa do sujeito. (Botassini, 2002, p.41)

Ainda segundo a pesquisadora, a razão para o maior preenchimento do sujeito no português encontra lugar a partir do processo de enfraquecimento da flexão verbal, como resultado da diminuição do uso de formas distintas entre singular e plural que gerou, como, por exemplo, a possibilidade de realização de sentenças como em: ele fala muito/ eles fala muito.

Dessa forma, mesmo tendo nosso foco voltado para as realizações específicas das formas “tu”, “você” e pronome nulo, sem levar em consideração a predominância, mas, sobretudo, compreendendo as possibilidades de uso legitimadas pelo sistema linguístico do PB de diversas formas de referência da segunda pessoa do singular na posição de sujeito, assumimos que as formas pronominais que preenchem explicitamente a segunda pessoa do singular, “tu”, “você”, “ce”, “ocê”, “senhor/senhora”, bem como o pronome nulo ou zero configuram um processo de variação na língua portuguesa.

A fim de comprovar tal assunção, vejamos a subseção, a seguir, que levanta uma série de estudos que se ocupam da discussão acerca da segunda pessoa do singular através do preenchimento pronominal e a subseção subsequente que apresenta os estudos acerca do não preenchimento do sujeito no português brasileiro. Apresentaremos, de forma sucinta, as pesquisas que tratam o fenômeno em variação na língua falada e escrita para traçar um panorama do atual *status* desse processo variacionista no PB.

2.3. ESTUDOS ANTERIORES: A REALIZAÇÃO DA SEGUNDA PESSOA DO SINGULAR NOS DADOS DO PB

A presente tese objetiva verificar como é realizada a segunda pessoa do singular na posição de sujeito no contexto de comunicação digital no português brasileiro, através do preenchimento explícito ou não de elementos pronominais. Assim, faz-se necessário apresentar os trabalhos que se dedicaram a esta temática, observando a língua através de dados da oralidade, a fim de averiguar se os resultados encontrados alcançam o uso no meio digital ou se possuem estruturas linguísticas distintas em virtude do meio pelo qual são realizadas. Para tal, descrevemos estudos em diversos estados, definindo como critério a diversidade e a distribuição geográfica da análise do fenômeno em variação, contemplando uma amostra de todas as regiões do país, bem como apresentamos os trabalhos que estudaram essa variação no estado de Alagoas.

2.3. 1. A variação nos dados do PB: o preenchimento pronominal do sujeito de segunda pessoa do singular

Nesta subseção, não objetivamos esgotar a apresentação de todas as pesquisas realizadas até o momento, mas sim expor uma amostra de alguns estudos sobre o processo de variação do preenchimento da segunda pessoa do singular a partir de pronomes, através de dados obtidos em trabalhos representantes de cada uma das regiões do país, buscando apresentar a realidade desse processo no PB.

2.3.1.1.As pesquisas na região Sul: Loregian-Penkal (2005) e Franceschini (2010)

Buscando discorrer sobre o preenchimento do sujeito através da alternância da segunda de “tu” e “você” em Santa Catarina, Loregian-Penkal (2005) analisou a fala de informantes de cinco cidades catarinenses. O trabalho da pesquisadora foi desenvolvido dentro do projeto VARSUL e levou em consideração as variáveis nível de escolaridade, faixa etária e gênero.

Através de sua análise, a autora aponta uma distribuição heterogênea entre as formas e afirma que levando em consideração as variáveis de ordem

social, informantes do sexo feminino e de localidades situadas no litoral, como Florianópolis e Ribeirão da Ilha, tendem ao maior uso da forma “tu”. Tais dados justificam-se pela recente ocupação dessas regiões pelos açorianos, sendo, assim um traço étnico que ainda é resguardado e, no caso da predominância da forma “tu” entre informantes mulheres deve-se ao fato de que a forma não seria desprestigiada, o que, conforme Labov explica, é uma tendência feminina a preferência pelas formas de prestígio.

Apesar de a pesquisadora focar especificamente na realização das formas “tu” e “você” enquanto variável dependente em estudo, o trabalho revela que em alguns contextos linguísticos, o pronome é apagado. Nesse caso, percebemos a possibilidade do uso do sujeito nulo, tanto com a flexão de segunda quanto de terceira pessoa nos dados apresentados na pesquisa.

Em suma, o trabalho revela que mesmo se tratando de cidades pertencentes ao mesmo estado, Santa Catarina, a distribuição do uso dos pronomes se deu de forma bem distinta, tendo maior frequência de uso do “você” na cidade de Lages e menor frequência na cidade de Blumenau. Além disso, descreveu uma distribuição equitativa entre as duas variáveis nos dados dos chapecoenses e uma maior frequência no uso do “tu” com flexão de segunda pessoa nas cidades litorâneas analisadas, Florianópolis e Ribeirão, dando um caráter bem heterogêneo a essa variação no estado, diferentemente do que acontece em outros estados como Rio Grande do Sul, com maciço uso de “tu” e no Paraná, com mais expressividade do “você” (Lorengian-Penkall, 2004).

Lorengian-Pekal (2005) ainda argumenta que apesar da máxima de plena substituição do “tu” pelo “você” enquanto pronome de segunda pessoa no português do Brasil, nas cidades analisadas isso não foi confirmado. Alerta, também, para a necessidade de estudos que façam uma reflexão sobre a língua a partir de estudos dialetais que possam analisar esses pronomes através de dados reais nas mais diversas localidades.

Já Franceschini (2010) analisou o impacto das variáveis sociais na variação entre os pronomes tu/você na cidade de Concórdia em Santa Catarina,

através de um *corpus* constituído por 12 entrevistas e considerando como variáveis independentes a faixa etária, o nível de escolaridade e o sexo do informante.

A partir da descrição dos dados, a autora apontou que o uso do pronome você foi de 48%, enquanto o pronome tu representou 52% da amostra. O programa utilizado para tratamento dos dados foi o Varbrul, que mostrou que as variáveis faixa etária e escolaridade foram consideradas significativas.

Na pesquisa, nos falantes com 50 anos ou mais o uso do tu foi favorecido e nos falantes mais jovens o inverso aconteceu. A autora, por fim, concluiu que a predominância do “você” nos dados dos informantes mais jovens pode sinalizar um processo de mudança linguística em tempo aparente.

2.3.1.2. As pesquisas na região Sudeste: Mota (2008), Silva e Gonçalves (2016) e Modesto (2006)

Na região Sudeste, damos destaque ao trabalho de Mota (2008) que analisou a variação “tu” e “você” e em São João da Ponte, Minas Gerais. A autora desenvolveu sua pesquisa a partir de uma amostra composta por 24 informantes (femininos e masculinos) e levou em consideração para além do gênero, o nível escolaridade e faixa etária como variáveis sociais.

A pesquisa teve seu *corpus* constituído a partir de entrevistas e de situações naturais de uso da língua. Uma estratégia utilizada pela pesquisadora foi a entrevista com temas familiares e não familiares aos entrevistados. A partir desses procedimentos metodológicos, os resultados apontaram para uma utilização maciça da forma “você”, 90%, em detrimento do pronome “tu”, 10%. Apesar do pequeno valor percentual, Mota (2008) investigou os fatores de ordem linguística, social e histórica que justificariam a ocorrência do “tu” no dialeto de São João da Ponte, tanto na posição de sujeito, quanto na posição de objeto, que contou com maior frequência de uso.

Para tal, elencou os fatores linguísticos função sintática da variante, tipo de frase e tempo verbal; além dos seguintes fatores extralinguísticos: idade, sexo, grau de intimidade entre os falantes, tipo de discurso, área geográfica de

atuação profissional do informante. Por fim, o estudo revelou que a faixa etária de informantes jovens (15 - 25 anos) e a localidade da zona rural são fatores condicionantes ao uso da forma “tu” e comprovou que o processo de variação entre “tu” e “você” está atrelado diretamente ao estilo formal e/ou informal e ao grau de intimidade entre os falantes.

Silva e Gonçalves (2016) desenvolveram seu estudo no município de Ressaquinha em Minas Gerais, Sudeste do país, a partir de dados de 26 informantes divididos em zona rural e zona urbana, focando especificamente na realização de “tu” ou “você”. O objetivo era identificar a procedência da realização do uso da forma “tu” em um estado que a maioria dos dados já colhidos revelam a preferência pela forma “você”, como o trabalho de Mota (2008) e Scherre *et al* (2015).

Através das análises dos dados, os autores confirmaram a existência da variação entre “tu” e “cê” e pontuam que há pouco uso da forma “você”. A esse resultado, são somadas as constatações de que o fenômeno ocorre tanto na zona rural quanto na urbana. Outro dado relevante é que os homens utilizam majoritariamente o “tu”, com 51 ocorrências dentro do *corpus* analisado e as mulheres registram apenas 36 ocorrências.

Os autores destacam que a comprovação da presença da forma “tu” na cidade de Ressaquinha é uma grande descoberta, tendo em vista que, de um modo geral, a fala mineira, como já mencionado aqui anteriormente, é caracterizada pelo forte uso do “você” na posição de sujeito e na de objeto. Um outro grande achado da pesquisa é que a variável dependente é composta pelas formas “tu”, “você”, “ocê” e “ce” nessa comunidade linguística. O trabalho de Silva e Gonçalves (2016) também revelou a presença dos correferentes de segunda pessoa (te, ti, contigo, teu e tua) que evidenciam o quanto é produtivo o uso do “tu” na cidade de Ressaquinha-MG.

Os dados orais colhidos na cidade de Santos, em São Paulo, foram objeto de estudo no trabalho de Modesto (2006). A pesquisa contou com dados obtidos através de 20 gravações e, para verificar o processo de variação entre “tu” e “você”, levou em consideração gênero, faixa etária e escolaridade como fatores

sociais; e função sintática da forma de tratamento, referenciação e monitoramento, enquanto fatores discursivo-pragmáticos, como variáveis linguísticas a serem analisadas.

Em linhas gerais, o trabalho de Modesto (2006) aponta que a preferência dos santistas é pela forma “você” em detrimento do uso da forma “tu” enquanto forma pronominal para o preenchimento da função de sujeito. O pesquisador destacou que quanto maior o grau de escolaridade, menor o uso do “tu”. O sexo do informante não foi significativo para distinção entre os usos, pois ambos revelam o maior uso do “você”. Sobre a referenciação, os dados revelam que, independentemente de direta ou indeterminada, o “você” é a forma mais utilizada. O grau de monitoramento foi condicionante para maior ou menor número de ocorrência das formas pronominais, sendo maior a utilização do “tu” em contextos de fala não monitorados. O autor destaca que os dados demonstram que quanto à expressividade, enquanto variável discursiva, o “tu” foi mais recorrente em enunciados com maior expressividade enquanto o “você” é favorecido em enunciados de menos expressividade.

2.3.1.3. As pesquisas na região Centro-Oeste: Scherre *et al.* (2011) e Andrade (2010)

Um grande referencial de pesquisa no Centro-Oeste brasileiro sobre a temática que tratamos até aqui é o trabalho de Scherre *et al.* (2011). Nele, os autores analisam dados colhidos em Brasília, nos anos de 2006 e 2007, através de uma compilação de vários trabalhos.

A pesquisa observou a variação entre “tu” e “você” em falantes de localidades distintas e levou em consideração o sexo e a faixa etária do informante. Diante disso, revelou que na Ceilândia, onde o quantitativo de filhos de nordestinos é maior, o uso do “tu” era consideravelmente maior do que no Plano Piloto, onde a influência da ocupação nordestina não foi expressiva. Aponta, ainda, que nas relações simétricas há um maior percentual de uso da forma “tu” em detrimento de “você”. A pesquisa conclui também que o “tu” brasiliense é tido como supra-regional, sem apresentar concordância de segunda pessoa.

Andrade (2010) também se dedicou a estudar as formas pronominais de segunda pessoa no Centro-Oeste e para isso, levou em consideração a variação entre as formas “você”, “cê” e “tu”, através de 43 informantes, divididos de formas distintas em quatro “regiões” de Brasília: Planalto, Lago Azul, Asa Norte e Sudoeste. No total das ocorrências analisadas, o “tu” mostrou-se mais produtivo, 48%, do que as variantes “você” e “ce” que empataram com percentual de 26%. Importante mencionar que os dados coletados também revelaram a utilização de pronomes nulos em mais de 120 ocorrências, contudo, foram descartadas e não foram consideradas no estudo. Os dados revelam a grande frequência de uso do “tu” entre as crianças e justificam, conforme também justifica por Scherre *et al.* (2011), como já mencionado, pela origem nordestina de seus familiares. A origem dos pais também mostrou-se um importante argumento para uso da forma “ce”, bastante utilizado pelos filhos de pais de origem mineira.

A partir desses resultados, a autora argumenta em torno do nascimento de um falar específico de Brasília, com crescente uso do pronome “tu” sem a concordância verbal correspondente. Defende, ainda, que essa variedade brasiliense já possui algumas características linguísticas em termos lexicais e gramaticais bem definidas e que a contribuição nordestina teria sido fundamental para isso, não apenas para os falantes filhos de migrantes nordestinos, mas pelas relações estabelecidas linguisticamente no uso linguístico cotidiano da cidade.

2.3.1.4. As pesquisas na região Norte: Costa (2013) e Martins (2010)

Subindo o mapa geográfico, rumo aos dialetos do Norte, temos a investigação proposta por Costa (2013) que, a partir das cidades de Boa Vista, Belém, Macapá, Manaus, Rio Branco e Porto Velho, delineia um panorama sobre o uso de “tu” e “você” como preenchedores da função de sujeito na fala das capitais nortistas. O autor investigou as duas formas pronominais através de dados de 48 informantes, divididos em 8 por cidade, com base no projeto Atlas Linguístico do Brasil, via entrevistas de fala espontânea. Como fatores linguísticos observados, o autor elencou a explicitação do pronome, concordância verbal e tempo verbal; no âmbito dos fatores extralinguísticos, a pesquisa observou escolaridade, faixa etária, localidade e gênero do informante.

Após a análise, o autor conclui que o processo encontra-se em uma variação estável, tendo predominância do “tu” nas cidades de Manaus, Boa Vista e Porto Velho e nas cidades de Macapá, Belém e Rio Branco a predominância da forma “você”. O trabalho apontou que as variáveis gênero e faixa etária não foram relevantes na escolha de uma ou outra variável dependente e que o fator tempo verbal foi condicionador da escolha pelo pronome “tu” com sentenças no tempo presente e do “você” no pretérito; a explicitação do pronome demonstrou a preferência da forma “tu” com concordância de segunda pessoa; e o modo verbal no imperativo favoreceu o uso do “tu” com flexão canônica de segunda pessoa.

Outro importante trabalho sobre esse processo de variação é do Martins (2010). Nele, se faz a análise da variação entre as formas “tu”, “você” e “senhor” a partir de dados de entrevistas de trinta informantes nascidos na cidade de Tefé, no estado do Amazonas. Para tal, o autor recorreu aos fatores escolaridade, faixa etária e gênero. Apesar do interesse inicial residir no estudo entre “tu”, “você” e “senhor”, a utilização do pronome zero foi diagnóstica como uma considerável variante. Do ponto de vista das variáveis extralingüísticas, foram observadas o tipo de referência (genérica e específica), o tipo de discurso (direto e relatado) e o paralelismo. O tipo de simetria na relação entre os interlocutores da pesquisa também foi observado, levando em consideração as relações de forma simétrica e assimétrica.

Após a análise dos dados, o autor conclui que o pronome “tu” é o preferido entre os falantes de Tefé e que é pouco observada a concordância verbal canônica. O trabalho também revela que é maior a freqüência do uso da forma “você” em relação à forma “senhor”. Apesar da alta freqüência do uso de pronomes zero, o autor os desconsiderou na análise. Em termos de relevância para uso de uma ou outra variante, a variável paralelismo obteve destaque quando se observa os dados com o pronome “tu” e o estudo também destaca a importância do contexto de maior intimidade para realização do uso do “tu”.

2.3.1.5. As pesquisas na região Nordeste: Alves (2010), Soares (1980) e Nascimento (2017)

Já observando os estudos que trabalham com dados do português falado no Nordeste, analisando o estado do Maranhão, Alves (2010) buscou verificar o uso das formas “tu” e “você” a partir das variáveis naturalidade, faixa etária, sexo e escolaridade. Como fatores de ordem linguística, a pesquisadora definiu concordância, tipo de referência e tipo de relato, através de um *corpus* constituído com dados do Atlas Linguístico do Maranhão, AliMA, projeto desenvolvido na Universidade Federal do Maranhão. Partindo do pressuposto de que no falar maranhense havia uma maior preferência pelo uso da forma “tu”, a pesquisadora iniciou sua investigação e após análise dos dados, constatou que, diferentemente do que era imaginado e inclusive difundido no senso comum, o percentual de 61,6% de uso foi da forma “você”.

Em relação às variáveis elencadas no trabalho, a naturalidade favoreceu a ocorrência do “tu” entre os informantes da zona rural e a do “você” nas áreas urbanas. Um outro fator que se mostrou condicionador nesse processo de variação entre as cidades pesquisadas no Maranhão foi a idade, tendo entre informantes mais jovens a prevalência da forma “tu” e os mais idosos tiveram maior uso da forma “você”. Contrariando a hipótese inicial da pesquisadora, o fator sexo não demonstrou ser condicionador da variação entre as formas nos dados analisados. E o nível de escolaridade mostrou-se como mais relevante nesse processo de variação.

No âmbito das variáveis linguísticas observadas no trabalho, Alves (2010) concluiu que apenas o fator tipo de relato enquanto discurso relatado próprio atuou diretamente nesse contexto de variação no português falado no Maranhão, trazendo o uso do “tu” com mais expressividade, em detrimento de relato de terceiros que recorriam ao uso do “você”. Em suma, os dados arrolados na pesquisa demonstram estatisticamente o uso do “você” como predominante e os fatores condicionantes idade e localidade são os mais influentes nesse processo de variação, bem como o fator linguístico relato revelou ser determinante no uso das formas estudadas.

Já na cidade de Fortaleza, no Ceará, o trabalho de Soares (1980) investigou o uso das formas “tu”, “você” e “senhor” a partir dos dados obtidos através de dois tipos de questionários e de conversas informais, totalizando 72

informantes que foram divididos em sete células, a partir das variáveis sexo e escolaridade.

Um achado importante revelado pela pesquisa é de que em contextos de relações assimétricas há variação entre “você” e “senhor”, e o “tu” é preferido quando o interlocutor representa uma classe inferior na relação; já nas relações simétrica há uma plena variação entre as três formas estudadas. Nesse sentido, o autor afirma que a situação discursiva, o papel social dos interlocutores, a faixa etária e o grau de intimidade são os condicionantes do processo de variação na fala em Fortaleza. O trabalho também revela que há omissão do sujeito pronominal que acaba por ser compensada pelo uso de vocativo. Por fim, destaca que o “você” possui o maior número de ocorrências nas situações mais formais e que o “tu” é utilizado de forma generalizada e que variáveis como nível de escolaridade, grau de formalidade e atenção à fala condicionam o uso ou não da concordância.

Já no estado da Bahia, Nascimento (2017) analisou os fatores condicionadores do uso de “tu” e “você”, baseando-se nos inquéritos do Programa de Estudos do Português Popular Falado de Salvador e de inquéritos gravados na comunidade de Amargosa. Como variáveis linguísticas, a autora selecionou tipo de frase, tipo de discurso e tipo de referência; a faixa etária e a escolaridade dos informantes foram selecionadas como variáveis sociais.

O trabalho apontou para o uso expressivo da forma “você”, com 92,2% de ocorrências e 7,8% de ocorrências da forma “tu”, sendo categórico o uso de “você” em Salvador e apresentando variação entre as formas de segunda pessoa na cidade de Amargosa. Nos dados colhidos na capital baiana, a autora explica que encontrou uma possível mudança em curso, tendo em vista a totalidade de uso da forma “você”, apresentando-se enquanto a forma de prestígio em detrimento da forma “tu”. Já na análise dos resultados encontrados no município de Amargosa, tendo como dado relevante o uso predominante da forma “tu” quando observada a variável tipo de discurso nos casos de relato próprio, com 62,5% de ocorrências.

A pesquisadora conclui que a forma pronominal “tu” demonstrou ser muito mais frequente na cidade de Amargosa, apesar de nem sempre ter sido possível registrar nas gravações, pois ao referirem-se a pessoas próximas, os informantes demonstravam o maior uso do “tu”, comprovando que o grau de intimidade entre os falantes é um dos fatores motivadores das formas pronominais de segunda pessoa. A não completa amostragem desse dado, segundo a autora, deve-se ao método da entrevista que pode ter influenciado no automonitoramento da fala dos informantes.

2.3.1.5.1. As pesquisas em Alagoas: Vitório (2018), Silva e Vitório (2019) e Silva (2019).

A variação entre as formas de segunda pessoa do singular na posição de sujeito ainda é pouco estudada no estado de Alagoas. Sendo assim, esta pesquisadora encontrou, até a conclusão da escrita desta tese, apenas duas autoras que dedicaram atenção a esse fenômeno no estado, a saber, Vitório (2018), Silva e Vitório (2019) e Silva (2019).

O trabalho de Vitório (2018) buscou estudar a variação tu/você na posição de sujeito na fala de maceioenses, através de uma amostra constituída por 72 entrevistas sociolinguísticas. A pesquisadora coletou seus dados em 2010 e para fins de estratificação, selecionou as variáveis sexo, faixa etária e escolaridade. As entrevistas revelaram o uso preferencial da forma “você”, com 98% das ocorrências e apenas 2% de uso da forma “tu”, demonstrando uma regra semicategórica dentro desse processo de variação.

Além dos dados colhidos nas entrevistas, desenvolveu um teste e o aplicou a 46 informantes que nasceram e foram criados na capital do estado alagoano. Considerando situações hipotéticas ambientadas na cidade, vivenciadas por um personagem fictício nomeado de João e caracterizado enquanto um estudante universitário, objetivou representar relações simétricas, assimétricas descendentes e assimétricas ascendentes. Através da realização do teste desenvolvido, a autora chegou aos seguintes dados: 65% de uso do “você”, 16% da forma “tu”, 5% de “CE” e 14% de outras formas, incluindo o pronome zero e a forma “senhor”.

Observando os dados da pesquisadora, revelou-se que no âmbito das relações simétricas, há uma notável preferência pelo “tu” na fala dos informantes de Maceió, diferentemente dos dados encontrados através da entrevista realizada, justificando a importância dos aspectos de maior intimidade e familiaridade como condicionantes da forma canônica “tu”. O trabalho de Vitório (2018) revela que o “você” pode funcionar como uma espécie de elemento coringa nas relações simétricas, assimétricas descendentes e assimétricas ascendentes, tendo seu número de ocorrências expressivamente maior em relação às outras formas. A autora destaca que nas relações assimétricas descendentes o pronome zero e as formas senhor ou senhora também foram encontrados. Por fim, pontua que apesar da maior frequência de uso do “você”, em contextos de mais intimidade, o “tu” é mais recorrente, justificando que ambas as formas justificam seu uso a parte do maior grau de intimidade e proximidade entre os informantes de Maceió.

Outro trabalho encontrado, que resultou no trabalho de dissertação de Silva (2019) analisou a variação entre as formas pronominais de segunda pessoa na cidade de Coité do Nóia, localizada no Agreste alagoano, através de dados de 36 informantes. Concordância verbal, paralelismo pronominal, tipo de relato e tipo de referência foram definidos como fatores de ordem linguística para o estudo e; sexo e faixa etária como fatores de ordem social. As relações entre os diálogos oriundos de relações simétricas e assimétricas também foram consideradas no estudo.

Após a análise dos dados através do sistema GoldVarbX, a autora descreve que na comunidade pesquisada há a variação entre “tu” e “você”, tendo 11% de ocorrências da primeira forma e 89% para as formas “você ou cê”. No que tange às variáveis sociais, a relação de idade entre os informantes mostrou-se significativa nos diálogos entre jovens-jovens, gerando elevação no uso do “tu” e nos diálogos entre adulto-adulto houve a predominância do uso do “você”.

Em termos estatísticos, a autora afirma que as variáveis mais significativas foram, paralelismo pronominal, sexo e faixa etária, tendo, respectivamente, o uso com maior frequência dos pronomes “tu” e “você” nos

contextos em que eles eram antecidos pelo mesmo pronome; nas relações entre os informantes do mesmo sexo, o uso do “tu” prevaleceu e nas relações de sexos diferentes, o “você” teve maior prevalência; o pronome “tu” predominou entre os informantes mais jovens e o “você” na fala dos informantes com mais idade.

A variável sexo não foi considerada como significativa na análise e as variáveis tipo de referência, tipo de relato, tipo de relação entre os informantes e as relações simétricas e assimétricas apresentaram nocaute⁴. Como pode ser observado, a análise demonstra a presença do processo de variação entre as formas de segunda pessoa na posição de sujeito em território alagoano, mas restringe-se a análise de uma única cidade, conforme objetivo da autora.

Ambas as pesquisadoras aqui já mencionadas desenvolveram uma outra pesquisa no estado alagoano. Dessa vez, a pesquisa de Silva e Vitória (2019) investigou a realização de “tu”, “você” e “ocê” na posição de sujeito no sertão de Alagoas, através de dados do Projeto Lusa (A língua usada no sertão alagoano) com 96 entrevistas e que levou em conta as variáveis extralinguísticas sexo/gênero, faixa etária e nível de escolaridade.

As autoras demonstram uma baixa frequência de realização do “tu”, com apenas três ocorrências nos dados investigados. Tal fato, fez com que o trabalho tenha enveredado para análise específica das formas “você” e “cê”, que tiveram respectivamente 473 e 33 realizações. Através da análise estatística realizada no GoldVarbX, detectaram que há a variação entre “você” e “cê” entre os sertanejos de Alagoas, mas a predominância de uso ficou resguardada à forma “você”, com 94% e apenas 6% de uso da variante “cê”. Sobre essa maciça preferência por “você”, as autoras alegam a possibilidade do “cê” ainda ser uma forma não foi completamente implementada, estando em processo, ainda, para uso efetivo na comunidade do sertão de Alagoas.

Como condicionantes do uso de uma ou outra forma em variação, o sistema utilizado para tratamento dos dados considerou significativamente

⁴ Temos por nocaute “um fator que, num dado momento da análise, corresponde a uma frequência de 0 a 100% para um dos valores da variável dependente” (GUY; ZILES, 2007, p.158)

apenas os fatores escolaridade e paralelismo formal. Através deles, evidenciaram que independente do nível de escolaridade, o “você” é mais frequente, contudo, quando do uso do “cê” o maior número de ocorrências reside entre os informantes de nível médio. O paralelismo formal demonstrou maior percentual do “você”, quando antecedido pela mesma forma e, ainda, o trabalho demonstrou que o maior percentual de “cê” também se deu quando antecedido por ele mesmo.

2.3.2. Uma visão geral: a proposta dos subsistemas de Scherre *et al* (2015)

Buscando fazer um mapeamento dos dados obtidos através de pesquisas sociolinguísticas sobre o processo de variação pronominal no preenchimento da segunda pessoa na posição de sujeito, com as formas “tu” e “você”, Scherre *et al* (2015), definem 6 subsistemas para o português brasileiro, apresentando o uso e a predominância de uma ou outra forma, como podemos observar a seguir:

1. VOCÊ: uso exclusivo das formas você/cê/ocê;
2. TU com concordância baixa: predomínio de TU acima de 60% com concordância abaixo de 10%;
3. TU com concordância alta: predomínio de TU acima de 60% com concordância entre 40% e 60%;
4. TU/VOCÊ com concordância baixa: uso médio de TU abaixo de 60% com concordância 10%;
5. TU/VOCÊ com concordância média: uso médio de TU abaixo de 60% com concordância entre 10% e 39%;
6. VOCÊ/TU sem concordância: uso de TU de 1% a 90%. (SCHERRE *et al*, 2015, p. 138).

Scherre *et al.* (2015) chegaram a esse resultado através da compilação de mais de 29 mil dados, oriundos de 60 amostras de fala obtidas através de entrevistas, conversas naturais, entrevistas geolinguísticas e em conversas uma amostra de conversa instigada por imagens.

De forma sistemática, a proposta partiu da análise de pesquisas que utilizaram prioritariamente a situação de entrevista sociolinguística, com a presença de um ou dois entrevistadores, com o objetivo de capturar a fala de

forma espontânea, entretanto, por não se tratar de um conjunto homogêneo de dados, alguns problemas foram detectados quando comparados e sistematizados.

Com base nisso, de acordo com Scherre *et al.* (2015, p. 135), “essas entrevistas podem não refletir toda a realidade dos usos dos pronomes de segunda pessoa, por, muitas vezes, evitarem constantes trocas de turno e/ou por não terem naturalmente toda a gama das relações interacionais”). As autoras destacam que apesar das possíveis dificuldades para registrar os pronomes de segunda pessoa, é visível que nos locais em que o pronome “tu” não é a forma preferencial da comunidade linguística, captá-lo para fins de registro e análise é ainda mais difícil, como em Brasília e no Rio de Janeiro. Destacam ainda que em lugares em que o “tu” é preferencial, sua captura é facilitada até mesmo em entrevistas, como pode ser visto no Rio Grande do Sul, por exemplo.

Para a realização do mapeamento, as formas “tu”, “você”, “ocê” e “cê” foram utilizadas, na perspectiva de comprovar que estes elementos pronominais não possuem o mesmo *status* em relação ao grau de naturalidade ou do tipo de simetria (simétrica/assimétrica), concluindo que qualquer forma pronominal de segunda pessoa pode ser empregada de forma distinta de acordo com a localidade, como demonstram as pesquisas até então.

Buscando atualizar a proposta de distribuição das formas pronominais de segunda pessoa do singular especificamente na região Nordeste, Scherre *et al* (2021) e traz algumas ponderações sobre os subsistemas definidos em Scherre *et al* (2015) e explora os aspectos de ordem interacional a depender do tipo de coleta de dados utilizado, por meio de análises recentes, realizadas até 2020. Assim, oferece um novo mapeamento e distribui as formas de segunda pessoa do singular do português falado nas capitais nordestinas da seguinte forma:

Tabela 02: Distribuição dos pronomes de segunda pessoa no PB

Tabela 1 – Média da distribuição de pronomes explícitos na segunda pessoa do singular no português brasileiro falado nas capitais dos estados do Maranhão (MA), Piauí (PI) e Ceará (CE), região Nordeste: médias de amostras diversificadas detalhadas nas Tabelas 1a, 1b e 1c

Capitais	Amostra/Projeto	Você	Cê	Ocê	tu sem concordância	tu com concordância	100%/ Total
São Luís, capital de Maranhão	Quatro amostras: Alima, Alib, Carneiro, Alves	21% (303)	2% (32)	0% (0)	66% (954)	11% (154)	(1.545)
Teresina, capital do Piauí	Uma amostra: Alib	92% (49)	?	?	8% (4)		(53)
Fortaleza, capital do Ceará	Três amostras: Alib, Soares, Norporfor	51% (925)	1,8% (33)	0,3 (1)	45% (816)	1,9% (35)	(1.810)
Natal, capital do Rio Grande do Norte	Dois amostras: Alib, Banco Conversacional de Natal	85,5% (370)	?	?	14,3% (62)	0,2 (1)	(433)
João Pessoa, capital da Paraíba	Dois amostras: ValPB, Alib	96% (2064)	?	?	3% (62)	1% (16)	(2.144)
Recife, Capital de Pernambuco	Uma amostra: Alib	86% (100)			14% (16)		(116)
Maceió, capital de Alagoas	Três amostras: Alib, Vitória 2010, Vitória 2017	90% (856)	2% (21)	?	8% (80)	0 (0)	844
Aracaju, capital de Sergipe	Uma amostra: Alib	96% (133)	?	?	4% (5)	?	(138)
Salvador, capital da Bahia	Três amostras: Alib, NURC/SSA, PEPP	99% (1.107)	?	?	1% (9)	Não há	(1.116)

Fonte: Scherre *et al* (2021, p 176)

Conforme apresentado, ao analisarem os dados das amostras elencadas, no âmbito do preenchimento do sujeito pronominal, as autoras definem que é possível verificar quatro blocos de manifestações distintas na região Nordeste, considerando as médias das capitais. Assim, Scherre *et al* (2021), propõe uma subdivisão:

Grupo 1- Equivalente a cidade de São Luís, capital do Maranhão, possui um sistema mais vigoroso de você ~ tu com concordância ~ tu sem concordância, ternário ou talvez binário;

Grupo 2- Correspondente às capitais Teresina-PI e Fortaleza-CE que manifestam você e tu preferencialmente sem concordância;

Grupo 3- Composto por Natal- RN, João Pessoa-PB, Recife-PE, Maceió-AL e Aracaju-SE que registram o uso preferencial pela forma pronominal você, mas possui, especificamente em relação ao uso do tu, registro de 16% de uso em Natal e, em Recife, registro de 14%.

Grupo 4- Representado pela capital baiana, Salvador, com status de capital da forma você no Nordeste brasileiro, que resguarda apenas 3% de tu nas amostra analisada. Vislumbra-se ainda a inserção futura da forma cê, tendo em vista que as formas devem variar de acordo com os aspectos de simetria entre os pares e não pares.

Como bem apresentado pelas autoras, é verificado uma nova distribuição das formas pronominais nas capitais nordestinas e existe um processo de regularidade no uso marginalizado da forma cê e sem registros de ocorrência da forma ocê.

Diante dos resultados apresentados nesta seção, buscando evidenciar a complexidade do sistema pronominal, em especial, as formas de realização explícita da segunda pessoa do singular, a partir do suporte dados pelos estudos anteriores aqui destacados, em nossa análise, almejamos descrever e analisar como se dá esse preenchimento pronominal bem como a ausência dele, entendendo esse binômio como um fenômeno variacionista, ampliando os dados e a abrangência da análise, trazendo os dados do português brasileiro coletados no ambiente digital do uso da língua.

A seguir, vejamos as pesquisas que se ocuparam com o não preenchimento dos pronomes na função de sujeito, focando na realização do sujeito nulo no português brasileiro.

2.4. ESTUDOS ANTERIORES: O SUJEITO NULO NO PORTUGUÊS BRASILEIRO

Como buscamos analisar o processo de variação entre o preenchimento e o não preenchimento do sujeito pronominal da segunda pessoa do singular, através do sujeito nulo, nesta subseção, apresentaremos alguns estudos que fazem essa discussão através das diferentes pessoas pronominais e destacaremos, em específico, os resultados encontrados no âmbito da segunda pessoa.

2.4.1. Duarte (1995)

Através do suporte teórico da Sociolinguística Variacionista, Duarte (1995) analisou a utilização do sujeito exposto por pronome pleno e do sujeito nulo através de quase 1800 dados. A autora investiga as ocorrências de pronomes pessoais sujeitos em contextos em que tradicionalmente seriam omitidos em favor da possibilidade de referência da pessoa do discurso através da desinência da forma verbal, tendo em vista que historicamente, como já mencionado anteriormente, o português brasileiro evitava o uso excessivo de pronomes pessoais na função de sujeito.

Duarte objetiva avaliar como esse princípio de evitar pronomes pessoais sujeitos mudou ao longo do tempo no português brasileiro, observando em quais contextos essa omissão é mais recorrente e como os fatores linguísticos e sociais afetaram esse processo de mudança.

A autora destaca que do conjunto de dados analisados, apenas 29% apresentam o sujeito nulo, enquanto o sujeito constituído por pronome pleno compreende um total de 71% das ocorrências. Sobre isso, Duarte *op cit* (p. 4) afirma que “embora não se possa dizer que perdemos a possibilidade de omitir o sujeito, observa-se clara preferência pelo uso da forma pronominal plena”.

Os dados analisados indicam que a mudança linguística ocorreu de forma mais acentuada na segunda pessoa do discurso, tendo em vista a substituição de “tu” e “vós” por “você” e “vocês”. Em relação às demais pessoas, a autora destaca que as ocorrências de sujeito nulo são cada vez menos recorrentes na primeira pessoa e que a terceira pessoa ainda é resistente ao processo de mudança quando possui um SN antecedendo. Para a autora, os dados revelam que essa mudança gradual do preenchimento do sujeito reverberou na perda do “Princípio Evite Pronome”.

Ao passo que analisa os aspectos linguísticos, aponta que as orações subordinadas que possuem sujeito correferente em relação às pessoas do discurso condicionam ao sujeito nulo. Já em relação ao tempo verbal, destaca que o pretérito imperfeito e o presente são os que mais resguardam as marcas

flexionais dos verbos; os contextos de pretérito perfeito do indicativo são os que mais apresentam sujeito nulo e, por fim, em relação ao tempo verbal, todos os tempos do modo subjuntivo desfavorecem o sujeito nulo.

Em síntese, no importante e referenciado estudo realizado, Duarte (1995, p. 29 e 30) destaca que, no português brasileiro “o sujeito nulo já não se encontra em distribuição complementar como pronome pleno. Ele é antes uma opção que se realiza cada vez menos em favor deste, cuja ocorrência, em momento algum, compromete a aceitabilidade de uma sentença”. Especificamente em relação à segunda pessoa, a autora afirma que o sujeito nulo sofreu uma queda em relação ao seu uso ao longo do tempo, sendo no século XIX praticamente categórico, para representar um percentual de ocorrências inferior a 25% no século XX.

2.4.2. Farias e Martins (2021)

Buscando analisar o processo de preenchimento do sujeito pronominal no Amazonas, Farias e Martins (2021) fazem um recorte temporal através de 131 anúncios da primeira década e da penúltima década do século XX, veiculados pelo Jornal do Commercio, considerado como o mais antigo e que ainda circula no estado, na tentativa de perceber indícios de um processo de mudança em curso em relação ao processo de preencher ou não o sujeito pronominal.

Para tal, selecionou as variáveis pessoa do discurso, forma de realização do pronome, tempo verbal, forma nominal do verbo, animacidade do sujeito e recorte temporal. Os dados foram tratados e analisados através do tratamento estatístico do Goldvarb X.

Os resultados revelaram, em primeira instância, que das 607 ocorrências analisadas, o preenchimento do sujeito ocorreu em apenas 23%, enquanto o não preenchimento chegou a 77% do valor total das ocorrências, o que distingue dos resultados apontados em diversos trabalhos como os de Duarte (1993, 1995, 2003), por exemplo, e que refuta a hipótese de mudança no PB em direção à classificação enquanto língua de sujeito preenchido. Sobre isso, os

autores argumentam que pode ser motivado devido ao apelo comercial tendo em vista que o jornal, através de seus anúncios publicitários, preza pelo apelo comercial, trazendo resistência ao processo de mudança paramétrica de um sujeito nulo para um preenchimento de sujeito pronominal.

Seguindo com a análise, afirmam que o Goldvarb detectou as variáveis recorte temporal e pessoa do discurso como as que mais condicionam o sujeito nulo no *corpus* em estudo. Em suma, no que tange à variável recorte temporal, os resultados confirmam a hipótese dos autores, tendo os dados do início do século XX com maior utilização do sujeito nulo em relação aos dados da penúltima década do mesmo século. Contudo, analisando especificamente os dados do segundo período analisado, tem-se a constatação de um processo de mudança mais lento, saindo de uma perspectiva de sujeito nulo total para parcial, não indicando, portanto, uma mudança categórica para o preenchimento do sujeito pronominal.

Para os autores, não há como confirmar um processo de mudança nesse fenômeno, mas, sim, indícios de que no português coexistem aspectos de línguas pro-drop e características verbais que não permitem a identificação da referência do sujeito, como a pessoa do discurso, que reverberam em uma língua de sujeito nulo parcial, pelo menos em relação ao gênero textual investigado.

Sobre a pessoa do discurso, os dados revelam que o sujeito nulo é condicionado de forma intensa pela 1ª pessoa. Em relação à 2ª pessoa, apesar dos poucos dados encontrados, apontam para o preenchimento do sujeito pronominal, principalmente nos dados referentes ao segundo período analisado, atribui-se tal constatação ao processo de simplificação das desinências morfológicas do verbo que misturam a 2ª e a 3ª pessoa. Por isso, seria necessário o uso do pronome explícito como uma marca distintiva entre eles. Os autores, acrescentam, ainda que a 3ª pessoa condiciona o sujeito nulo nos dois momentos temporais analisados, com ênfase para a predominância da 3ª do plural no primeiro período e da 3ª do singular no segundo.

Em suma, comparando os dados dos dois períodos estudados, foi verificado que, no gênero analisado, ainda é observado um processo gradual de mudança, tendo o sujeito nulo, ainda, seu lugar reservado no português, não sendo substituído plenamente pelo sujeito pronominal explícito.

2.4.3. Silva (2021)

Com o objetivo de analisar a omissão e a expressão de sujeito no português atual nos contextos de sentenças negativas simples canônicas, Silva (2021) analisa um corpus do projeto “LínguaPoa”, constituído a partir de oito entrevistas orais realizadas entre os anos de 2015 e 2018, na cidade de Porto Alegre, no Rio Grande do Sul.

Ao analisar os dados que totalizam 590 ocorrências, a autora encontrou 53% de uso do sujeito nulo e 47% de sujeito pleno. Nessa amostra, 542 ocorrências eram no singular e faziam referência à 1ª, 2ª ou 3ª pessoa do discurso, as outras 48 ocorrências correspondiam ao plural nas três pessoas do discurso.

A 1ª pessoa do singular despontou como o maior número de ocorrências, tendo a mesma distribuição tanto enquanto sujeito nulo como de sujeito pleno e isso se deve ao fato da coleta de dados ter sido realizada através das entrevistas, com perguntas solicitando o ponto de vista do informante sobre algum tema.

No caso da segunda pessoa do singular, a autora encontrou basicamente o mesmo número de ocorrências entre as formas de sujeito pleno e sujeito nulo, sendo 11 e 12 respectivamente. A terceira pessoa também apresentou ocorrências de sujeito nulo tanto no singular quanto no plural, totalizando, respectivamente 74 ocorrências de nulo e 35 de sujeito pronominal explícito.

Observando fatores de ordem semântica, também observados na análise, tem-se que a noção de referencialidade tende a ser condicionante no uso de uma ou outra forma. Sobre isso, os resultados apontaram que em caso de

referente [+humanos] e [+gên sem.] os sujeitos tendem a ser explicitamente expressos e que, por outro lado, quando em traços opostos [-humanos] e [-gên sem.] o sujeito tende a ser nulo.

Em síntese, a análise verificou que em contextos de sentenças negativas simples, há preferência pelo sujeito nulo e isso justifica-se pelo fato de que em contextos nos quais o advérbio de negação encontra-se à esquerda, privilegia-se o não preenchimento do sujeito pronominal, logo acarreta diretamente a limitação de uso do sujeito pronominal.

Por fim, com base nos resultados, Silva assume que as sentenças negativas simples favorecem a manutenção do sujeito nulo e que a mudança no que tange à realização dos sujeitos plenos afetou todas as pessoas do discurso, mas de forma distinta, tendo sido afetadas primeiramente a 1ª e a 2ª pessoa e, só em seguida, a terceira.

2.4.4. Nunes (2000)

O trabalho desenvolvido por Nunes (2000) discute o preenchimento ou não do sujeito através de informantes da cidade de João Pessoa, na Paraíba. À luz da perspectiva variacionista laboviana, a autora aponta que as variáveis pessoa do discurso, tempo verbal, tipo sintático da oração e faixa etária agem como condicionadoras do preenchimento ou não do sujeito pronominal.

Em linhas gerais, a autora apresenta a maior predominância do sujeito preenchido através de formas pronominais em detrimento do não preenchimento. Através da análise dos dados, detectou-se 70% de preenchimento e 30% de não preenchimento de sujeito pronominal da fala capital paraibana.

De acordo com os resultados encontrados no trabalho, na fala pessoense, foi observado que o antecedente pronominal e o nível de escolarização, também arrolados como variáveis na pesquisa, não demonstraram significância no que tange ao preenchimento ou não do sujeito pronominal. Contudo, as demais variáveis selecionadas favorecem o seu preenchimento.

Em especial, foi observado que ao longo do período, o falante escolhe usar ou não o sujeito pronominal e que, após a escolha da pessoa verbal, há maior predominância do preenchimento do sujeito quando ele é referente à segunda pessoa do singular, através da forma “você”. Já a forma “tu”, canônica como referente da segunda pessoa do singular não foi verificada como relevante para a análise.

Também foi constatado que tanto no singular como no plural, as formas de primeira pessoa “eu” e “nós” também favorecem esse preenchimento. Dessa forma, a autora pontua que esse processo revela um estágio de mudança avançado. A terceira pessoa do discurso, na pesquisa, foi encarada como inibidora do fenômeno.

O tempo verbal atuou como a variável que mais favoreceu o preenchimento ou não do sujeito pronominal. Assim, a autora concluiu que os verbos no pretérito e no presente condicionam o uso do pronome explícito e que todos os tempos do subjuntivo, bem como a forma nominal do infinitivo favorecem a utilização do sujeito nulo.

Ao se debruçar sobre a relação do tipo de oração, observou-se que em orações subordinadas, absolutas e principais, há o predomínio do sujeito nulo e que os demais tipos de sentenças atuam no condicionamento do preenchimento do sujeito pronominal.

Em relação à faixa etária, a autora afirma que há possibilidade de verificar um processo de mudança em relação ao que os trabalhos da região Sudeste apontam, pois os falantes com mais de 50 anos favorecem o sujeito pleno e os falantes classificados como jovens (15 a 25 anos) e como adultos (26 a 49 anos), favorecem o sujeito nulo.

2.4.5.Santos (2014)

Buscando observar a discussão acerca do preenchimento ou não do sujeito pronominal em Alagoas, apresentamos a pesquisa de Santos (2014) que objetivou verificar o preenchimento do sujeito na fala de crianças e adolescentes de Maceió, capital alagoana, especificamente, das que vivem em

entidades filantrópicas, através de entrevistas e narrativas, na tentativa de verificar a tendência ao não preenchimento dessa posição. Para tal, a autora selecionou um conjunto de fatores para examinar de que forma eles interagem no processo de preenchimento ou não do sujeito, a saber, concordância verbal, tempo verbal e natureza do sujeito, escolaridade, faixa etária e sexo.

Ao se debruçar sobre os dados coletados, observou que na fala das crianças e adolescentes das entidades analisadas, existe de forma visível um processo de variação entre o sujeito preenchido e não preenchido, tendo resultados de 59% e 41% respectivamente.

De acordo com o programa Goldvarb X, as variáveis elencadas como estatisticamente significativas nesse processo de variação foram o tempo verbal, a natureza do sujeito, a concordância verbal e o nível de escolaridade.

Com relação ao tempo verbal, a autora destaca que ele foi a primeira selecionada no Goldvarb como relevante para a variação analisada. Os dados revelaram que o tempo presente favorece o uso do sujeito preenchido e que, por outro lado, o tempo passado favorece o uso do sujeito não preenchido. No tempo presente, recebe destaque o uso maior das formas pronominais de primeira e terceira pessoa do singular e de uma maior adequação à regra padrão de concordância verbal. No âmbito do tempo passado, registrou-se o maior uso de marcas morfológicas de terceira pessoa do plural e do pronome inovador “a gente”.

Ao avaliar a variável natureza do sujeito, os fatores foram: SPron de 1ª pessoa do singular e do plural, SPron de 3ª pessoa do singular e do plural, SPron com “a gente” e SNom. Os resultados apontaram que o preenchimento do sujeito é favorecido quando é constituído por um sintagma nominal, tanto no singular quanto no plural. Todos os outros fatores favoreceram o uso do sujeito nulo, tendo destaque a 1ª pessoa como a que mais condiciona o uso do não preenchimento. Ainda nessa variável, a autora esclarece que os pronomes de segunda pessoa “você” e “vocês” representaram um número mínimo de ocorrências e por isso foram descartados da análise, mas destaca que pelo que foi observado a maioria relacionava-se com o preenchimento do sujeito.

No que tange à variável concordância verbal, os resultados apresentaram que quando a regra de concordância padrão não é obedecida, a tendência é o preenchimento do sujeito. Por outro lado, o fator + concordância foi diretamente motivado pelo grande número de ocorrências com a primeira pessoa do singular, o que favorece, como já mencionado, a preferência pelo não preenchimento do sujeito.

O nível de escolaridade também foi considerado relevante nesse processo de variação. O que os dados mostraram é que ao fim do Ensino Fundamental, os falantes tendem a preencherem mais o sujeito do que os que estão no início desse ciclo formativo. Ainda foi detectado que os estudantes que estão concluindo o EF fazem uso de diversos sujeitos pronominais, já os do início, optam pela primeira e terceira do singular e que estes também são os que menos utilizam a concordância padrão e os elementos verbais no tempo presente.

Conforme visto no conjunto de estudos apresentados nesta seção, o sujeito nulo é analisado através de diversos olhares e, conseqüente, apresenta resultados distintos, principalmente, quando se observa sua realização a partir de dados de fala ou de escrita, além do resultado advindo da correlação com os mais diversos fatores linguísticos e/ou sociais. Assim, a partir desse panorama de discussão acerca do preenchimento *versus* não preenchimento do sujeito pronominal, objetivamos relacionar tais achados com os encontrados, especificamente, acerca da segunda pessoa do singular no âmbito da comunicação digital. Para tal, na próxima seção, apresentaremos essa modalidade de uso da língua, suas características específicas, a importância da realização de pesquisas nesse ambiente e alguns estudos que se ocuparam de fenômenos linguísticos realizados no contexto da comunicação digital.

3. A COMUNICAÇÃO DIGITAL: QUE MODALIDADE É ESSA?

O surgimento dos primeiros computadores aconteceu quase que ao mesmo tempo nos Estados Unidos, na Alemanha e na Inglaterra, e é datado da primeira metade da década de 1940. No primeiro momento, o uso ficou reservado aos militares para serem usados em cálculos científicos, passando a se disseminar apenas na década de 60. Segundo Lévy (2000, p.31) os computadores eram grandes, frágeis e isolados em salas refrigeradas, utilizados como grandes máquinas de calcular. Ele afirma que neste início a informática servia apenas para cálculos científicos, estatísticas de Estados e empresas e outros trabalhos pesados de gerenciamento.

É notável que o desenvolvimento da tecnologia, em especial, da Internet, encurtou as distâncias entre as pessoas, pois a comunicação passou a existir de forma mais facilitada através de outra modalidade de uso da língua. As redes sociais têm grande contribuição na constituição desse novo paradigma, pois as interações nelas realizadas são capazes de garantir propósitos comunicativos diversos.

É cada vez maior o número de pessoas que tem acesso e utiliza cotidianamente as diversas ferramentas digitais, como smartphones, computadores, tablets e demais dispositivos, para fins de comunicação. As facilidades do acesso, que vão desde o grande e quase ilimitado alcance até a rapidez da comunicação, são tentadoras e talvez até indispensáveis nos dias de hoje, fazendo com que a adesão a essa nova forma de comunicação esteja se alastrando rapidamente e ganhando cada vez mais usuários.

A partir do momento em que a internet surge, naturalmente, o conceito de meios de comunicação é revisto. No Brasil, assume-se que é a partir da década de 1990 que a internet consolida-se e passa a integrar o dia-a-dia da população. Segundo o Comitê Gestor de Internet no Brasil (2015), metade das residências brasileiras possuem acesso à internet, um número gigante e crescente a cada ano.

Com base nessas constatações, percebemos que a comunicação digital surge como uma nova possibilidade de interação entre os seres humanos. E sobre ela, outras várias nomenclaturas vão sendo referenciadas com o objetivo de definir e legitimar essa nova modalidade, como 'discurso mediado por computador' (Herring, 2001) ou 'discurso escrito interativo' (Androutsopoulos, 2011). Para fins didáticos, nesta tese, recorreremos ao termo popularmente utilizado e assumiremos 'comunicação digital'.

Lévy (2020, p. 81) argumenta que a comunicação digital é ainda mais interativa do que a comunicação telefônica, por exemplo, pois apesar de ambas serem possíveis graças ao processo de avanço tecnológico, a primeira implica, na mensagem, na imagem da pessoa e na da situação, aspectos que inevitavelmente estão em jogo no processo de comunicação.

Compreendendo que as demandas sociais também são motores propulsores no engajamento de mais e mais pessoas no mundo virtual, e considerando a estreita ligação entre a sociedade e linguagem, faz-se necessário enxergar que a interação verbal ganhou um novo *lócus* na história recente da humanidade. Os fenômenos linguísticos se realizam no contexto variável dos acontecimentos sociais e, sendo prática humana, revelam o uso particular que grupos ou classes de homens fazem. (Cavalcanti e Catanduba, 2004).

Obviamente, é inegável que a comunicação digital só pôde ser desenvolvida a partir do surgimento das redes de computadores. Sobre isso, Lévy (1999, p. 31 e 32) comenta que:

No final dos anos 80 e início dos anos 90, um novo movimento sociocultural originado pelos jovens profissionais das grandes metrópoles e dos campi americanos tomou rapidamente uma dimensão mundial. Sem que nenhuma instância dirigisse esse processo, as diferentes redes de computadores que se formaram desde o final dos anos 70 se juntaram umas às outras enquanto o número de pessoas e de computadores conectados a inter-rede começou a crescer de forma exponencial. [...] as tecnologias digitais surgiram, então, como a infra-estrutura do ciberespaço, novo espaço de comunicação, de sociabilidade, de organização e de transação, mas

também novo mercado da informação e do conhecimento.

Sendo reconhecido como um novo espaço de comunicação, a rede mundial de computadores acabou por gerar novas formas de uso da língua, trazendo consigo um compartilhamento de características já conhecidas nas modalidades de escrita convencional e da oralidade, mas com diversas inovações inerentes à modalidade que então surgiu.

De acordo com Vieira (2016):

[...] Do ponto de vista terminológico, quando se menciona a expressão “comunicação digital” significa falar simultaneamente em um conjunto de quatro elementos fundamentais: (1) a Internet, que diz respeito à rede de computadores conectados e capazes de se comunicar; (2) a World Wide Web, que não é a Internet em si, mas uma parte desta, ou seja, trata da dimensão lógica mais visível, especificamente o conjunto de páginas com interface gráfica que funciona neste maquinário digital; (3) os dispositivos móveis e outros gadgets baseados em bits (como câmeras digitais, scanners, filmadoras etc.); (4) a interação entre os diversos atores através dessa estrutura. (Vieira, 2016, p. 12 e 13)

Ainda de acordo com Vieira (2016), a comunicação digital foi popularizada graças à Internet e ao recente processo de massificação de smartphones, conseqüentemente, tem impactado a vida de cidadãos, grupos, organizações e governos, trazendo novos aspectos a serem consideradas nas práticas sociais.

A partir do final da década de 1990, o uso da internet ainda era tímido, contudo, os primeiros estudos da Linguística surgiam apontando para fenômenos linguísticos observados nesse novo ambiente de comunicação. De acordo com Tagliamonte (2015, p 02), “o termo comunicação mediada por computador apareceu pela primeira vez em Kiesler *et al* (1984:1123), cujo objetivo era analisar a ascensão e disseminação da internet e comunicação

baseada em rede⁵". Mais adiante, Herring (1996) tem seu destaque com a publicação de *Computer-mediated Communication e Linguistic, Social and Cross-cultural Perspectives*. Esses primeiros estudos revelam os primeiros olhares quanto à linguagem no âmbito da comunicação digital. Posteriormente, outras análises seguiram buscando dar conta, por exemplo, de elementos condicionadores de fenômenos estruturais das línguas no contexto da comunicação digital.

A partir disso, a inquietação do campo da Linguística busca dar conta da caracterização da comunicação digital diante da escrita convencional e da oralidade. De acordo com Crystal (2005):

a comunicação mediada por computador é mais do que um agregado de características faladas e escritas. Porque faz coisas que nenhum desses outros meios faz, tem de ser vista como uma nova espécie de comunicação. É mais do que um híbrido de fala e escrita, ou o resultado do contato entre dois veículos existentes há muito. **Os textos eletrônicos, de qualquer tipo, não são a mesma coisa que as outras formas de texto.** Eles demonstram fluidez, simultaneidade (ao estarem disponíveis em um número indefinido de máquinas) e não se degradam com cópias; transcendem as limitações tradicionais de disseminação do texto; e possuem fronteiras permeáveis (por causa do modo como um texto pode ser integrado a outros ou exibir links para outros). (Crystal, 2005, p. 90, *grifo nosso*)

Questões diversas são propostas numa primeira instância na tentativa de compreender a relação das propriedades da modalidade digital com as das demais modalidades de uso da língua ou se esse novo ambiente de uso da língua carrega consigo características específicas, não encontradas nem na fala nem na escrita fora do ambiente de interação via Internet. Squires (2010) afirma que:

⁵ The term computer-mediated communication first appeared in Kiesler et al. (1984:1123), whose goal was to analyze the social psychological implications of the rise and spread of the internet and network-based communication.

os novos estudiosos da comunicação mediada por computadores ficaram intrigados com o fato de que o uso da CMC parece abranger características da fala e da escrita: por ser informal, síncrona e efêmera- “como a fala”- e/ou editável, baseado em texto e assíncrono- “como a escrita”⁶. (SQUIRES, 2010, p. 462 ; tradução nossa)

De acordo com Androutsopoulos (2011, p 149), “a escrita no ambiente digital relaciona-se diretamente com três temas: oralidade, compensação e economia”. No âmbito da oralidade, verifica-se que a comunicação digital manifesta no texto escrito aspectos oriundos da língua oral. Já no que tange ao caráter da compensação, discute-se que os aspectos entoacionais e a falta de expressões faciais fazem com que outros recursos sejam acionados para facilitar a interação, como fontes, emojis e demais recursos computacionais. O aspecto da economia linguística é evidenciado através da utilização de diversos recursos para fazer com que a comunicação aconteça de forma eficiente, a partir do acionamento do menor número de recursos linguísticos que precisam ser acomodados no número de caracteres disponíveis em cada tipo de ambiente digital e adequados à agilidade da comunicação digital.

Androutsopoulos (2011, tradução nossa) destaca que “Toda a escrita em rede é realizada em tecnologias digitais que permitem a comunicação privada ou pública, troca assíncrona ou quase síncrona entre indivíduos e grupos em vários aplicativos ou plataformas”⁷. A autora (*op cit*) ainda argumenta que:

“o que há de novo na “linguagem da nova mídia” não é apenas um número de construções ou estruturas inovadoras, mas novos recursos e estratégias para produção linguística escrita e construção de significado,

⁶ Early CMC scholars were intrigued by the fact that CMC usage seems to encompass canonical features of both speaking and writing: it may be informal, synchronous, and ephemeral—“like speech”—and/or editable, text-based, and asynchronous—“like writing.”

⁷ All networked writing is carried out on digital technologies that enable private or public, asynchronous or near-synchronous exchange among individuals and groups on various applications or platforms.

da grafologia à estrutura do discurso. (Androutopoulos, 2011, p.153, *tradução nossa*)⁸.

Como pode ser visto, o campo de investigação linguística nas mídias digitais apesar de recente, já aponta para certo alinhamento no que tange ao seu horizonte e definições. Apresentando-se como mais um ambiente de manifestação das estruturas linguísticas nos mais diversos níveis e trazendo diversos pontos a ser analisados.

Cavalcanti & Catanduba (2004), especificamente sobre a comunicação nas redes sociais, apontam que:

Como quando se escreve nas redes sociais há um misto de fala e de escrita, ou seja, escreve-se representando uma possível fala, que é, na maioria das vezes, coloquial, este exemplo **não configura um uso de internetês, mas sim, o uso da língua numa variante não padrão**. O que caracteriza o internetês é o seu léxico e a forma como as palavras são registradas, obedecendo, inclusive, a uma regularidade nas variações propostas. Fugir a regras básicas e convenções cristalizadas é o mesmo que desobedecer às normas da língua padrão na modalidade escrita formal. (2004, p.6, *grifo nosso*)

Com o suporte dessas definições, temos um norte a percorrer durante a investigação aqui empreendida, logo, a partir dessa caracterização apresentada, bem como da tentativa de responder às inquietações reveladas na introdução desta tese, a subseção a seguir apresentará um pequeno conjunto de estudos acerca dos processos de variação linguística no contexto da comunicação digital, a fim de situar a produção realizada até o momento e localizar nosso estudo dentro desse panorama.

3.1. Os fenômenos linguísticos na comunicação digital

A escassez de trabalhos que tratem especificamente de processos linguísticos de ordem estrutural na comunicação digital nos inquieta e motiva a

⁸ What is new in 'new media language' is not just a number of innovative constructions or structures, but new resources and strategies for written language production and meaning making, from graphology to discourse structure.

investigar o que acontece internamente durante as interações linguísticas via Internet.

Pinheiro (2023), ao fazer uma revisão de literatura científica sobre as estruturas linguísticas na comunicação digital, com foco nos fenômenos de variação linguística, aponta para a mesma dificuldade em termos de produção acadêmica acerca da temática.

A premissa de que onde há interação humana, há uso linguístico, independentemente de onde ela ocorra, nos leva a compreender esse universo da comunicação digital como um vasto território de estudos e possibilidades investigativas. Nesse sentido, observar o uso da língua na comunicação digital também é revelador das características inerentes a esse sistema vivo e heterogêneo.

Para Squires (2016, p.2):

Assumimos que a língua na comunicação mediada por computadores (CMC) varia; a língua na CMC é representada e representa; a língua na CMC muda e é mudada. Essas premissas nos permitem as perguntas muito mais interessantes de como, abordadas por um leque de perspectivas analíticas, incluindo a variação linguística quantitativa, a mudança diacrônica, o contato linguístico, a ideologia linguística, a identidade sociolinguística, as redes sociais e estilo.

Acompanhando os trabalhos que se ocupam dessa temática, as questões sobre o tipo, características e inovações desse novo uso da língua sempre são levantadas em maior ou menor intensidade. De acordo com Tagliamonte (2015, p. 01, tradução nossa), a grande questão seria “que tipo de linguagem as pessoas usam quando se comunicam com os outros usando mediação baseada em dispositivo? [...]”⁹.

A tentativa de responder a esse tipo de questionamento pode ser feita a partir de diversas perspectivas. Os estudos buscam averiguar os processos

⁹ What type of language do people use when they communicate with each other using device-based mediation, a phenomena referred to as CMC?

linguísticos, inclusive os de variação, no meio digital a partir de chats, de conversas em aplicativos de mensagens, de postagens em redes sociais, de usos linguísticos específicos de pessoas ou personagens em plataformas digitais, etc. Assim, buscando apresentar um breve quadro de pesquisas que exemplifiquem o foco em processos de variação, a seguir, traremos os trabalhos de Othero (2018), Silva e Pinheiro (2020) e Pinheiro (2021).

3.1.1. Othero *et al* (2018)

Objetivando analisar o processo de retomada anafórica do objeto direto no português brasileiro, os autores investigaram um *corpus* constituído de histórias em quadrinhos infantis e um outro *corpus* composto por postagens na rede social Twitter. Ambos os *corpora* trazem marcas de oralidade, buscando uma aproximação com a fala e objetivam comparar duas hipóteses sobre a distribuição entre objetos nulos e pronomes e, ainda, determinar a distribuição entre objetos nulos, clíticos pronominais e pronomes plenos de 3ª pessoa na retomada anafórica de objetos diretos, com o intuito de entender se o processo distributivo desses elementos se assemelha mais aos encontrados nos estudos da fala ou da escrita.

A escolha pelos quadrinhos deu-se em virtude dos diálogos que podem ser facilmente observados nos balões que indicam as falas dos personagens, já as postagens realizadas no Twitter foram selecionadas por apresentarem, de um modo geral, uma escrita com marcas de oralidade em virtude do baixo grau de monitoramento dos internautas.

Ao comparar os dois *corpora*, os autores confirmam que ambos possuem marcas de oralidade e que buscam, apesar de serem língua escrita, representar a fala. Especificamente, na retomada anafórica de objeto direto de 3ª pessoa, foram as histórias em quadrinhos que apresentaram o padrão mais próximo da língua escrita, tendo um percentual de ocorrências de clíticos pronominais de 3ª pessoa e de objeto nulo próximos aos encontrados nas pesquisas da fala no PB. A análise trouxe, ainda, um grande número de ocorrências de retomadas anafóricas com pronomes plenos, que ainda são estigmatizados na escrita com monitoramento. Tais achados levam ao entendimento de que, possivelmente, o

uso de pronomes plenos pode estar sendo uma possibilidade de retomada anafórica pronominal com vistas a uma não estigmatização no português brasileiro.

Nos dados revelados no *corpus* da rede social Twitter, os autores concluem que se encontram características mais próximas da fala vernacular do português do Brasil. As ocorrências de retomadas anafóricas não revelaram o uso de clítico pronominal nessa condição, mas sim, a predominância do uso do objeto nulo e de pronomes plenos, conforme também é visto na fala do PB.

Em suma, o estudo empreendido apontou que a retomada anafórica do objeto direto de 3ª pessoa é realizada a partir de cinco tipos de estratégias nos corpora averiguados: 1) a retomada anafórica em *corpus* escrito com característica de fala é frequentemente realizada pelo objeto nulo; 2) A utilização de pronome pleno é outra estratégia recorrente no *corpus*, e tal achado ainda não tinha sido apresentado em estudos de corpora escrito; 3) O clítico é uma estratégia apenas na história em quadrinhos que é afetada pelo padrão da escrita, mas na plataforma Twitter encontra-se em desuso, como na fala em PB; 4) Os dados tratados estatisticamente demonstraram que tanto a animacidade quanto o gênero semântico favorecem a retomada por pronome pleno ou objeto nulo; 5) O fator animacidade mostrou-se influente na retomada anafórica pronominal em ambos os corpora, equiparando-se aos achados de outros estudos com dados de fala no PB.

3.1.2. Silva e Pinheiro (2020)

Com o propósito de estudar a variação da 1ª pessoa do singular e o objeto direto de 3ª pessoa em referência anafórica, na web, na fala e na escrita convencional, os autores investigam conversas privadas de Whatsapp, aplicativo de mensagens instantâneas, e Blogs de viagens, que são diários virtuais com linguagem informal nos quais são compartilhadas as experiências em viagens através da web.

Para tal, o trabalho busca contrastar os resultados encontrados na análise sobre a variação da 1ª pessoa do singular através dos blogs, de Lima (2014),

com os resultados de Paredes Silva (1988) que analisou o mesmo fenômeno através de cartas pessoais do século XX.

Ao comparar os dados, Silva e Pinheiro (2020) afirmam que a ausência de sujeito de 1ª pessoa nas cartas analisadas por Paredes Silva também foi encontrada nos blogs de viagens estudados por Lima (2014). Nesse caso, nas duas pesquisas em questão, a conexão discursiva foi fator condicionante para o não preenchimento do sujeito. Assim, a conexão de grau 1, sendo considerada como ótima, propiciou o maior de número de ocorrências da ausência de sujeito de primeira pessoa, em virtude de que quanto mais a informação está conectada à sua menção prévia, menos foi necessário explicitar a informação mais adiante.

Com isso, chegaram ao resultado de cerca de 80% dos dados com o pronome zero em ambos os estudos, sendo especificamente 77% nas cartas pessoais e 78% nos blogs.

No âmbito da investigação sobre a variação do objeto direto de 3ª pessoa, Silva e Pinheiro (2020) propuseram-se a contrastar os resultados encontrados por Pinheiro (2016 e 2017) na perspectiva da Sociolinguística laboviana.

No trabalho de 2016, Pinheiro demonstrou que a anáfora zero é a variante mais proeminente entre um grupo de informantes formado por jovens cariocas em ambiente socioeducativo. No trabalho, o autor destaca que 50% das ocorrências são correspondentes ao uso da anáfora zero em detrimento das demais possibilidades de preenchimento do objeto direto de 3ª pessoa. Entre os fatores condicionadores desse uso, estão: 1) a distância de apenas uma oração entre as menções; 2) a manutenção da função sintática do referente; e 3) o traço [- animado] do referente.

Já em Pinheiro (2017), o objetivo foi investigar o objeto direto de 3ª pessoa em referência anafórica em conversas privadas no aplicativo Whatsapp entre dois jovens interlocutores brasileiros e universitários. Esse trabalho analisou a variação entre a anáfora zero, nome e pronome reto e aponto a preferência pelo uso da primeira variante em 64,2% dos dados. Os fatores que influenciaram esse número foram: o traço [- animado] do referente; a distância

de apenas uma oração entre as menções e a manutenção da função sintática do referente, respectivamente.

O estudo de Silva e Pinheiro (2020) ainda compararam esses dados com os achados em Averbug (2020) que analisou o mesmo fenômeno em variação nos dados de redações de universitários e encontrou a predominância de uso de clítico, 40%, e apenas 23% de anáfora zero, diferentemente das pesquisas anteriores.

Comparando os resultados encontrados, os autores concluem que os dados do Whatsapp e os de fala, apesar das distinções entre as classes sociais dos informantes, compartilham mais semelhanças entre si do que os de escrita convencional. Logo, mesmo o Whatsapp sendo uma ferramenta que funciona como um chat de mensagens escritas, a realização do fenômeno analisada tende a seguir a seguir o funcionamento observado na língua oral.

3.1.3. Pinheiro (2021)

O trabalho de Pinheiro (2021), resultado de sua dissertação de mestrado, propõe uma comparação entre o processo de variação do objeto direto de 3ª pessoa em chats privados e em cartas pessoais, a partir do suporte teórico da Sociolinguística Variacionista (Labov, 2008 [1972]), tendo em vista que esses dois gêneros discursivos carregam características da escrita informal e da conversação de forma espontânea. O autor chama a atenção para o fato de que apesar das semelhanças entre os gêneros, algumas características são bem distintas, como os suportes que os veiculam e o (as)sincronismo de sua manifestação.

Para tal, o autor constitui seu *corpus* a partir de mensagens em chats privados na plataforma Messenger, no ano de 2012, e em cartas pessoais escritas entre 1979 e 1984, e comparou esses dados com amostras de fala do português nos mesmos períodos. Os fatores linguísticos selecionados foram a animacidade do referente, a dia distância referencial e a função sintática do antecedente.

Através da análise feita a partir dos resultados estatísticos obtidos através do Goldvarb (2001), concluiu que os chats privados compartilham características das produções de fala informais no português brasileiro. Os dados revelaram que há a predominância da anáfora zero de forma majoritária, seguida do sintagma nominal anafórico, do pronome pleno e do clítico acusativo, achados semelhantes aos já encontrados em estudos da modalidade falada do português no Brasil.

Os grupos de fatores mostraram-se relevantes estatisticamente no que tange à influência na realização da anáfora zero, sendo o traço [– animado] nos referentes, distâncias referenciais menores e a função sintática de objeto direto no antecedente os que favorecem essa realização, conforme também já verificado em amostras de diversos trabalhos da modalidade oral do PB.

Os dados das cartas pessoais revelaram divergência quando comparados aos dados de fala, tendo grande uso de clítico acusativo que é normalmente associado aos contextos de escrita formal. Curiosamente, a anáfora zero foi a variante com maior percentual de uso, traço apontado na escrita mais informal desde o fim do século passado.

Somente o traço [– animado] nos referentes, elencado como um dos fatores observados, demonstrou favorecer a anáfora zero, confirmando o que outras pesquisas já apontaram. A distância referencial mostrou-se relevante apenas quando a anáfora zero e o clítico acusativo foram considerados de forma conjunta, a partir de menores distâncias entre as menções a um mesmo elemento referente.

Por fim, o autor conclui que os dados referentes às cartas pessoais analisadas distanciam-se do que é encontrado no PB falado no que tange à variação do objeto direto de 3ª pessoa. Logo, na comparação entre as mensagens dos chats privados e as cartas pessoais, as primeiras se aproximam mais das produções orais no português do Brasil.

Após apresentar a caracterização e conceitos da comunicação digital, bem como os mais recentes trabalhos que analisam fenômenos morfossintáticos

nessa modalidade de uso da língua, percebemos quão rica é essa “nova” realidade de manifestação linguística. A língua, mesmo em um contexto “novo” continua manifestando sua heterogeneidade e colocando em relevo processos de variação tanto quanto podem ser observados na fala e/ou na escrita convencional. Com isso, temos que o universo dos usos linguísticos materializados via dispositivos de comunicação digital, através de redes sociais e demais ambientes, deve ser explorado no âmbito da investigação linguística, pois suas características e possibilidades ainda precisam ser visualizadas como mais um canal real de manifestação da língua, principalmente no âmbito da investigação acerca da segunda pessoa do singular.

A seguir, traremos o aporte teórico-metodológico no qual nossa tese está ancorada.

4. A SOCIOLINGUÍSTICA VARIACIONISTA DE LABOV

A presente seção apresenta os pressupostos da teoria laboviana que ancoram nossa tese, bem como demonstra o modelo metodológico variacionista que foi utilizado para constituição do *corpus* da pesquisa, a definição das variáveis independentes, os procedimentos de coleta e análise dos dados.

4.1. O ESCOPO TEÓRICO DA SOCIOLINGUÍSTICA LABOVIANA

A Sociolinguística surge enquanto uma teoria que postula a heterogeneidade linguística de forma sistematizada, reivindicando destaque para o estudo da fala e trazendo à tona as variações que existem na língua através da inter-relação com fatores de ordem extralinguística. Nessa perspectiva teórica, a língua é dotada de “heterogeneidade sistemática”, fato que permite a identificação e demarcação de diferenças sociais na comunidade, constituindo-se como parte da competência linguística dos indivíduos, o domínio de estruturas heterogêneas (Weinreich; Labov; Herzog, 2006, p.10).

Dessa forma, refuta a noção de homogeneidade linguística e dá visibilidade ao que é inerente a qualquer língua humana: a variabilidade.

Indo contra a ideia de homogeneidade linguística difundida e defendida por Saussure (2006) e do conceito de falante ideal cunhado por Chomsky (1965, 1997), Labov postula que a língua não é propriedade do indivíduo, mas sim da comunidade, fato que o leva a crer que o novo modo de fazer linguística é “estudar empiricamente as comunidades de fala” (Labov, 2008, p.259).

Os estudos baseados na perspectiva sociolinguística surgiram na década de 1960, e tem seu marco a partir da conferência *Sociolinguistics Conference*, organizada por William Bright, na Universidade da Califórnia em Los Angeles, que contou com a presença de William Labov e objetivava discutir a relação entre língua e sociedade. Para Bright, era necessário relacionar as variações linguísticas observáveis em uma comunidade às variações existentes na estrutura social desta mesma sociedade.

A iniciativa de pesquisa de William Bright contribuiu para os primeiros estudos da Sociolinguística. Pela primeira vez, de fato, os fatores sociais, como

a identidade do falante no contexto da comunicação, passaram a ser considerados para o cenário da pesquisa linguística.

Apesar de a sociolinguística ser considerada como a primeira teoria a incorporar o lado social da língua aos seus estudos, Labov (1972) aponta que Saussure, precursor do estruturalismo considerado como iniciador da linguística, já apontava para a face social da linguagem, porém deixava-o de lado:

Segundo Saussure, langue “est la partie sociale du langage...elle n’existe qu’en vertu d’une sorte de contrat passé entre les membres de la communauté”[“é a parte social da linguagem...ela não existe fora de um tipo de contrato estabelecido entre os membros da comunidade”](1962:321). Por esta razão, a Escola de Genebra saussuriana é frequentemente mencionada como a escola social da linguística (Labov, 2008, p. 217)

Embora o lado social da linguagem já tivesse sido identificado como importante, os estudos estruturalistas não o consideraram, pois a preocupação teórica estava em analisar o que era a forma e não a substância da língua. Para tal, a língua deveria ser estudada a partir de suas relações internas, enquanto sistema que é formado por um conjunto de unidades que obedecem a certos princípios de funcionamento no qual o valor de cada elemento passa a existir a partir de suas relações internas ao próprio sistema.

Segundo Lucchesi (2004a, p. 157):

[...] o modelo teórico estruturalista não podia incorporar como objetos de sua reflexão sistemática nem a prática lingüística concreta, nem o processo sócio-histórico de constituição da língua. Baniam-se, desse modo, todas as questões relativas ao caráter social e histórico da língua dos principais cenários teóricos da ciência da linguagem.

Foi Bright quem chamou a atenção para a importância da diversidade linguística nos estudos da língua, para a real necessidade de estudar a língua a partir de sua inter-relação com os fatores de ordem social, no entanto, não chegou a definir uma metodologia capaz de incluir essa diversidade em seus trabalhos. Segundo Monteiro (2000), Bright definiu os fatores condicionantes do

fenômeno da diversidade linguística como aqueles ligados ao falante, mas ainda atrelava à sociolinguística um papel secundário ou complementar à Linguística, Sociologia e Antropologia.

Nesse sentido, tendo em vista a importância da definição de modelo metodológico que pudesse ser utilizado na análise de forma a abranger a diversidade linguística, dando o caráter teórico científico independente, o título de pai da Sociolinguística foi atribuído a William Labov, que definiu e orientou os procedimentos para o desenvolvimento dos estudos na área.

Labov (2008[1972]) desenvolveu estudos sobre a centralização dos ditongos (ay) e (aw) na ilha de Martha's Vineyard (1963) e sobre a estratificação social do < r > em lojas de departamento em Nova York (1966). A partir dos resultados destes estudos, o autor propôs um modelo teórico que considerava o lado social da linguagem, nestes estudos ele fazia a relação do social com o linguístico, assim observando e mais tarde expondo que os fatores sociais condicionavam as mudanças fonológicas na língua. Ele constatou que o uso dos ditongos au e ay servia para os falantes se identificarem como nativos, contrapondo com as formas linguísticas padronizadas utilizadas pelos turistas que visitavam a ilha, o que mostrou que o uso da variante pelos nativos, considerada estigmatizada em relação à forma padrão, servia para a construção de sua identidade social, como descendentes dos Yankees, o grupo étnico que colonizou a ilha no século XVII.

O modelo laboviano propôs o estudo da língua através de um método quantitativo, realizado através de procedimentos específicos para coleta de dados e constituição do *corpus* e consequente análise. Graças a essa definição, a Sociolinguística ganha o devido *status* científico em 1964.

Labov (2008) argumentando sobre o termo sociolinguística, explicou que:

Por vários anos, resisti ao termo sociolinguística, já que ele implica que pode haver uma teoria ou prática linguística bem-sucedida que não é social [...]. Apesar de um considerável volume de atividade sociolinguística, uma linguística socialmente realista parecia uma perspectiva remota nos anos de 1960. (Labov, 2008, p. 13)

Dessa forma, temos que é a partir de Labov, que o objeto de estudo da sociolinguística é a língua falada, observada, descrita e analisada em situações

concretas de uso. Tal definição foi decisiva para o novo olhar sobre os estudos linguísticos, trazendo o que é diverso e mutável da língua como cerne da questão.

Para Weinreich, Labov e Herzog (2006 [1968], p. 13):

A base do conhecimento intersubjetivo na linguística tem de ser encontrada na fala – a língua tal como usada na vida diária por membros da ordem social, este veículo de comunicação com que as pessoas discutem com seus cônjuges, brincam com seus amigos e ludibriam seus inimigos.

A variação linguística é vista como inerente à língua. Labov (1972, p.3) afirma que a ideia de variação sistemática é motivada por pressões sociais que “continuamente operam sobre a língua, não devendo, pois, ser estudada fora do contexto social”.

Para Tarallo (1990, p.06), a teoria da variação linguística trata-se de um modelo teórico-metodológico que assume o “caos” linguístico como objeto de estudo”. Como esse modelo, por princípio, não admite a existência de uma ciência da linguagem que não seja social. Sendo assim, para Labov (2008, p. 215):

A língua é uma forma de comportamento social: declarações deste tipo podem ser encontradas em qualquer texto introdutório. Crianças mantidas em isolamento não usam a língua; ela é usada por seres humanos num contexto social comunicando suas necessidades, ideias e emoções. (LABOV, 2008, p. 215)

A teoria sociolinguística ficou conhecida também como variacionista, pois o modelo Laboviano postula regras categóricas nas quais há observação de uma(s) variante(s) em oposição à outra(s) testando seu uso probabilisticamente, desenvolvendo assim um modelo de regra variável, com o qual é possível verificar se o uso de uma variável vai depender de fatores extralinguísticos e linguísticos mostrando as relações que são estabelecidas entre eles. Para Labov (2008), a variação não ocorre ao acaso, ela é motivada por um conjunto complexo de parâmetros, por condicionamento ou variáveis que favorecem ou inibem o emprego de variantes.

Labov (1978) define que dois enunciados que se referem ao mesmo estado de coisas com o mesmo valor de verdade constituem-se como variantes de uma mesma variável. Complementando tal noção, Mollica (2015, p 10- 11) explica que variantes são as diversas formas alternativas que configuram um fenômeno variável, tecnicamente chamado de variável dependente. Ainda segundo a autora:

Uma variável é concebida como dependente no sentido que o emprego das variantes não é aleatório, mas influenciado por grupos de fatores (variáveis independentes) de natureza social ou estrutural. Assim, as variáveis independentes ou grupos de fatores podem ser de natureza interna ou externa à língua e podem exercer pressão sobre os usos, aumentando ou diminuindo sua frequência de ocorrência. (MOLLICA, 2011, p.11)

Desse modo, pode-se observar que o uso de uma determinada variável não vai ser aleatório, pelo contrário, ele vai ocorrer em situações regulares tanto linguisticamente quanto socialmente, assim ocorrendo uma regularidade em meio à heterogeneidade. E graças a essa heterogeneidade, que é sistemática, há a identificação e demarcação de diferenças sociais na comunidade, constituindo-se como parte da competência linguística dos indivíduos, o domínio de estruturas heterogêneas (Weinreich; Labov; Herzog, 2006, p.101).

Labov, ([1972] 2014, p. 238), ainda destaca que:

A existência de variação e de estruturas heterogêneas nas comunidades de fala investigadas está certamente bem fundamentada nos fatos. É a existência de qualquer outro tipo de comunidade de fala que deve ser posta em dúvida. (...) a heterogeneidade não é apenas comum, ela é o resultado natural de fatores linguísticos fundamentais

Para o estudo sociolinguístico, a investigação começa a partir da observação das variáveis na comunidade linguística. Dessa forma, tem-se que a sociolinguística estuda a língua em uso em uma determinada comunidade linguística. Para Mollica (2013, p. 9), a Sociolinguística Variacionista é uma

ciência “presente num espaço interdisciplinar, na fronteira entre língua e sociedade, focalizando precipuamente os empregos linguísticos concretos, em especial os de caráter heterogêneo”. Essa língua é heterogênea, ou seja, não é falada da mesma forma por todos os membros da comunidade. Cada comunidade de fala possui características linguísticas que a distingue das outras e constituída por indivíduos que não falam do mesmo modo, mas que se relacionam intra e extralinguisticamente.

Esse modelo de investigação objetiva alcançar as respostas para responder a questão central do processo de mudança linguística, a partir de dois importantes princípios teóricos:

- (i) o sistema linguístico que serve a uma comunidade heterogênea e plural deve ser também heterogêneo e plural para desempenhar plenamente as suas funções rompendo-se assim a tradicional identificação entre funcionalidade e homogeneidade;
- (ii) os processos de mudança que se verificam em uma comunidade de fala se atualizam na variação observada em cada momento nos padrões de comportamento linguístico observados nessa comunidade, sendo que, se a mudança implica necessariamente variação, a variação não implica necessariamente mudança em curso (Weinreich; Labov; Herzog 1968).

Diante disso, o trabalho da sociolinguística é identificar categorias independentes que condicionem o uso ou o não-uso de uma variante que podem ser internas ou externas a língua, os fatores internos são estruturais, próprios do sistema linguístico, e os externos são sociais. Ambos vão influenciar ou não para o uso de uma ou outra variante dependente. Nesse sentido, para Naro (2003, p. 16), a questão central para a Teoria da Variação é a avaliação do *quantum* com que cada categoria postulada contribui para a realização de uma ou de outra variante.

Dito de outra forma, a Sociolinguística parte do pressuposto de que “toda variação é motivada, isto é, controlada por fatores de maneira tal que a heterogeneidade se delinea sistemática e previsível”, Mollica (2015, p. 10).

Ancorado neste arcabouço teórico, a presente tese busca analisar a variação entre as variáveis preenchimento pronominal e sujeito nulo de segunda pessoa do singular na posição de sujeito na comunicação digital feita por empresas alagoanas, através de postagens em seus perfis oficiais na rede social

Instagram e, para tal, recorreremos à análise a partir da motivação sistemática de fatores linguísticos que são condicionantes ao uso de uma ou outra(s) variável(is) em questão, como veremos mais adiante.

4.2. A METODOLOGIA SOCIOLINGUÍSTICA

Na presente subseção, descrevemos o arcabouço metodológico da pesquisa sociolinguística e os procedimentos específicos para a realização da pesquisa aqui proposta. Além disso, apresentamos o processo de coleta de dados, codificação e tratamento através do software R. Trazemos, ainda, o *locus* da pesquisa, as informações detalhadas sobre as empresas que serviram como fonte para coleta de dados e as variáveis linguísticas selecionadas para o estudo.

O modelo teórico-metodológico da Sociolinguística permite que analisemos quais os fatores sociais e linguísticos proporcionam o uso ou o não-uso das variantes linguísticas. Em uma pesquisa científica de cunho sociolinguístico, os procedimentos metodológicos são de suma importância, uma vez que utilizados de maneira correta conferem qualidade e sucesso à pesquisa.

Os estudos científicos seguem métodos que se adequam à finalidade dos mesmos. As pesquisas variacionistas, em específico, seguem um método de quantificação de dados, proposto por Labov (2008), com o objetivo de analisar a probabilidade do uso de variantes linguísticas considerando fatores linguísticos e extralinguísticos como condicionantes ou não destas variantes, observando se elas são apenas variações ou mudanças linguísticas. Desse modo, a metodologia sociolinguística permite e tem como objetivo analisar e sistematizar as variantes linguísticas existentes em uma mesma comunidade de fala (Tarallo, 1991).

Para Tarallo (1991), é o objeto de estudo, ou seja, o fato linguístico que mostrará qual o método a seguir. Nesse sentido, o autor afirma que:

O fato sociolinguístico, o dado de análise, é ao mesmo tempo a base para o estudo linguístico: o acervo de informações para fins de confirmação ou rejeição de hipóteses antigas sobre a língua e também para o levantamento de novas hipóteses. (Tarallo, 1991 p.18)

Amparados nesses cuidados procedimentais que garantem a qualidade para a pesquisa sociolinguística, nas subseções seguintes, descreveremos os procedimentos metodológicos definidos especificamente para nossa investigação.

4.2.1. Os procedimentos metodológicos adotados para o estudo

Na presente subseção, apresentamos os procedimentos metodológicos realizados especificamente para a investigação proposta nesta tese. Assim, a seguir, explicitamos as variáveis selecionadas para o trabalho, o *lócus* da pesquisa, os critérios para a escolha das empresas que serviram como suporte para coleta dos dados, os critérios de classificação e constituição do *corpus* analisado e o software R que foi utilizado para tratamento estatístico dos dados analisados.

4.2.1.1. As variáveis linguísticas elencadas

Entendendo que Sociolinguística compreende o estudo da língua a partir de suas relações com fatores linguísticos e de ordem social, os estudos empreendidos a partir dessa perspectiva buscam observar de que modo as variáveis linguísticas e/ou extralinguísticas condicionam e explicam os fenômenos em variação. Para tal, uma série de variáveis são elencadas de modo a quantificar e analisar a interferência delas no que tange a um fenômeno linguístico.

De acordo com Naro (2012):

O problema central que se coloca para a Teoria da Variação é a avaliação do *quantum* com que cada categoria postulada contribui para a realização de uma ou de outra variante das formas em competição. No uso real da língua, que constitui o dado do linguista, seja na forma falada ou na forma escrita, tais categorias se apresentam sempre conjugadas; na prática, a operação de uma regra variável é sempre o efeito da atuação simultânea de vários fatores (Naro, 2012, p.16-17).

Ainda nesse sentido, para Mollica (2015, p 27), as variáveis, tanto linguísticas quanto não linguísticas, não agem isoladamente, mas operam num conjunto complexo de correlações que inibem ou favorecem o emprego de formas variantes semanticamente equivalentes. Dessa forma, tem-se que os fatores como nível de escolaridade, estratificação social, idade, região do falante, sexo, entre outros são potencialmente influenciadores no uso de uma forma linguística. Guy (2007, p.33-34) afirma que:

A análise da regra variável é um tipo de análise multivariada altamente empregada em estudos de variação linguística hoje em dia. Seu propósito é separar, quantificar e testar a significância dos efeitos de fatores contextuais em uma variável linguística. Esses fatores condicionantes podem ser tanto sociais (...), ou linguísticos.

Para análise aqui proposta, focamos nosso olhar, especificamente, nas variáveis linguísticas, tendo em vista que analisamos postagens empresariais no meio da comunicação digital, não sendo possível, portanto, a observação de variáveis sociais e dando um caráter mais verticalizado nas questões de ordem linguística envolvidas nessa modalidade de uso da língua. Assim, para observar o papel da influência das variáveis linguísticas em relação ao fenômeno de variação do preenchimento sujeito de segunda pessoa do singular e, considerando alguns resultados de estudos anteriores que destacam os potenciais motivadores de uso de uma ou outra variável dependente, elencamos: 1) tipo de sentença, sendo interrogativas e não interrogativas; 2) tempo verbal, 3) modo verbal e 4) paralelismo.

4.2.1.1.1. Variável tipo de sentença

Trabalhos anteriores que se debruçaram sobre a análise da segunda pessoa do singular no português já mencionam que o tipo de sentença pode condicionar a realização da segunda pessoa na posição de sujeito, tanto quanto

à escolha do preenchimento via elemento pronominal quanto na realização do sujeito nulo, como Mota (2008) e Silva (2021), por exemplo.

Nesse sentido, para nossa investigação, consideramos como variáveis, especificamente, as sentenças interrogativas, aquelas em que o emissor faz uma pergunta ao interlocutor, podendo ser direta ou indireta; e as não interrogativas, que podem ser declarativas, imperativas, negativas, optativas ou exclamativas. Nesse caso, a escolha por privilegiar a especificidade das sentenças interrogativas dá-se por ser um recurso argumentativo bastante utilizado na comunicação digital de marcas com seus clientes, na tentativa de “fisgar” o interlocutor e estabelecer um diálogo a partir da resposta solicitada.

Nesse caso, assumindo que a segunda pessoa irá emergir em contextos de diálogos diretos entre os interlocutores, em específico, na comunicação digital de cada uma das empresas com seu consumidor, enquanto interlocutor, percebemos que tais variáveis podem ser analisadas enquanto contextos de ordem linguística que podem condicionar a escolha entre as variantes de segunda pessoa do singular na posição de sujeito, como nos exemplos¹⁰ a seguir:

1- INTERROGATIVAS

a- Já pensou em dar um up no seu banheiro? (C¹¹)

b- E sabe onde você encontra esse produto incrível? (O)

2- NÃO INTERROGATIVAS

a- Com nossos serviços personalizados, você pode ter certeza de que terá uma safra de sucesso este ano. (O)

b- Marque aqui aquele professor que você adora e merece todas as homenagens nesse dia. (S)

4.2.1.1.2 Variável tempo verbal

¹⁰Todos os exemplos desta seção foram retirados do *corpus* do trabalho.

¹¹ As letras em cada um dos exemplos representam o código de identificação de cada uma das empresas que compuseram nosso *corpus*.

Considerada como relevante em estudos anteriores, como os de Magalhães (2001) e Loregian-Penkak (2004), entre outros, e entendendo que a relação da colocação pronominal com o verbo é algo intrinsecamente estabelecido pela língua, selecionamos como variável linguística o tempo verbal nas sentenças: passado¹² e não passado¹³.

É sabido que o verbo é a base geral das diversas relações sintáticas, e dada a sua importância dentro da língua, consideramos importante verificar se a pressão interna dada pela carga semântico-funcional do verbo pode, ou não, condicionar o uso de uma outra forma variante no preenchimento ou não da segunda pessoa do singular.

Sendo assim, buscamos averiguar se e de que forma o tempo da categoria verbal pode condicionar a escolha entre as variantes aqui analisadas, conforme sentenças como as exemplificadas abaixo:

3-PASSADO:

a- Você sabia que na Carajás é possível realizar sua compra com praticidade e economia? (C)

5- NÃO PASSADO:

a- Diz aí pra gente qual sabor do Refresco Coringa você vai saborear nesse domingo? (G)

b- Se você está procurando aquele prato com cara- e sabor!- de restaurante chique, bem-vindo a esse nhoque. (S)

4.2.1.1.3. A variável modo verbal

Como já mencionado em outros trabalhos como o de Lorengial-Pekal (1996), os modos verbais parecem incidir como condicionadores das formas de segunda pessoa do singular na posição de sujeito. Assim, com o objetivo de

¹² Nessa variável, agrupamos todos os tempos dos verbos com noção de passado, ou seja, pretérito perfeito, mais que perfeito e imperfeito, independente do modo verbal.

¹³ Aqui, agrupamos todos os tempos dos verbos com noção de presente e de futuro, independente do modo verbal.

verificar como ele pode atuar dentro do contexto da comunicação digital, nossos dados foram agrupados e analisados como demonstrado a seguir:

13-INDICATIVO

a-*Já imaginou (C)*

14- SUBJUNTIVO

a-*Se você souber (C)*

15- IMPERATIVO

a-*E corre pra Coagro (O)*

16- INFINITIVO

a- *Mais chances de ganhar (G)*

4.2.1.1.4. A variável paralelismo

A terceira variável selecionada foi o paralelismo que deve ser visto como a manutenção de uma mesma forma ao longo de um encadeamento discursivo. Scherre (1998, p.301) aponta que o efeito do paralelismo deve ser visto como uma “tendência geral de formas gramaticais particulares que ocorrerem juntas”, a autora ainda comenta que “sua forma de atuar, cria uma harmonia discursiva formal, tornando o discurso mais coeso”.

Scherre (*op cit*) destaca que o processo de repetição de variantes de uma mesma variável dependente tem se mostrado uma importante restrição no estudo de diversos fenômenos em várias línguas, podendo ocorrer no plano discursivo, oracional, sintagmático e ainda no plano da palavra.

Os fatores dessa variável elencados para a presente investigação foram os apresentados e exemplificados a seguir:

6- OCORRÊNCIA DE FORMA DE SEGUNDA PESSOA ÚNICA/ISOLADA

a- *Como você utiliza o leite de coco coringa? (G)*

7- MANUTENÇÃO DA MESMA FORMA DE SEGUNDA PESSOA

a- *Você sabia que na Coagro você encontra os melhores produtos da marca Jacto? (O)*

8- NÃO MANUTENÇÃO DA MESMA FORMA DE SEGUNDA PESSOA

a- *Se você é um apreciador de café de verdade, precisa experimentar nosso café em grãos. (G)*

Dessa maneira, para essa análise multivariacional, todos os dados foram codificados conforme o fator em que elas se enquadravam, ou seja, formas isoladas, paralelas ou não paralelas.

4.3. O *lócus* da pesquisa: a rede social Instagram

Para nossa investigação, definimos a rede social Instagram como *lócus* da pesquisa, nela, selecionamos quatro perfis públicos e oficiais de empresas do estado de Alagoas, buscando verificar como se dá a realização da segunda pessoa do singular na posição de sujeito em suas postagens que são dirigidas de forma bastante informal ao seu consumidor, enquanto interlocutor.

As diversas transformações de ordem social e econômica que aconteceram no mundo moderno geraram novos conceitos de comunicação entre as empresas e seu público consumidor. Em especial, com avanço das tecnologias digitais, o cenário corporativo percebeu a necessidade de inserir-se no universo das redes e mídias sociais na tentativa de estabelecer uma relação mais próxima com seu cliente e, assim, através de uma comunicação mais prática e direta, obter bons resultados comerciais.

De um modo geral, as redes sociais têm sido cada vez mais utilizadas para comunicação ao redor do mundo e o Brasil sempre se destaca por ser um dos países que mais utiliza as redes de um modo geral. Segundo Torres (2018),

essas redes são nomeadas dessa forma porque é algo compartilhado e aberto para interação e colaboração de todos, e ainda, porque são mídias que funcionam como meios de transmissão de informação e de conteúdo.

Para Silva *et. al.* (2009):

Redes sociais incitam mudanças profundas na maneira como os indivíduos se comunicam e relacionam, possibilitando às pessoas construir, interagirem e comunicarem entre si. Isso tem gerado mudanças profundas, que são experimentadas em todos os aspectos da vida social (Silva *et. al.*, 2019, p.9).

Através de um perfil em rede social, todo usuário consegue realizar interações através de sua rede e, ainda, instituir outras redes interacionais. Segundo o que Vaz (2010) defende:

A internet deixou de ser uma mídia para ser um ambiente. Uma brecha virtual no espaço-tempo no qual temos experiências de entretenimento, de troca e acúmulo de informações, de comunicação e de compras. Torna-se a cada dia em uma maneira de exercermos cada vez mais a nossa própria cidadania, a nossa própria condição humana na era da informação e conhecimento (Vaz, 2010, p. 415).

Em específico, a rede social Instagram é uma das maiores redes sociais da atualidade, com cerca de 2 bilhões de usuários, e permite a interação entre pessoas, bem como entre empresas e seu público consumidor. Apenas no Brasil, soma cerca de 99 milhões de pessoas que utilizam e acessam essa rede todos os dias, sendo o segundo maior país em número de usuários, ficando atrás apenas dos Estados Unidos e tendo no público mais jovem a maior fatia de uso e consequente popularidade.

O Instagram foi lançado em outubro de 2010 e foi desenvolvido por dois engenheiros de programação, o norte-americano Kevin Systrom e o brasileiro Mike Krieger. Sendo sucesso desde seu lançamento, alcançou importantes marcas, como, por exemplo, tornou-se o aplicativo mais baixado na Apple Store já no dia de seu lançamento.

Nele, os usuários compartilham conteúdos pessoais e profissionais, através de postagens (ou apenas posts) com mensagem escrita, vídeos e fotos, gerando assim arquivos que ficam armazenados em perfis e socializados na rede. Mesmo tendo a funcionalidade de acesso via computador, o principal acesso ao aplicativo é via aparelhos móveis, como o celular. Desde sua criação, essa era a intenção e atribui-se a essa facilidade, o enorme sucesso dessa rede social que permite a mobilidade e rapidez na interação independente do local em que esteja o usuário.

Em virtude de sua popularidade, tornou-se um grande aliado das empresas na busca do estreitamento da relação com seu consumidor, sendo responsável por divulgar imagem de produtos ou serviços, preço e demais informações relevantes e, ainda, ser capaz de ser um canal direto de comunicação entre o consumidor e a empresa e vice-versa, aumentando a visibilidade das marcas e gerando mais resultados.

4.4. A coleta dos dados e a constituição do *corpus*

O estudo na perspectiva variacionista se apoia sobre o paradigma quantitativo (Baylley, 2002; Guy, 1993), logo, para atingir o objetivo de apresentar de forma sistemática a variação ora analisada, buscamos coletar um conjunto de dados de forma criteriosa. A abordagem quantitativa permite que o linguista obtenha um panorama real sobre o fenômeno em estudo, através do estabelecimento de correlações entre as estruturas da língua e aspectos variáveis de ordem linguística e social.

Como já mencionado anteriormente, nosso intuito é observar o processo de variação entre o preenchimento e não preenchimento da segunda pessoa do singular na posição do sujeito, no âmbito da comunicação digital, através de postagens de empresas do estado de Alagoas em seus perfis na rede social Instagram.

Assim, o primeiro critério utilizado foi a origem alagoana de cada uma das empresas. Esse recorte da localização geográfica das empresas foi uma escolha metodológica de forma a trazer um panorama desse processo da comunicação

digital efetuada no estado, apresentando um cenário de forma mais unificada, diminuindo as distorções das marcas linguísticas de outros estados; bem como da possibilidade de observar os dados arrolados e comparar com resultados obtidos com os das pesquisas já existentes sobre o fenômeno na língua oral ou na escrita convencional.

Em específico, entre um grande conjunto de empresas do estado, a motivação para escolha das empresas selecionadas para busca de dados e constituição de nosso *corpus* deu-se pelo fato de estarem alocadas na lista entre as maiores empresas de Alagoas, levando em conta seu faturamento anual, segundo os dados disponibilizados pela plataforma Econodata (2023) que disponibiliza dados de empresas de todo país, a partir de recursos de pesquisa e segmentação de leads, através de dados extraídos da internet.

Outra escolha metodológica foi a de pertencerem ao ramo varejista, pois, nele, as empresas comunicam-se diretamente com seu consumidor final, seu interlocutor, e não com distribuidores ou outras empresas de um modo geral. Considerando que esses grupos empresariais possuem seu público majoritariamente no estado, acreditamos que a comunicação com seu consumidor, interlocutor, se dá através de usos linguísticos que denotem a ideia de proximidade, no estabelecimento de uma relação simétrica. Sobre a noção de relação simétrica, assumimos a definição de Briz (2004) que afirma que:

As relações simétricas são aquelas em que existe, ou se percebe, igualdade funcional e de papéis sociais entre os participantes da interação (idade, gênero ou profissão). As interações de maior proximidade ocorrem quando os interlocutores têm mais experiências ou saber compartilhado, **maior grau** de contato (físico ou ocular) e **de compromisso afetivo**. (Briz, 2004, p. 80, *grifo nosso*).

Especificamente no cenário do marketing atual, as empresas relacionam-se digitalmente e para tal utilizam propriedades linguísticas que viabilizem a identificação de seus clientes, enquanto comunidade linguística, com a marca/empresa.

Para Labov (1972), comunidade de fala é aquela que compartilha normas e 'atitudes' sociais perante uma língua. Ainda segundo o autor:

A comunidade de fala não é definida por nenhuma concordância marcada no uso de elementos linguísticos, mas sim pela participação num conjunto de normas compartilhadas; estas normas podem ser observadas em tipos de comportamento avaliativo explícito e pela uniformidade de padrões abstratos de variação que são invariantes no tocante a níveis particulares de uso. (Labov, 2008, p . 150).

A respeito disso, ao explicar sobre comunidade linguística, Alkimim (2003, p. 32) enfatiza que “Ao estudar qualquer comunidade linguística, a constatação mais imediata é a existência de diversidade ou da variação. Isto é, toda comunidade se caracteriza pelo emprego de diferentes modos de falar”. Dessa forma, temos que as variedades linguísticas são próprias de toda comunidade linguística e nosso intuito aqui é verificar como isso se processa no meio da comunicação digital.

Como condição para que as empresas fizessem parte desse conjunto que serve como base para extração dos dados, além de preencherem os dois já mencionados anteriormente, selecionamos as que possuem perfil público e oficial na rede social Instagram e que, também, buscam efetuar a comunicação de forma ativa e contínua, com postagens constantes, predominantemente dirigindo-se à segunda pessoa do singular, estabelecendo diálogo com seu cliente.

Obedecendo a esses critérios, selecionamos quatro empresas. A seguir, traremos uma caracterização geral do estado, apresentamos e situamos as empresas elencadas para nossa investigação.

4.5. Sobre as escolhas: o estado e as empresas alagoanas

O estado de Alagoas fica situado no Nordeste do Brasil, possui uma área territorial de cerca de 27.830, 661 km² e é composto por 102 municípios, com

população estimada em 3.365.351 habitantes, segundo o IBGE (2021), ocupando a 25ª posição no ranking de território geográfico do país.

Em específico, no que tange à economia, o estado possui atividades na indústria, pecuária, agricultura, extração de petróleo e turismo. No ramo industrial, sobressaem-se as que trabalham no subsetor de produtos químicos, beneficiamento da cana-de-açúcar para produção de açúcar e álcool, processamento de alimentos e cimento. Destacam-se o fumo, a cana-de-açúcar, o coco, arroz, abacaxi, mandioca, milho e feijão como os principais produtos cultivados em solo alagoano. As criações de equinos, bovinos, bubalinos, caprino, suínos e ovinos são as principais no território. Alagoas também possui reservas minerais de gás natural, salgema e petróleo que movimentam fortemente a economia.

Em relação à população, em Alagoas, 51,55% são mulheres e 48,45% são homens. Grande parte dos habitantes residem na área urbana, cerca de 73,6% e apenas 26,4% na zona rural.

Conhecendo essa caracterização do estado e reconhecendo o grande número de empresas existentes, o recorte metodológico de escolha obedeceu aos critérios já mencionados anteriormente, chegando às seguintes empresas: Carajás, Sococo, Grupo Coringa e Coagro.

A seguir, apresentaremos algumas informações relevantes sobre elas, extraídas diretamente do site oficial de cada uma, buscando caracterizá-las e apresentá-las:

A empresa Carajás ocupa a 24ª posição no ranking da Econodata (2023) em relação ao seu tamanho e faturamento anual que chega a mais de R\$340,000,000. Surgiu em 1973, em Maceió, Alagoas, e foi fundada por Álvaro Mendonça que iniciou a história da família no ramo da construção, através de uma Madeireira. Após cerca de quase três décadas, mais especificamente no início dos anos 2000, a empresa foi remodelada e expandida, dando origem à Carajás Home Center, que teve a frente os irmãos Abílio Alves, Alex Mendonça, Alexandre Mendonça e Antoine Mendonça.

Através de parcerias com as principais indústrias do segmento, a Carajás trabalha oferecendo produtos para o Lar, que vão desde material de construção até itens de decoração. A empresa reúne cerca de três mil colaboradores e além

de 4 unidades em Alagoas, possui outras lojas espalhadas pelo Nordeste. Atualmente, é considerada a 5ª maior empresa Home Center do Brasil, segundo o ranking da revista Anamaco de 2023.

Além do site (<https://www.carajas.com.br/>), a empresa comunica-se digitalmente com seus clientes, diariamente através de postagens, no perfil oficial no Instagram (@carajasonline), que soma mais de 500 mil seguidores, como pode ser visto a seguir:

Imagem 01: Perfil oficial do Instagram- Carajás



Fonte: Instagram

Ao longo do período de constituição do *corpus*, que será detalhado na subseção seguinte, selecionamos as postagens que geraram os dados para nossa análise.

A segunda empresa alagoana selecionada foi a Sococo, criada na década de 1960, na cidade de Maceió-AL, pelos irmãos Tavares. Os fundadores eram portugueses, radicados na capital do estado e conhecidos como comerciantes que trabalhavam com a negociação de coco ralado.

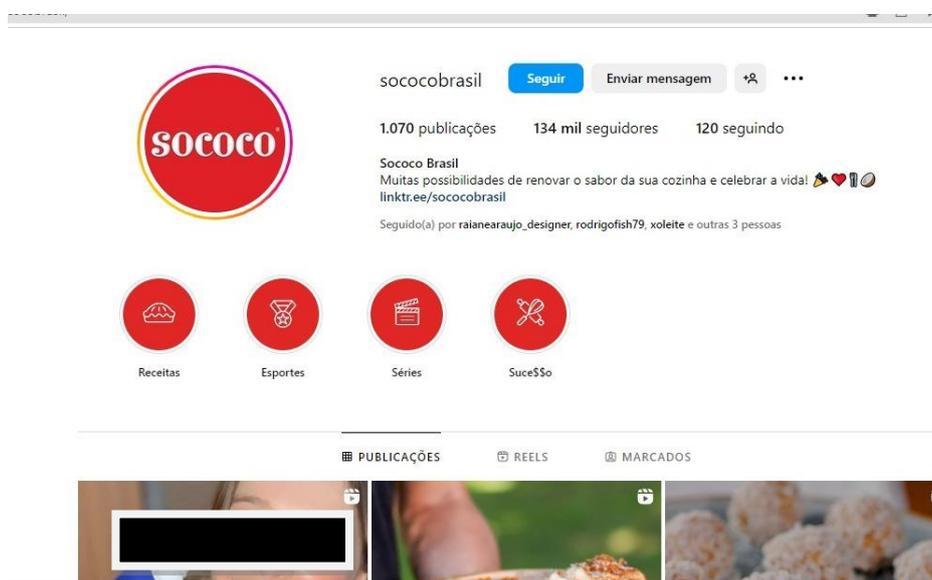
Através da fundação da empresa, objetivavam produzir e distribuir produtos derivados do coco para o Brasil e o exterior. Na época, esses produtos eram consumidos timidamente em virtude da baixa oferta no mercado e da produção quase que exclusiva apenas na região Nordeste, de maneira informal nos quintais das casas, principalmente na área litorânea.

Tendo em vista a mudança do cenário comercial dos produtos, a empresa cresceu e tornou-se líder de mercado. Atualmente, todas as etapas da cadeia produtiva dos derivados do coco são feitas pela empresa, desde o plantio e cultivo em coqueirais próprios até a comercialização. A Sococo conta com mais de cinco mil colaboradores atualmente e busca inovar na pesquisa e criação de novos produtos para otimizar sua matéria-prima.

Atualmente, a empresa ocupa a 32ª posição no ranking das maiores empresas do estado, tendo faturamento anual superior a R\$ 214,000,000, segundo o Econodata (2023).

A marca dialoga constantemente com seu público consumidor através das mídias digitais em seu site (<https://www.sococo.com.br/>) e em seu perfil na rede Instagram (@sococobrasil). Na rede social, acumula mais de 130 mil seguidores e diariamente interage com seu consumidor através de suas postagens:

Imagem 02¹⁴: Perfil oficial do Instagram- Sococo



Fonte: Instagram

¹⁴ Como apresentado nesta imagem e como segue nas demais no decorrer desta tese, os rostos ou demais traços que possam identificar humanos, bem como seus nomes ou identificações serão borrados através de faixas pretas, com o intuito de preservar as identidades que por ventura apareçam no conjunto de dados capturados, tendo em vista que o objetivo da pesquisa é tratar da comunicação efetuada pela(s) empresa(s) e não por indivíduos.

A terceira empresa elencada para nosso trabalho foi o Grupo Coringa, que possui faturamento anual de mais de R\$ 400,000,000, ficando na 22ª posição entre as maiores do estado, conforme aponta o Econodata (2023). Nasceu na década de 1960, em Arapiraca, Alagoas, através de um grupo que desenvolvia atividades com o beneficiamento de fumo para cigarros manuais. Já em 1970, inovou e ampliou sua área de atuação com a inserção de alimentos.

O Grupo Coringa cresceu e deu origem a uma ampla indústria que oportunizou a ampliação de sua capacidade de produção, juntamente com a inovação em equipamentos e tecnologia, dando destaque ao processo de refinamento de milho, tendo os flocos de milho como seu principal produto. Além dos alimentos derivados do milho, a empresa atua com outros produtos como café, corantes, leite de coco, arroz, creme de arroz e muitos outros, totalizando cerca de 30 produtos no total.

Um novo parque fabril foi adquirido em 1990, também em Arapiraca e em 2010 mais um parque industrial foi inaugurado, dessa vez na cidade de Luiz Eduardo Magalhães, na Bahia, ampliando sua participação no mercado.

Mantendo sua sede administrativa na cidade alagoana de Arapiraca, a empresa acumula mais de 53 anos de atuação, gerando empregos e desenvolvimento econômico para a região.

No âmbito da comunicação com seu consumidor, a empresa fornece informações e possibilidade de atendimento em seu site (<https://beacons.ai/grupocoringa>) e mantém ativa sua conta na rede social Instagram através de seu perfil oficial (@grupocoringa):

Imagem 03: Perfil oficial do Instagram- Grupo Coringa



Fonte: Instagram

A quarta empresa que utilizamos para nossa pesquisa foi a Coagro, conhecida como líder do ramo de produtos agropecuários, resguarda a 37ª posição entre as maiores empresas de Alagoas e, segundo o EconoData (2023), tem faturamento de mais de R\$ 145,000,000 ao ano.

Ela foi fundada em 1992, pelo engenheiro agrônomo Ricardo Dantas, como uma loja de varejo agropecuário na cidade de Arapiraca, interior de Alagoas. Em 1994, fortalecida no mercado, surge enquanto a Distribuidora Coagro.

Hoje em dia, a empresa trabalha com produtos agrícolas e veterinários e mantém sua sede administrativa em Arapiraca. Continua investindo em tecnologia e desenvolvimento de habilidades técnicas de suas equipes que atuam em suas 20 filiais distribuídas no Nordeste.

Através de seu perfil oficial (@grupocoagro) na rede social Instagram, que acumula cerca de 14 mil seguidores, a empresa estabelece comunicação com seu público consumidor constantemente:

Imagem 04: Perfil oficial do Instagram- Grupo Coagro



Fonte: Instagram

Na subseção a seguir, apresentaremos os procedimentos de coleta dos dados e constituição da amostra.

4.6. A amostra

Uma das características da pesquisa sociolinguística variacionista é o tratamento de um conjunto de dados representativos, analisados de forma quantitativa, a partir de um modelo sistemático de coleta de dados para constituição de uma amostra significativa do fenômeno em análise.

Assim, buscando cumprir essas exigências metodológicas, a amostra analisada nesta tese foi extraída das quatro empresas alagoanas já mencionadas acima e a coleta e seleção dos dados que constituem nosso *corpus* considerou o mesmo espaço temporal para coleta. Logo, coletamos os dados oriundos de postagens veiculadas em cada uma das páginas oficiais das empresas, no intervalo entre fevereiro de 2022 e maio de 2023.

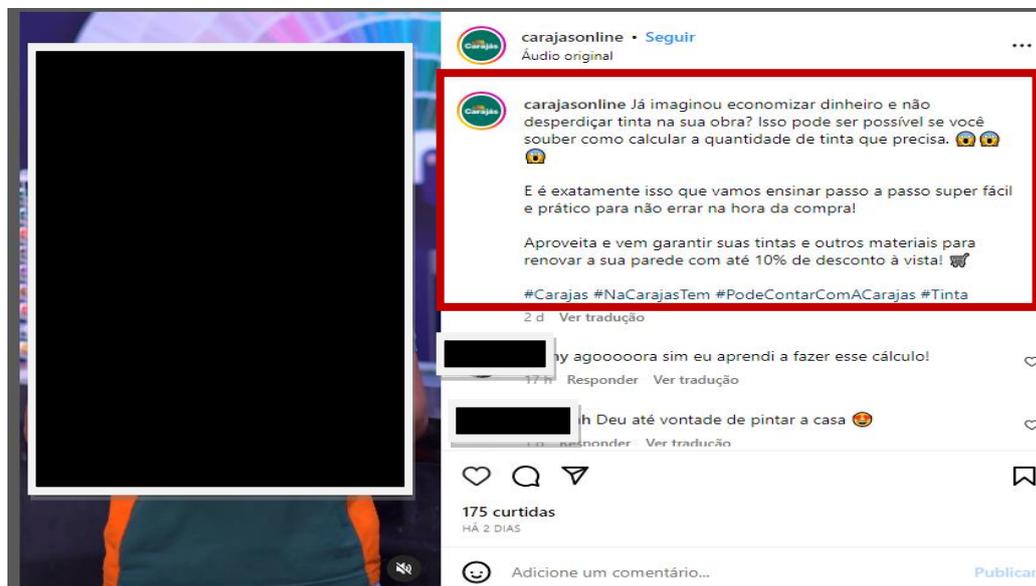
A partir de capturas de imagens diretas de cada uma das postagens, retiramos os dados linguísticos que continham a realização da segunda pessoa, de forma explícita, através de pronome de segunda pessoa na posição de sujeito, ou não explícita, demarcadas pela supressão pronominal ou do que assumimos aqui como pronome zero ou sujeito nulo com referência direta ao interlocutor do discurso na comunicação digital. Assim, selecionamos postagens como as exemplificadas abaixo:

Imagem 05: Postagem- COAGRO



Fonte: Instagram

Imagem 06: Postagem- Carajás



Fonte: Instagram

No conjunto que constitui nossa amostra, selecionamos tanto o material linguístico escrito na descrição da imagem, quanto o material linguístico que contenha a realização do fenômeno estudado na imagem da postagem, quando encontrado, como pode ser visto no próximo exemplo:

Imagem 07: Postagem- Coringa



Fonte: Instagram

A partir dessa coleta sistematizada, os dados foram transcritos e codificados através de uma representação específica de cada uma das empresas para fins de organização e controle. Dessa forma, as empresas foram representadas com os seguintes códigos: Carajás (C), Sococo (S), Coagro (O) e Grupo Coringa (G). É importante destacar que nossa análise não considera cada empresa como uma variável, mas sim, toma os seus dados como um conjunto representativo da comunicação digital de empresas alagoanas.

Após transcrição e codificação dos dados, foram alocados em planilha Excel para contagem e demarcação do conjunto de variáveis que foram analisadas. Em suma, o quantitativo total compreende 487 contextos oracionais para nossa análise, como se observa na tabela a seguir:

TABELA 03: QUANTITATIVO GERAL DOS DADOS

Empresas	Variante 1 Você	Variante 2 Sujeito Nulo	Row Total
Carajás	58 31.4%	127 68.6%	185 38.0%
Coagro	49 38.0%	80 62.0%%	129 26.5%
Grupo Coringa	49 44.1%	62 55.9%	111 22.8
Sococo	21 33.9%	41 66.1%	62 12.7%
Total	177	310	487

Fonte: Elaborado pela autora a partir do software R

Como pode ser visto na tabela acima, os dados coletados em cada um dos perfis não compõem um mesmo quantitativo numérico e isso não representa um problema para análise, pois tratamos todos os dados como um único conjunto representativo, como já mencionado anteriormente e, primordialmente, por serem distintos em virtude da periodicidade e tamanho das postagens e, ainda, do tipo de produtos ou postagem que vinculam em suas redes.

Seguindo com os procedimentos metodológicos, após identificação e coleta dos dados, foram identificadas cada uma das variáveis linguísticas elencadas para nossa análise, já apresentadas anteriormente na subseção 3.2.11, a saber: 1) tipo de sentença, sendo interrogativas e não interrogativas; 2) tempo verbal; 3) paralelismo e 4) modo verbal. Fechando assim todo processo de caracterização de cada um dos dados em análise, codificamos cada uma da seguinte forma:

Tabela 04: Codificação das variáveis

Tipo de sentença	Tempo Verbal	Paralelismo	Modo Verbal
I-Interrogativa	p- Passad o	M-Manutenção das formas	IND- Indicativo
N- Não interrogativa	n- Não passad o	U-Ocorrência única/isolada	INF- Infinitivo
		D- Distinção entre as formas	SUB- Subjuntiv o
			IMP- Imperativ o

Fonte: Elaborada pela autora

Para a organização das ocorrências e de cada uma das variáveis linguísticas adotadas para o trabalho, preenchemos uma planilha no programa Excel com os seguintes dados: ocorrência transcrita tal qual a postagem coletada, identificação da empresa, variável dependente, variáveis linguísticas independentes e “palavra” que representa explícito ou implicitamente a referência à segunda pessoa do singular na posição de sujeito.

O passo seguinte foi tratar os dados estatisticamente, a partir de seus percentuais e pesos relativos em termos de ocorrências, através do software R, que será descrito na subseção seguinte.

4.7. O tratamento dos dados através do software R

Tendo em vista que a metodologia laboviana prima pelo viés quantitativo da análise, o sistema para avaliar os dados estatisticamente é extremamente importante para sistematização dos achados da pesquisa.

Através dos resultados apresentados numericamente, quer seja percentualmente e/ou estatisticamente, é possível estabelecer os que são consideravelmente relevantes e os resultados menos significativos que por ventura vão emergir do tratamento dado aos dados colhidos. De acordo com Scherre e Naro (2008) :

[...] Os resultados numéricos obtidos pelos programas só têm valor estatístico. O seu valor linguístico é atribuído e interpretado pelo linguista. Se o linguista for bom, certamente os resultados lhe permitirão refutar ou não as hipóteses estabelecidas quando da análise dos dados linguísticos. (Scherre e Naro, 2008, p. 162)

Dessa forma, em termos estatísticos, podemos comprovar quais variáveis são condicionadoras ou não dentro do processo de variação entre o preenchimento ou não da segunda pessoa na posição de sujeito na comunicação digital de empresas em Alagoas.

Nas últimas décadas, a ciência linguística vem se beneficiando do tratamento matemático dado aos fenômenos linguísticos. Em especial, a Sociolinguística, justamente por sua abordagem descritiva e quantitativa de tratamento dos dados é uma das correntes que mais recorrem ao tratamento dado pelos programas computacionais de análise estatística. O interesse por análises a partir de variáveis e a necessidade de agrupamento de um conjunto representativo, em termos quantitativos, de dados fazem com que os sistemas computacionais sejam cada vez mais eficientes e necessários para garantia do rigor científico.

Para nosso trabalho, definimos o software R como o responsável pelo tratamento estatístico dos dados, por ser fácil acesso, moderno, gratuito e capaz de atender às demandas necessárias nesta pesquisa. Em suma, de acordo com Aquino (2013):

O R possui uma enorme quantidade de procedimentos estatísticos em milhares de pacotes livremente disponíveis na internet e que podem ser carregados opcionalmente. [...] Existem dezenas de interfaces para o R. Algumas exigem a memorização de numerosos comandos; outras oferecem uma interface com vários botões e itens de menu clicáveis (Aquino, 2013, p. 03)

Um estudo sociolinguístico, a partir do tratamento de dados através do sistema R, pode observar e mensurar o peso das variáveis envolvidas no processo de variação linguística.

Especificamente, em relação aos métodos utilizados para nossa análise no R, através da interface Rstúdio que funciona a partir de scripts computacionais com funções específicas para probabilidade, tratamento estatístico, geração de dados e, através dessa interface, vários cálculos podem ser realizados. Os gráficos apresentando os resultados em nossa análise foram gerados pelo pacote “ggplot2” e também por meio da geometria “geom_col”.

O teste estatístico de Análise de Variância (ANOVA) foi utilizado com o objetivo de conferir se há diferença significativa entre as variantes a partir das variáveis analisadas. Os dados também foram observados a partir do modelo de regressão multinível com efeito misto que as variáveis observadas possuem uma hierarquia, ou seja, são dependentes.

Através do R, conseguimos visualizar o percentual das ocorrências de cada uma das variantes, bem como o p-valor da variável independente testada. Nesse caso, quanto menor for o p-valor da variável em análise, maior será o nível de relevância estatisticamente falando, por outro lado, quanto maior for o p-valor, menos significativa essa variável será.

Dessa forma, nossa análise está pautada nos achados gerados durante as rodadas realizadas no sistema, avaliando o papel de cada uma das quatro variáveis elencadas no trabalho e que nos trouxeram valores estatísticos que serão arrolados em nossa discussão que segue na próxima seção.

5. APRESENTAÇÃO E ANÁLISE DOS DADOS

Na presente seção, voltaremos nossa atenção para a apresentação, descrição e análise dos dados colhidos nas postagens das empresas alagoanas, selecionadas para nossa investigação, acerca da variação no preenchimento ou não do sujeito de segunda pessoa do singular. Definiremos os termos percentuais de uso das variantes dependentes e independentes elencadas para nosso trabalho, traremos alguns dados a título de exemplificação dos quantitativos apresentados e faremos a discussão e análise dos achados a partir do *corpus* utilizado.

5.1. AS VARIÁVEIS DE SEGUNDA PESSOA DO SINGULAR NA POSIÇÃO DE SUJEITO: VISÃO GERAL

Após a coleta dos dados, obtidos através de prints das postagens das empresas, iniciamos a contabilização dos dados relevantes para nosso estudo. Nesse caso, primeiramente, selecionamos, para nosso *corpus*, as sentenças em que os participantes utilizaram forma pronominal expressa ou nula fazendo referência à segunda pessoa do singular¹⁵.

Ao avaliarmos a distribuição total das ocorrências em nossos dados, nos deparamos com um achado bastante significativo: em todos os contextos de segunda pessoa do singular na posição de sujeito, apenas a forma pronominal “você” foi encontrada na representação de pronome pleno. A forma “tu”, bem como todas as demais que poderiam referenciar a segunda pessoa do singular na posição de sujeito não emergiram em nosso conjunto de dados.

Sobre o não aparecimento da forma de preenchimento “tu” em nossos dados, podemos relacioná-lo com o que Celso Cunha e Luís Felipe Lindley Cintra (2001) afirmam sobre o emprego da forma “tu” no Brasil:

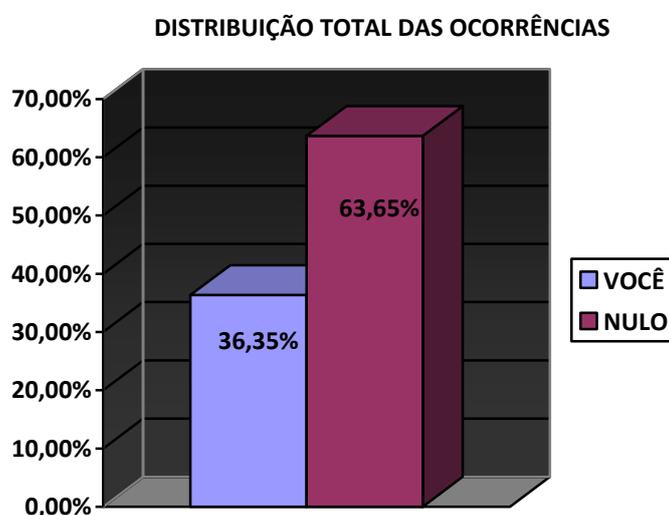
(...) restringe-se ao extremo Sul do país e a alguns pontos da região norte, ainda não suficientemente delimitados. Em quase todo o território brasileiro, foi ele substituído por você como forma de intimidade. Você

¹⁵ Destacamos que, embora não seja objeto de nossa pesquisa, as demais formas de segunda pessoa (tu, ocê, cê, senhor) também não foram localizadas na posição de objeto.

também se emprega, fora do campo da intimidade, como tratamento de igual para igual ou de superior para inferior. (Cunha e Cintra , 2001, p. 305).

Assim, após essa constatação, tivemos a seguinte caracterização entre as variantes dependentes: Variante 1 -Você- 177 ocorrências e Variante 2- Sujeito Nulo- 310 ocorrências. Vejamos os percentuais no gráfico a seguir:

Gráfico 01- Total das ocorrências



Fonte: Elaborado pela autora

Conforme Bagno (2012, p. 750), “o pronome tu, tendo sofrido a concorrência desleal com *você*, acabou se confinando a determinadas variedades regionais e/ou determinadas camadas sociais”. Esse argumento pode ajudar a sustentar o não aparecimento do “tu” nos dados arrolados em nossa investigação. Para o autor:

“O índice **tu**, por ser menos difundido, é uma forma **marcada**, ao passo que **você** é **não marcada**. Por isso, toda a publicidade, todos os meios de comunicação e todas as instituições oficiais empregam **você** na interlocução virtual com a cidadã e o cidadão comuns” (BAGNO, 2012, p. 752, grifos do autor).

Nesse sentido, observando comparativamente a grande quantidade de realizações da forma “você” em detrimento de nenhuma realização da forma “tu”, os dados levantados no campo da comunicação digital comprovam esse processo de substituição e de força com o qual o “você” lança-se no sistema pronominal do português para além da fala e da escrita convencional. Tal achado pode ser enxergado como uma pista do processo de mudança no qual o pronome “você” passa a ocupar, definitivamente, o lugar de referência à segunda pessoa em contextos de comunicação “menos monitorada” ou menor nível de formalidade.

Observemos, agora, postagens analisadas com as duas variantes encontradas:

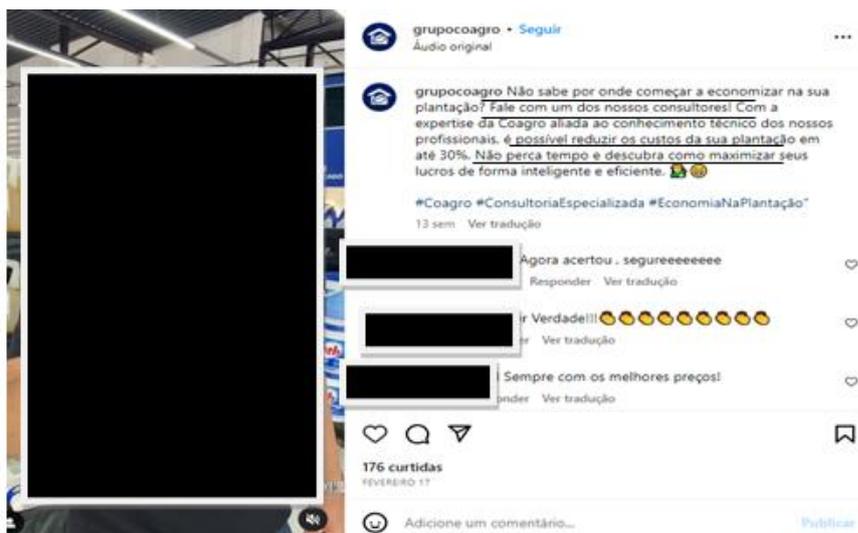
IMAGEM 08: SUJEITO PRONOMINAL DE SEGUNDA PESSOA



Fonte: Instagram

Como podemos observar, tanto na legenda da postagem em si quanto na imagem, o sujeito foi preenchido com o pronome “você”. Vejamos, agora, exemplo de postagem com a utilização do sujeito nulo:

IMAGEM 09: OCORRÊNCIAS DE SUJEITO NULO



Fonte: Instagram

Como vimos na imagem 09, o sujeito nulo foi utilizado em diversas oportunidades para referenciar a segunda pessoa do singular, nesse caso, o cliente/consumidor. Observamos que a comunicação realizada digitalmente pela empresa alcança seu objetivo ao passo que gera comentários acerca do que é informado. Portanto, apesar do sujeito não ser viabilizado expressamente através de pronomes de segunda pessoa, o sujeito nulo não prejudica a comunicação, gerando sentenças perfeitamente gramaticais no português brasileiro e sendo reconhecida como tal pelos interlocutores, gerando interação.

Para além da ausência de realizações da forma “tu”, nossos dados revelam que até mesmo o pronome “você” não estaria sendo a opção mais recorrente de uso para a função de sujeito, perdendo força e dando espaço para o sujeito nulo no português brasileiro, especificamente, quando observado esse gênero da comunicação digital em estudo.; tendo o pronome nulo um percentual de realizações bem superior, 63,65% em relação ao você, 36,35%, como vimos.

O processo de variação entre o preenchimento pronominal “você” e o sujeito nulo na comunicação digital pode ser encarado como um dado relevante no que tange à distribuição das possibilidades de referência à segunda pessoa. De um modo geral, as diversas pesquisas já realizadas acerca da variação na segunda pessoa do singular voltam o olhar, de modo específico, para a variação

entre as formas pronominais explícitas, não observando e até descartando os dados que apresentavam esse resultado, como Martins (2010) e Silva (2019), por exemplo. Contudo, nas postagens analisadas, merece destaque essa realização da segunda pessoa de forma não explícita através de um pronome.

Sobre o não preenchimento do sujeito, as pesquisas acerca do sujeito nulo no PB revelam certa dificuldade no tratamento da segunda pessoa do discurso, tendo em vista os aspectos metodológicos que dificultam a coleta dos dados e, ainda, sobretudo, o processo de inter-relação entre a 2ª e 3ª pessoas no que tange à morfologia verbal.

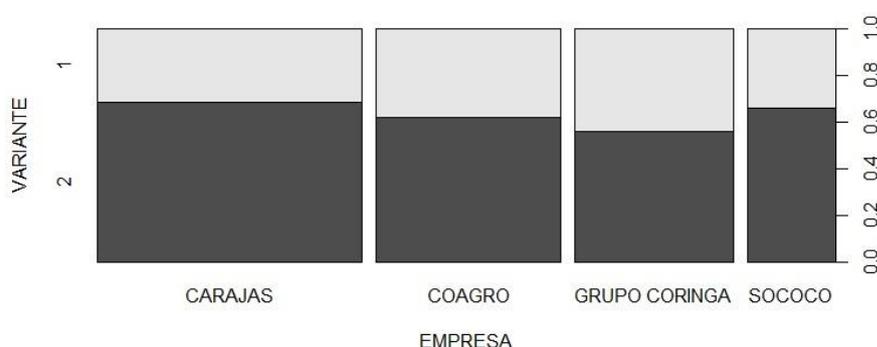
Segundo Duarte (1993), a 2ª pessoa lidera a mudança no caso do não preenchimento do sujeito pronominal, tendo sido a primeira pessoa do discurso a ser afetada pela mudança e a que revela percentuais mais significativos na transição do sujeito nulo para o sujeito explícito. Contudo, nossos dados refutam essa afirmação, para o gênero analisado na comunicação digital, tendo maior expressividade a forma nula no conjunto das ocorrências analisadas.

Refletindo sobre isso, observemos o que Duarte apresenta no estudo realizado em 1995, acerca do preenchimento do sujeito pronominal explícito:

que o português do Brasil perdeu a propriedade do Princípio 'Evite Pronome' levando-se em conta o enfraquecimento da flexão, responsável pela identificação da categoria vazia sujeito. Como consequência dessa perda, a língua caminha na direção das línguas que têm a posição sujeito preenchida (Duarte, 1995: 30 e 134).

Ao voltarmos nosso olhar para os dados analisados, é interessante observar que os resultados refutam a proposta de Duarte no que tange à realização do sujeito de segunda pessoa. É visível que a variação aqui encontrada acontece de forma sistemática e recorrente, pois quando observamos os dados de cada uma das empresas, mesmo que de forma isolada, todas mantêm o maior número de realizações da variante “sujeito nulo” em detrimento do pronome “você”, conforme pode ser visto a seguir:

GRÁFICO 2: DISTRIBUIÇÃO DAS VARIANTES POR EMPRESA¹⁶



Fonte: Software R

A predominância do não preenchimento do sujeito com elemento pronominal e, conseqüentemente, uso do sujeito nulo, apontada através dos resultados de nosso trabalho, segue os achados do trabalho de Silva (2021) no que se refere ao quantitativo de maior número de ocorrências do sujeito não explícito morfologicamente dentro do *corpus* analisado.

Labov (2003) define que há três tipos de regras linguísticas: uma regra que é categórica, com frequência de realização de 100% de uma única forma; uma regra que é semicategórica, com 95% a 99% de frequência; e uma regra variável, com frequência entre 5% a 95%. Em suma, seguindo a proposta laboviana e observando os resultados encontrados no que tange à distribuição geral, nosso fenômeno em estudo na comunicação digital compreende uma regra variável entre o preenchimento do sujeito pronominal com a forma “você” e o sujeito nulo no âmbito da segunda pessoa do singular, com maior frequência do não preenchimento do sujeito pelo elemento pronominal. Tal resultado é fruto da relação direta de uma característica licenciada pela língua com as características da comunicação digital, como a dinâmica da rápida interação e a supressão de elementos linguísticos para uma comunicação mais ágil, acessível e que pode oscilar entre o formal e o informal, conforme apontado em Pinheiro (2021).

¹⁶ A variante 1 corresponde ao pronome “você” e a 2 ao sujeito nulo de segunda pessoa do singular em toda análise.

Na próxima subseção, apresentaremos os resultados encontrados através da análise de cada uma das variáveis independentes elencadas para a presente investigação.

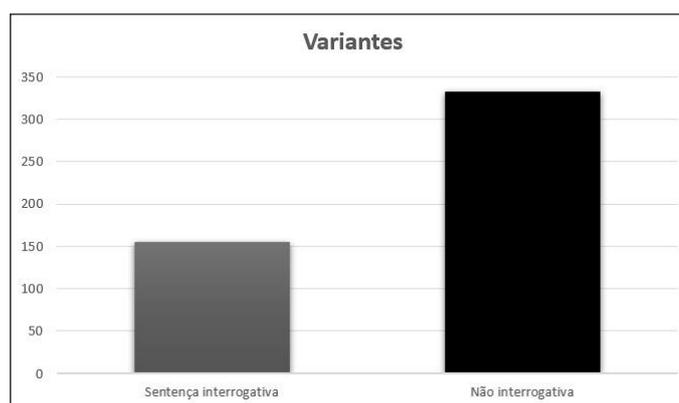
5.2. O PAPEL DAS VARIÁVEIS INDEPENDENTES

5.2.1. Variável tipo de sentença

Buscando avaliar o papel da variável tipo de sentença, como já mencionado, selecionamos os fatores sentenças interrogativas e não interrogativas e através do software R, obtivemos resultados importantes.

Em termos de distribuição geral, o *corpus* apresentou um maior percentual de sentenças não interrogativas. Tal resultado já era esperado, tendo em vista que as não interrogativas englobam todos os demais tipos de sentenças, como informamos anteriormente. A seguir, observemos a distribuição geral dos tipos de sentenças:

GRÁFICO 3 : DISTRIBUIÇÃO TOTAL DOS TIPOS DE SENTENÇAS



Fonte: Software R

O conjunto de sentenças que constitui nosso *corpus* teve 155 do tipo interrogativa e 332 do tipo não interrogativa. Observamos alguns exemplos das ocorrências analisadas:

IMAGEM X: VARIANTE VOCÊ EM SENTENÇA INTERROGATIVA



Fonte: Instagram

Agora, vejamos um exemplo de contexto interrogativo com sujeito nulo na imagem abaixo:

IMAGEM X: VARIANTE SUJEITO NULO EM SENTENÇA INTERROGATIVA



Fonte: Instagram

Ao passo que adentramos no universo de ocorrências de cada um dos tipos de sentenças, verificamos que a variante 2, equivalente ao sujeito nulo, é a mais recorrente em ambos os fatores. Logo, o não preenchimento do sujeito pronominal de segunda pessoa tende a ocorrer tanto nas sentenças interrogativas como nas não interrogativas, como pode ser visualizado no gráfico abaixo:

GRÁFICO 4: DISTRIBUIÇÃO DAS VARIANTES POR TIPO DE SENTENÇA



Fonte: software R

Para visualizar de maneira ainda mais clara, observemos a tabela a seguir com os valores percentuais e quantidade de ocorrências:

TABELA 4: PERCENTUAIS DE REALIZAÇÃO DAS VARIANTES NO TIPO SENTENÇA

Fatores	Variante 1 Você	Variante 2 Sujeito nulo	Row Total
I	75 48.4%	80 51.6%	426 87.5%
N	102 30.7%	230 69.3%	332 68.3%
Total	177	310	487

Fonte: Elaborado pela autora a partir do software R

Contudo, analisando os percentuais apresentados, percebemos que as sentenças interrogativas manifestam muito mais o uso do pronome pleno em termos percentuais, 48.4%. Ainda assim, os valores apontam uma discreta

distinção, tendo pouco menos de 3% de diferença entre o sujeito nulo e o sujeito pronominal explícito “você” quando observamos apenas os contextos sentencias interrogativos.

No caso das sentenças não interrogativas, os resultados apontam para uma diferença muito mais significativa em termos de distribuição das variantes, tendo o sujeito nulo mais do que o dobro de número de ocorrências em relação ao você, 30,7% e 69.3% respectivamente. Nesse caso, concluímos que o contexto de sentenças não interrogativas favorece o uso do sujeito nulo.

A análise de Duarte e Reis (2018, p. 177), no que se refere ao processo de mudança no português entre o pronome nulo e o pronome pleno, indica que a “2ª pessoa se apresenta como um caso de mudança quase concluída, com sujeitos nulos restritos a contextos em que são pragmaticamente identificados, [...] como perguntas”. Entretanto, nossos resultados não confirmam tal afirmação na amostra analisada, pois os valores percentuais tão próximos indicam que, de fato, o processo de variação ocorre de forma bem distribuída entre as variantes. Já no âmbito das não interrogativas fica claro o maior número de ocorrências do sujeito nulo.

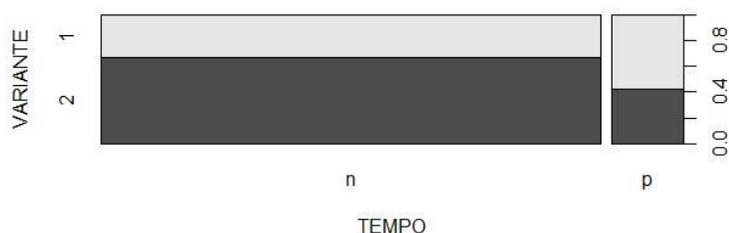
5.2.2. Variável tempo verbal

Seguindo com a avaliação do papel das variáveis no processo de variação entre o sujeito pronominal de segunda pessoa do singular e o sujeito nulo, temos a variável tempo verbal, tida, em estudos como o de Santos (2014) e Almeida e Assunção (2008), como relevantes para o estudo desses processos de variação.

De acordo com Lyons (1977, p.677), tempo verbal é uma categoria dêitica, bem como uma categoria que também pode ser depreendida da sentença. Logo, enquanto um aspecto constitutivo importante do elemento verbal- que funciona no português como referência de pessoa em contextos de sujeito nulo- torna-se pertinente avaliá-lo como um possível condicionante no processo de investigação que propomos.

Para análise dessa variável, selecionamos os fatores passado (p) e não passado (n). No corpus de nossa investigação, encontramos o seguinte resultado no que tange à distribuição das ocorrências:

GRÁFICO 5: DISTRIBUIÇÃO DAS VARIANTES POR TEMPO VERBAL



Fonte: Instagram

Ao examinarmos cada um dos fatores, temos na variável passado a predominância de ocorrências de preenchimento do sujeito com pronome “você” e já na variável não passado ocorre o inverso, é o sujeito nulo que lidera o número de ocorrências. Vejamos alguns contextos dessas ocorrências:

IMAGEM X: VARIANTE VOCÊ NO TEMPO VERBAL PASSADO

carajasonline • Seguir

carajasonline Você já ouviu falar na decoração sertão? Minha gente, ela é a cara do Brasil! 🇧🇷

Essa tendência de decoração busca valorizar as raízes do nosso país através de móveis de madeira, tons mais terrosos e esverdeados, itens com bambu, fibras naturais e muito mais.

🛒 Já aproveita e enche o seu carrinho para renovar a decoração da sua casa do jeitinho que sempre quis!

#PraTodosVerem: o post carrossel conta com 5 imagens, onde a capa tem uma cadeira de madeira com almofadas em um ambiente com cortina, papel de parede com folhas, lâmpadas e o texto "Decoração do sertão: dicas para fazer em casa. A segunda imagem é de uma estante feita de madeira com almofadas com tons terrosos e estampas, além de ter cestos organizadores com toalha dentro. Na terceira imagem, há um vaso de plantas com o sol destacando o objeto com o texto "Vasos de plantas também são ótimos aliados na hora de compor a decoração do sertão!". Na quarta imagem, há uma cadeira de madeira com uma almofada estampada com o texto

106 curtidas
JANEIRO 28

Adicione um comentário... Publicar

Fonte: Instagram

IMAGEM X: VARIANTE SUJEITO NULO NO TEMPO VERBAL NÃO PASSADO



Fonte: Instagram

Observando as possibilidades de realização de segunda em cada um dos fatores temporais selecionados, observamos que a distribuição das ocorrências de segunda pessoa na comunicação digital não acompanha o que outros trabalhos pontuam, apresentando características específicas pra essa forma de uso da língua. A seguir, observemos os resultados percentuais revelados em nossa amostra:

TABELA 5: PERCENTUAIS DE REALIZAÇÃO DAS VARIANTES POR TEMPO VERBAL

Fatores	Variante 1 Você	Variante 2 Sujeito nulo	Row Total
n	142 33.3%	284 66.7%	426 87.5%
p	35 57.4%	26 42.6%	61 12.5%

Total	177	310	487
-------	-----	-----	-----

Fonte: Elaborado pela autora a partir do software R

O trabalho de Nunes (2000) aponta que os contextos de passado revelam mais ocorrências de sujeito nulo, contudo, em nossa amostra na comunicação digital, especificamente na segunda pessoa do singular, o fator passado apresentou um percentual de preenchimento pronominal superior ao do apagamento do sujeito. Já o fator que reúne os demais tempos verbais apresentou um percentual com o dobro de realização do sujeito nulo.

Santos (2014) destaca que nos dados analisados em sua pesquisa, o fator tempo verbal passado favoreceu o sujeito nulo, enquanto que o presente propiciou o maior preenchimento por pronome, especificamente na 1ª e 3ª pessoa, mas a autora não analisou a segunda pessoa. De toda forma, acreditamos que podemos inferir que, no âmbito do preenchimento do sujeito via elemento pronominal, nossos resultados se distanciam.

Em suma, os resultados oriundos da análise de nosso *corpus* não congregam com os apontados em outros trabalhos. Nesse sentido, torna-se plausível destacar que isso ocorre em virtude da modalidade de comunicação e da pessoa do discurso colocada aqui em evidência.

5.2.3. Variável modo verbal

A escolha pela análise da variável modo verbal se deu em virtude de apresentar distinções no que tange à realização da segunda pessoa, especialmente entre as formas pronominais, como nos trabalhos de Lorengian-Penkal (2005) e Costa (2013). Contudo, como o elemento verbal está intimamente ligado à realização do sujeito nulo, decidimos ampliar o escopo de análise dessa variável com o intuito de observar de que forma ela se relaciona com o preenchimento do sujeito pronominal e do sujeito nulo de segunda pessoa.

Outro aspecto relevante para essa escolha se deu em virtude da relação direta de alguns modos verbais com o próprio gênero analisado, as postagens de empresas na rede social Instagram. Nesse caso, as marcas linguísticas que

induzem o interlocutor, cliente/consumidor, poderiam apresentar distinções relevantes para a análise aqui proposta, como é o caso do uso do modo imperativo, por exemplo. Sobre isso, de acordo com Rosa (2003), os imperativos podem assumir uma ideia de ordem, de indicação ou de sugestão sobre o que e qual a melhor forma de realizar uma determinada ação.

Vejam a distribuição dos modos verbais de nossa amostra na tabela seguinte:

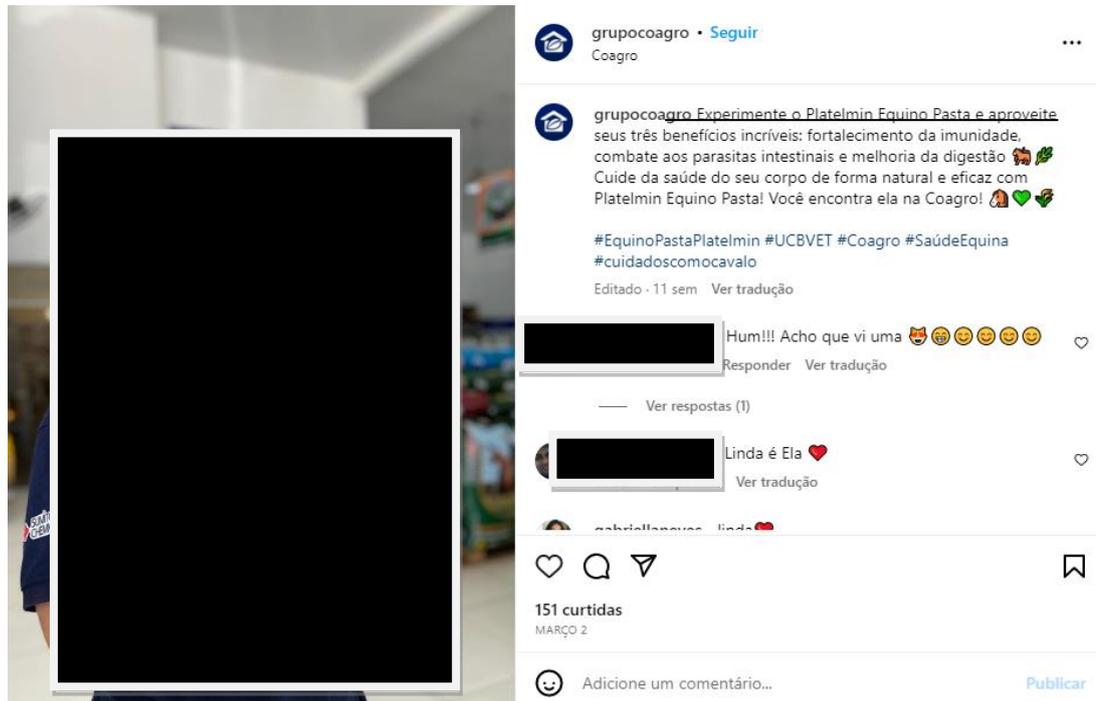
TABELA 6: PERCENTUAIS DE REALIZAÇÃO DAS VARIANTES POR MODO VERBAL

Fatores	Variante 1 Você	Variante 2 Sujeito nulo	Row Total
INDICATIVO	146 62.7%	87 37.3%	233 47.8%
IMPERATIVO	4 3.1%	123 96.9%	127 26.1%
SUBJUNTIVO	24 22.9%	81 77.1%	105 21.6%
INFINITIVO	3 13.6%	19 86.4%	22 4.5%
Total	177	310	487

Fonte: Elaborado pela autora a partir dos resultados do software R

Como visualizado acima, as variantes 1 e 2 são distribuídas em todos os modos verbais. Porém, a distribuição interna em cada um dos fatores difere bastante entre eles. O modo indicativo lidera com o maior número de ocorrências e apresenta o maior uso do preenchimento do sujeito com pronome “você”. Os demais modos verbais trazem a predominância do sujeito nulo. Entre eles, damos destaque ao modo imperativo que apresenta de forma muito significativa a predominância do sujeito nulo com 96.9% das ocorrências em detrimento de apenas 3.1% de preenchimento do pronome pleno nessa posição com referência de segunda pessoa. O modo subjuntivo ocupa a terceira posição em relação ao número de ocorrências e, por último, temos o infinitivo com apenas 4.5% de ocorrências no total da amostra. A seguir, vejamos algumas ocorrências que representam esses achados em nosso *corpus*:

IMAGEM X: VARIANTE SUJEITO NULO NO MODO VERBAL IMPERATIVO



Fonte: Instagram

IMAGEM X: VARIANTE VOCÊ NO MODO VERBAL SUBJUNTIVO



Fonte: Instagram

Conforme já mencionado, é interessante avaliar como o modo imperativo mantém a supremacia do recurso do sujeito pleno com praticamente 97% das ocorrências. Apesar de não ser o mais utilizado ao longo dos dados do *corpus*, possui destaque pelos resultados apontados. Na gramática de Bechara (2003, p. 237), define que “O imperativo em português só tem formas próprias para as segundas pessoas; as pessoas que faltam são supridas pelos correspondentes do presente do subjuntivo”, contudo apesar dessa condição explicitada, os dados da nossa amostra figuram com sua maioria das ocorrências de sujeito nulo com marcas verbais de terceira pessoa, resultado do processo pelo qual essas duas formas se inter-relacionam ao longo do tempo, a partir da inserção do “você” no paradigma pronominal do português.

Segundo Scherre (2005), nos diálogos ou instruções no singular, de um modo geral, dirigidos a uma só interlocutor, na língua falada corriqueiramente em Brasília, Minas Gerais, Goiás e Rio de Janeiro, aproximadamente 90% dos enunciados imperativos são expressos pelas formas hoje associadas à 3^o pessoa do modo indicativo. Ainda sobre o imperativo, avaliando sua realização especificamente no Nordeste, região que engloba nossa amostra, Bagno (2001) indica que:

é muito mais provável escutarmos um diga, venha, corra, faça na fala de um baiano, de um pernambucano ou de um paraibano do que as formas diz, vem, corre, faz (embora, como venho insistindo, elas também ocorram, já que a pesquisa linguística trabalha sempre com maior ou menor frequência e nunca com formas de uso absoluto e exclusivo. (Bagno, 2011, p. 567)

Ainda de acordo com o autor, nos locais onde o uso do “você” é predominante, o imperativo, de um modo geral, é associado ao indicativo, que também faz referência ao “tu” no imperativo. Em síntese:

Por outro lado, nas regiões onde se usa amplamente o você, a forma do imperativo mais freqüente é a que, supostamente, deveria corresponder ao uso de tu: vem,

faz, diz, pega, toma etc. Por isso, um mineiro ou paulista pode dizer, tranquilamente: —Eu já falei, agora fala vocêll, e não —fale vocêll como exigem as regras tradicionais. (Bagno, 2011, p. 568).

Em relação ao subjuntivo, Nunes (2000) destaca que durante a sua análise sobre a realização do sujeito nulo na fala em João Pessoa, os dados apontaram que no modo subjuntivo, independente do tempo verbal, o sujeito nulo foi desfavorecido, dando lugar ao preenchimento do sujeito com pronome. Diferentemente nos nossos achados que apresentam uma variação de mais de 50% de sujeito nulo em relação ao pronome explícito.

5.2.4. Variável paralelismo

A quarta e última variável selecionada para nossa investigação é o paralelismo entre as formas. Referenciada como uma importante variável tanto no processo de variação entre formas pronominais de segunda pessoa que preenchem a posição de sujeito como no estudo do sujeito nulo (Scherre, 1998; Costa, 2013; Santos , 2012; Martins, 2010; Silva, 2019).

De acordo com Silva (1999):

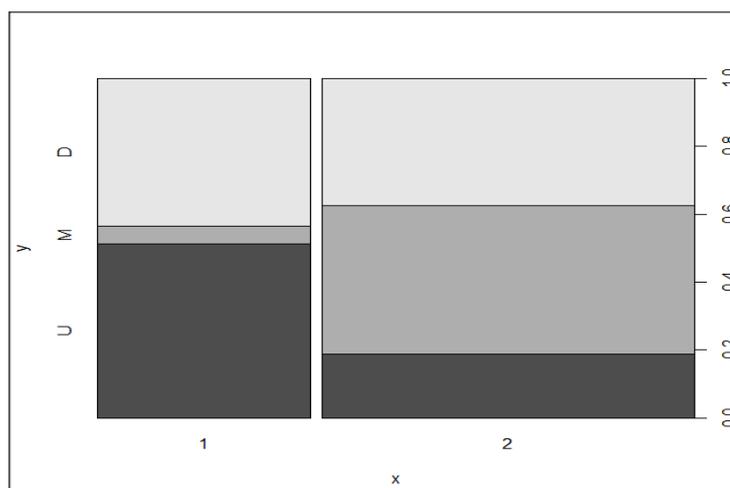
“De modo geral, o paralelismo configura-se como um fenômeno discursivo assinalado pela presença de repetições nos níveis fonológico, morfológico, lexical, sintático e semântico. Mas a literatura linguística associa-o, de modo específico, com a ideia de repetição de estruturas” (Silva 1999, p. 69).

A partir de Oliveira (2009, p. 7), temos que “o paralelismo formal indica se o antecedente mais próximo influencia de alguma forma a escolha de uma das variantes”. Ainda nesse sentido, segundo Scherre (1998, p.30), “A própria repetição das variantes de uma mesma variável dependente no discurso tem se evidenciado como uma restrição importante na análise de fenômenos variáveis de todos os subsistemas linguísticos em diversas línguas”. A partir dessas postulações, buscamos averiguar de que forma o antecedente de segunda

pessoa influência ou não no processo de variação entre o preenchimento pronominal e o sujeito nulo na amostra em estudo.

Em nossa análise, selecionamos os fatores manutenção das formas (M), ocorrência de forma única/isolada (U) e formas distintas (D). Abaixo, vejamos como as ocorrências das variantes foram distribuídas em nosso *corpus*:

GRÁFICO 6: DISTRIBUIÇÃO DAS VARIANTES POR PARALELISMO



Fonte: Software R

Como pode ser observado, percebemos distinções em relação a quantidade de ocorrências entre as variantes em cada um dos fatores em análise. Tal achado nos indica que no âmbito das realizações do fator (M), é prodominante o sujeito nulo e, por outro lado, nas observadas pelo fator (U), a preferência é pela variante preenchimento pronominal. Já o fator (D) revela valores menos discrepantes de cada uma das variantes.

A título de exemplo, observemos as imagens a seguir com alguns dados extraídos da amostra que exemplificam tais contextos de realização. A primeira imagem apresenta ocorrência de formas não paralelas, tendo no mesmo plano discursivo tanto o preenchimento do sujeito através do “você”, quanto à realização do sujeito nulo. Vejamos:

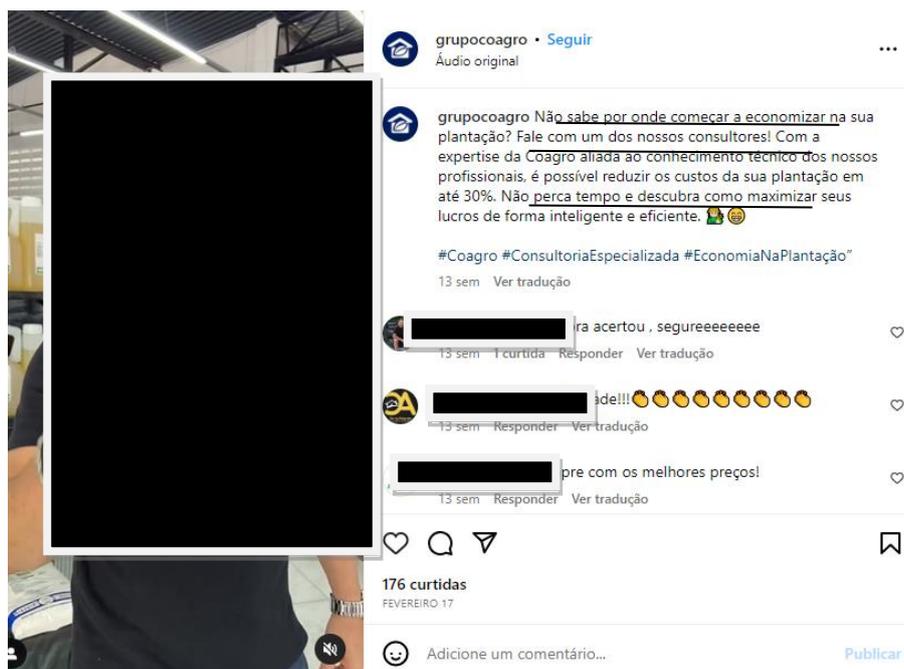
IMAGEM X: FORMAS NÃO PARALELAS NA VARIÁVEL PARALELISMO



Fonte: Instagram

A tendência do paralelismo ajuda a eliminar conflitos interpretativos tanto no plano da oração quanto no do discurso no sentido de reafirmar a sua significação. Abaixo, acompanhemos um exemplo de contexto de realização de formas paralelas:

IMAGEM X: FORMAS PARALELAS NA VARIÁVEL PARALELISMO



Fonte: Instagram

Agora, vejamos um exemplo de contexto de realização única no plano discursivo, através da imagem a seguir:

IMAGEM X: FORMA ÚNICA/ISOLADA NA VARIÁVEL PARALELISMO



Fonte: Instagram

Na imagem, verificamos que em contexto de uso isolado, o sujeito nulo foi utilizado, mas em diversos outros dados, encontramos o preenchimento do sujeito através do “você”. Dito isto e observando os exemplos acima, percebemos que os três fatores selecionados para o tratamento da variável

paralelismos são produtivos no campo da comunicação digital, de forma semelhante ao que ocorre com outros fenômenos variacionistas e em outras modalidades da língua.

O recurso do paralelismo nas línguas parece revelar uma inclinação mecânica para manutenção de estruturas paralelas que recebem influência de questões de estilo e descartando a existência de uma predisposição cognitiva determinada para se continuar falando do mesmo assunto, conforme as ideias de Weiner e Labov (1983). Por outro lado, em nossa amostra, em termos gerais, na distribuição dos dados em relação a cada um dos fatores analisados, temos a predominância de formas distintas com 39.6% dos dados totais, seguida pela realização de realização da segunda pessoa de forma única/isolada, com 30.6%, e 29.8% de manutenção do tipo de realização. Com o intuito de visualizar esses percentuais e observar a incidência de cada uma das variantes nos fatores em análise, observemos a tabela 7:

TABELA 6: PERCENTUAIS DE REALIZAÇÃO DAS VARIANTES NA VARIÁVEL PARALELISMO

Fatores	Variante 1 Você	Variante 2 Sujeito Nulo	Row Total
NÃO PARALELAS	77 39.9%	116 60.1%	193 39.6%
PARALELAS	9 6.2%	136 93.8%	145 29.8%
ISOLADAS	91 61.1%	58 38.9%	149 30.6%
Total	177	310	487

Fonte: Elaborado pela autora a partir dos resultados do software R

Ao avaliarmos os percentuais de ocorrências, percebemos que, em relação aos contextos de formas não paralelas e paralelas, a variante sujeito nulo é predominante, apresentando, na segunda, percentuais bem superiores em detrimento do preenchimento pronominal. Por outro lado, em planos discursivos com uma única realização do sujeito, a preferência é da variante “você”.

Conforme aponta Santos (2014), corroborando, através dos seus resultados, com o trabalho de Silva (2007), “ é relativamente pequena a

frequência em que esses pronomes preenchidos são usados repetidamente e, diferentemente da orientação da gramática tradicional, não são casos de destacamento do sujeito e muitos menos de tentativas de evitar a ambigüidade”. Nossos resultados apontam para o mesmo norte e indicam a preferência do sujeito nulo nesses contextos. Sobre isso, buscando explicar essa possibilidade, Scherre (1998, p.30) destaca que “repetem-se também variantes zero – menos codificação explícita leva a menos codificação explícita”.

Nos casos de uso isolado, a preferência pelo pronome explícito não traria “desgaste” para a forma, pois seu uso é pontual e, portanto, não repetido ao longo do plano discursivo. O gênero e a modalidade de comunicação aqui analisada podem servir como aspectos decisivos para essa escolha, de forma a referendar o chamamento direto do interlocutor para a informação veiculada.

De toda forma, conforme observado nas demais análises por variável, o sujeito nulo predominou no quantitativo geral de nossa amostra. Na próxima subseção, avaliaremos a significância de cada uma das variáveis, trazendo mais recursos para descrição e explicação do nosso fenômeno.

5.3. Avaliando os resultados

Nas subseções anteriores, apresentamos os dados através de seus percentuais de realização, buscando demonstrar a forma como cada uma das variáveis distingue o uso entre o preenchimento pronominal de segunda pessoa do singular e o sujeito nulo, averiguando cada um dos fatores arrolados.

Nesta, recorreremos aos valores de significância de cada variável, apontando quais foram ou não selecionadas pelo software R como significantes.. Como mencionados na seção metodológica desta tese, quanto menor for o p-valor de uma variável, maior será sua significância e, conseqüentemente, quanto maior for o p-valor, menor será a significância de uma variável.

No âmbito dos estudos da Sociolinguística, e em outros estudos das mais diversas áreas, o nível de significância que pode ser considerado significativo é o que possui resultado inferior à 0.05, tendo em vista que quanto mais próximo de 0 esse número estiver, conseqüentemente, teremos uma maior distinção

entre os modelos analisados e a própria significância do modelo, através da complexidade na qual seus fatores estão envolvidos.

Após a rodada dos dados na análise de variância, que observa o coeficiente de correlação interclasse, chegamos aos valores que excluíram as variáveis modo verbal, tempo verbal e tipo de sentença, por apresentarem resultados superiores a 0.05. Elas e suas significâncias são demonstradas na tabela 7:

TABELA 7: SIGNIFICÂNCIA DAS VARIÁVEIS EXCLUÍDAS

Variáveis excluídas	Significâncias
Modo verbal	0.86982
Tempo verbal	0.8239
Tipo de sentença	0.6362

Fonte: Elaborada pela autora a partir dos resultados do Software R

Como visto, três das quatro variáveis analisadas não apresentaram valores de significância relevantes estatisticamente conforme modelo de análise feito no R, contudo, permaneceram em nosso estudo em virtude da necessidade de observar o papel de cada uma delas na tentativa de encontrar pistas sobre “um início ou fim de uma mudança linguística, ou a especialização de significância referencial, de funções discursivas ou de funções pragmáticas” (Scherre; Naro, 2010, p.153) acerca do processo em variação aqui investigado.

Confirmando o que foi pontuado pelos autores, apesar dos valores de significância, os valores percentuais trazem resultados pertinentes, principalmente em se tratando da modalidade de comunicação digital que ainda é tão pouco explorada enquanto mais uma real e produtiva área de investigação linguística.

Em relação ao resultado da significância da variável paralelismo, temos o seguinte:

TABELA 8: SIGNIFICÂNCIA DA VARIÁVEL PARALELISMO

Variável significativa	Significância
Paralelismo	0.02

Fonte: Elaborada pela autora a partir dos resultados do Software R

O resultado da significância da variável paralelismo se deu em virtude do peso relativo, tendo o fator “manutenção das formas” como condicionador da variação com peso relativo de (0.90), gerando o valor de (0.02) como a significância dessa variável, conforme apresentado na tabela a seguir:

TABELA 8: SIGNIFICÂNCIA DA VARIÁVEL PARALELISMO

	Total	%	PR	p-valor
D	193	60.1	0.30	0.423
M	145	93.8	0.90	0.213
U	149	38.9	0.21	0.235
Total	487			

Fonte: Elaborada pela autora a partir dos resultados do Software R

A partir desses valores, temos que a variável paralelismo atua de forma expressiva no processo de variação entre o preenchimento do sujeito pronominal e do sujeito nulo de segunda pessoa do singular no âmbito da comunicação digital, através da amostra analisada.

Em síntese, através do nosso *corpus*, comprovamos nossa hipótese de que há variação entre o preenchimento do sujeito por elemento pronominal e sujeito nulo na comunicação digital realizada por empresas alagoanas na rede social Instagram. Tal resultado corrobora com os já apontados em outros trabalhos que focaram na análise do português através da fala ou da escrita convencional.

Contudo, sobre os elementos pronominais que preenchem a segunda pessoa do singular, nossos resultados foram categóricos em relação ao uso da forma “você”, não sendo encontrado nenhum outro elemento pronominal com essa referência em nossa amostra. Esse achado revela que, embora os diversos estudos apresentem variação entre as formas pronominais de segunda pessoa no português brasileiro, como visto através das pesquisas descritas na seção 2.4. desta tese, na modalidade de comunicação digital, essa variação pronominal não foi comprovada.

A exclusividade da forma “você” nos dados caminha na direção de estudos (Scherre, 1995, Scherre, 2021) que afirmam que, no português

brasileiro, o uso preferencial da segunda pessoa pronominal é realizado pelo “você”, em virtude de fatores históricos que incorporaram esse novo elemento ao paradigma dos pronomes pessoais no português brasileiro. Destacamos que os dados analisados correspondem às postagens realizadas por empresas naturais do estado de Alagoas e, nesse sentido, a incorporação de características linguísticas relacionadas a essa comunidade de fala também foram verificadas em relação ao uso massivo do “você”, em detrimento de outros elementos como “tu” e “ocê”, por exemplo.

Os trabalhos que observaram o preenchimento pronominal com referência de segunda pessoa do singular em Alagoas apontaram que, em amostras de fala da capital alagoana, obtidas através de entrevistas, 98% das ocorrências foram com a forma “você” e apenas 2% com a forma “tu” (Vitório, 2018); No Agreste alagoano, os dados também revelam a predominância da forma “você” com 89% em detrimento de 11% de “tu” (Silva, 2019); No sertão alagoano, no conjunto de dados oriundos de 96 entrevistas, apenas três ocorrências de “tu” foram encontradas, tendo todo o resto da amostra com apenas ocorrências de “você” e “cê” (Silva e Vitório, 2019). Com isso, temos que na modalidade de comunicação digital, a predominância da forma “você”, enquanto pronome de segunda pessoa do singular, revela similaridades com os demais dados arrolados em pesquisas sobre o português de Alagoas, nos quais o “você” ganha espaço e assume a referência de segunda pessoa do singular quase que de forma exclusiva, indicando um processo que caminha para uma mudança linguística.

Especificamente em relação à variante sujeito nulo, nossos resultados apontaram para sua maior distribuição no total da amostra, tendo 63,65% das realizações. Tal constatação se dá como um novo achado para os estudos sobre a temática, tendo em vista que a modalidade da comunicação digital possui especificidades que trazem ao relevo a intersecção entre as propriedades linguísticas e as específicas do gênero e do meio de veiculação da língua. Em especial, como as empresas dirigem-se ao cliente/consumidor, a referência à segunda pessoa fica explícita, pois nesse contexto o interlocutor é exclusivo, sendo o foco da comunicação, não cabendo, portanto, outra interpretação de pessoa, independentemente de marcas morfológicas verbais de 3ª pessoa figurarem nas ocorrências de sujeito nulo, por exemplo.

A maior incidência de sujeito nulo na comunicação digital demonstra que essa “nova” modalidade de uso da língua traz à tona um recurso que já não mais é tão observado em dados de fala ou de escrita convencional, mas que é totalmente licenciado pelo ambiente de uso e interação. Ao que parece, o sujeito nulo satisfaz a necessidade de referência de segunda pessoa do singular e, os casos do preenchimento pronominal podem ser encarados como uma estratégia de desfazer ambiguidades ou dar ênfase ao sujeito.

Conforme apresentamos, as variáveis selecionadas (tipo de sentença, tempo verbal, modo verbal e paralelismo) trouxeram uma interessante perspectiva de discussão dos dados. Com a grande incidência de sujeito nulo, também foi revelado que alguns fatores internos às variáveis, em termos percentuais, são mais relevantes para seu uso ou pela possibilidade de sujeito expreso pronominalmente através do “você”, como apresentado nas subseções anteriores. Em destaque, levantada como a variável de maior significância, através dos valores apontados no software R, a variável paralelismo mostrou-se diretamente relacionada com as possibilidades de ocorrência de uma ou outra forma, no contexto de uso de formas paralelas.

Nesse sentido, faz-se necessário destacar que a “nova” forma de comunicação traz consigo estruturas “inovadoras” e, conseqüentemente, novas estratégias de produção linguística, novos usos para os recursos fornecidos pela língua, não apenas os já observados na fala e na escrita convencional. De toda forma, ainda preserva o caráter heterogêneo, com processos de variação nos mais distintos níveis de análise, que refletem características estruturais desse tipo de comunicação e que merecem, cada vez mais, serem investigados.

6. CONSIDERAÇÕES FINAIS

Nossa tese propôs investigar a realização da segunda pessoa do singular na posição de sujeito na comunicação digital realizada na rede social Instagram, através de postagens em perfis públicos e oficiais de quatro empresas alagoanas, a partir da hipótese de que há um processo de variação entre o preenchimento do sujeito pronominal e o sujeito nulo.

Partindo dos resultados já apontados em trabalhos anteriores, que analisam esse processo de variação na fala e na escrita do português brasileiro, buscamos averiguar de que forma ele se realiza no âmbito da modalidade de comunicação no meio digital.

Recorremos à Teoria da Variação, proposta por William Labov (1972 e seguintes) e investigamos a atuação de um conjunto de variáveis linguísticas que poderiam favorecer o uso de uma ou outra variante em estudo, a saber: tipo de sentença (interrogativa e não interrogativa), tempo verbal (passado e não passado); modo verbal (indicativo, subjuntivo, imperativo e infinitivo pessoa); paralelismo (formas isoladas, formas paralelas e formas não paralelas). A amostra analisada compreende 487 sentenças, capturadas, na rede social Instagram das empresas, no período de fevereiro de 2022 a maio de 2023. Após coleta dos dados, o tratamento em termos percentuais e estatísticos foi feito através do software R.

Após a coleta dos dados, rodamos os dados no software R e o primeiro resultado apontado foi o de que a variante sujeito nulo predomina em termos percentuais de ocorrências e que todo preenchimento pronominal é realizado de forma categórica pelo pronome “você”. Dessa forma, comprovamos um processo de variação entre o preenchimento do sujeito e o sujeito nulo na comunicação digital das empresas alagoanas.

Observando a variável tipo de sentença, tivemos o maior número de sentenças não interrogativas. O fator “sentença interrogativa” manifestou o favorecimento do sujeito nulo, com quase 40% a mais de ocorrências, em detrimento do sujeito pronominal. Já as sentenças não interrogativas manifestaram um equilíbrio em termos percentuais entre as duas possibilidades de realização da segunda pessoa do singular na posição de sujeito. De toda forma, essa variável registra o maior percentual de sujeito nulo em ambos os

fatores, comprovando que independentemente do tipo de sentença, a variação é comprovada.

Ao analisarmos a variável tempo verbal, verificamos o maior número das ocorrências do fator “não passado”, tal resultado confirmou nossa hipótese, tendo em vista que esse fator compreende um conjunto maior de tempos verbais, bem como assinala para o contexto da comunicação digital das empresas que buscam manter a interação com o interlocutor, no âmbito do presente, apresentando produtos e serviços; bem como no futuro, criando desejos de consumo para que seu cliente seja alcançado. Dentro desse conjunto de ocorrências do fator “não passado”, a forma pronominal “você” foi favorecida, tendo 66.7% do total das realizações. O fator “passado”, de modo inverso, condicionou o maior uso do preenchimento do sujeito de forma explícita.

A terceira variável que analisamos foi o modo verbal, dividindo-o em quatro fatores correspondentes aos quatro modos verbais do português. O modo indicativo despontou como o que compreende o maior número de ocorrências na amostra. Ao averiguar cada um dos fatores, concluímos que os modos imperativo, subjuntivo e infinitivo favoreceram o uso do sujeito nulo, enquanto que o indicativo condicionou o maior uso da forma “você” na posição de sujeito. O modo imperativo, em especial, apresentou quase 97% do total de uso do sujeito nulo, mesmo intercalando desinências verbais de 2ª e 3ª pessoa, e isso pode ser explicado pela própria natureza de modo verbal que se propõe a designar ações ou fazer pedidos diretamente ao interlocutor- cliente, no caso de nossa amostra- não tendo a necessidade de colocação expressa do pronome de segunda pessoa.

Já a variável paralelismo, atuou muito fortemente, tendo para além de resultados percentuais, valor de significância de 0.02, dado pelo software R. Nesse contexto de análise, foram selecionados os fatores “manutenção das formas/ paralelas”, “formas distintas/não paralelas” e realização isolada/única. No fator que observou a realização de forma isolada, tivemos o uso majoritário do preenchimento pronominal, já nos demais fatores, foi o sujeito nulo que foi favorecido. Em especial o peso relativo (0,90) demonstrado na análise das formas paralelas foi definidor para o valor tão importante em termos de significância dessa variável.

Sobre a significância, embora as três primeiras variáveis não tenham apresentado um valor significativo nesse aspecto, revelou dados percentuais relevantes para a descrição e compreensão do fenômeno em análise, não sendo, portanto, descartados, mas, sim, explorados ao longo da análise como indícios de um processo de plena variação na comunicação digital motivado não apenas por características específicas do português, mas de propriedades inerentes à interação linguística realizada digitalmente e, ainda, mesclando características típicas da fala e da escrita convencional nesse “novo” ambiente de uso da língua.

Em suma, acreditamos que nossos objetivos foram alcançados e trouxemos à tona um contexto linguístico pouco estudado até o momento, mas vasto e profícuo campo de análise. Esperamos ter contribuído para uma maior compreensão acerca dos fenômenos em variação da segunda pessoa, bem como do universo da comunicação digital. Por fim, em pesquisas futuras, pretendemos continuar desenvolvendo investigações que correlacionem os usos pronominais e o meio digital, ampliando o horizonte de análise e a discussão acerca do tema.

REFERÊNCIAS

- ALMEIDA, N. M. de. **Gramática normativa da língua portuguesa**. São Paulo, Saraiva, 1989.
- ANDROUTSOPOULOS, J. Language change and digital media: a review of conceptions and evidence. In: COUPLAND, N.; KRISTIANSEN, T. (ed.). **Standard languages and language standards in a changing Europe**. Oslo: Novus Press, 2011. p. 145-159.
- ALKIMIM, T. Sociolingüística. In: MUSSALIM, F.; BENTES, A. C.(orgs). **Introdução à Linguística: domínios e fronteiras**. 3.ed. SP:Cortez, 2003.
- ANDRADE, C. Q. **Tu e mais quantos? – a segunda pessoa na fala brasiliense**. 139f. Dissertação (Mestrado em Linguística), Departamento de Linguística, Universidade de Brasília, Brasília, 2010.
- BAGNO, M. **Gramática Pedagógica do Português Brasileiro**. Parábola editorial, São Paulo, 2012.
- BARON, N. Enunciados segmentados em MIs. Tradução de Tania G. Shepherd. In: SHEPHERD, T. G.; SALIÉS, T.G. (org.). **Linguística da internet**. São Paulo: Contexto, 2013.p. 125-155.
- BAYLEY, R. The quantitative paradigm. In: CHAMBERS, J.K.; TRUDGILL, P.; SCHILLING-ESTES, N. (eds.), **The Handbook of Language Variation and Change**, p. 117-141. Malden, MA: Blackwell, 2002.
- BECHARA, Evanildo. **Moderna gramática portuguesa**. Rio de Janeiro: Lucerna, 2004.
- BECHARA, E. **Moderna gramática portuguesa**. 37 ed. Rio de Janeiro: Lucerna, 2006.
- BIDERMAN, M. T. C. **Teoria linguística**. São Paulo: Martins Fonte, 2001.
- BORTONI-RICARDO, S. M.. Um modelo para a análise sociolingüística do português do Brasil. In: BAGNO, Marcos (Org.). **Linguística da norma**. São Paulo, Loyola, 2002. p. 333-350.
- BOTASSINI, J. O. M. **A elipse do sujeito pronominal: uma análise variacionista**. São Paulo: Clara Luz, 2002.
- BRIZ, A. Cortesía verbal codificada y cortesía verbal interpretada em la conversación. In: BRAVO, D.; BRIZ, A. **Pragmática sociocultural – Estudios sobre el discurso de cortesía em español**. Barcelona: Ariel, 2004, p. 67-93 .
- CAMACHO, R. G. **Da Linguística formal à Linguística social**. São Paulo: Parábola, 2013
- CASTILHO, A. T. de. **Nova gramática do português brasileiro**. São Paulo: Contexto, 2010.
- CEGALLA, D. P. **Novíssima gramática da língua portuguesa**. 41. ed. São Paulo: Editora Nacional, 1998.

CHOMSKY, N. **Aspects of the Theory of Syntax**. Cambridge: The MIT Press, 1965.

COAN M.; FREITAG, R. **Sociolinguística variacionista: pressupostos teóricos metodológicos e propostas de ensino**. Revista Domínios de linguagem. Disponível em: Acesso em: 02 de abril de 2021. Acesso em: 02 de abril de 2021.

COELHO, I. L. *et al.* **Sociolinguística**. Florianópolis: LLV/CCE/UFSC, 2010. 172p. ISBN 978-85-61482-25-1. Disponível em: . Acesso em: 13 Junho 2019.

COSTA, L. B. da. **Variação dos pronomes “tu”/“você” nas capitais do Norte**. 2013. 94 p. Dissertação (Curso de Pós-Graduação em Linguística da Universidade Federal do Pará), Pará, 2013.

CRYSTAL, D. **A revolução da linguagem**. Rio de Janeiro: Jorge Zahar Ed., 2005. Título original: *The Language Revolution*, 2004.

CUNHA, C. F. da. & CINTRA, L. F. L. **Nova Gramática do Português Contemporâneo**. Rio de Janeiro: Nova Fronteira, 2008.

DUARTE, M. E. L. Do pronome nulo ao pronome pleno. In: ROBERTS, Ian; KATO, Mary (Org.). **Português brasileiro: uma viagem diacrônica homenagem a Fernando Tarallo**. Campinas: Ed. da UNICAMP, 1993. p. 107-128.

DUARTE, M. E. L. **A perda do Princípio "Evite Pronome" no Português Brasileiro**. 1995. Tese (Doutorado em Estudos da Linguagem) – Instituto de Estudos da Linguagem, Universidade de Campinas, Campinas, 1995.

DUARTE, M. E. L. Termos da Oração. In: Brandão, Sílvia F. & Vieira Sílvia R. (Org.). **Ensino de Gramática: descrição e uso**. SP: Contexto, 2007. PP. 185-204.

DUARTE, M. E. L.; REIS, P. R. **Revisitando o sujeito pronominal vinte anos depois**. ReVEL, vol. 16, n. 30, 2018.

DUARTE, M. E. L. A evolução na representação do sujeito pronominal em dois tempos. In: PAIVA, M. da Conceição; DUARTE, M. Eugênia L. (Orgs.). **Mudança linguística em tempo real**. Rio de Janeiro: Contra Capa: Faperj, 2003, 115-128.

FARACO, C. A. (1996). **O tratamento de você em Português: uma abordagem histórica**. Fragmenta, Curitiba, Ed. da UFPR, n. 13, p. 51-82.

FRANCESCHINI, L.T. **Variação pronominal nós/a gente e tu/você em Concórdia – SC**. Curitiba, 2010.

GUY, G. R. The quantitative analysis of linguistic variation. In: PRESTON, D. (ed.), **American Dialect Research**, p. 223-249. Amsterdam: Benjamins, 1993.

GUY, G.; ZILLES, A.. Sociolinguística Quantitativa. São Paulo: Parábola, 2007. HERRING, S. C. Computer-mediated discourse. In: SCHIFFRIN, D.; TANNEN, D.; HAMILTON, H. E. (org.). **The Handbook of Discourse Analysis**. Massachusetts: Blackwell Publishers, 2001. p. 612-634.

IORDAN, I. **Manual de linguística românica**. Editorial: Gredos, Madrid, 1972.

LABOV, W. **Sociolinguistic Patterns**. Philadelphia: University of Pennsylvania Press, 1972.

LABOV, W. **Padrões Sociolinguísticos**. Tradução de Marcos Bagno, Maria Marta Pereira Scherre, Caroline Rodrigues Cardoso. São Paulo: Parábola Editorial, 2008.

LÉVY, P. **A cibercultura**. São Paulo: Editora 34, 2000.

LÉVY, P. **Cibercultura**. 34 ed. São Paulo, 1999.

LOPES, C. R. dos S.; DUARTE, M. E. L. De “Vossa Mercê” a “Você”: análise da pronominalização de nominais em peças brasileiras e portuguesas setecentistas e oitocentistas”. In: BRANDÃO, S. F.; MOTA, M. A. (org.). **Análise contrastiva de variedades do Português: primeiros estudos**. Rio de Janeiro: In - fólio, 2004. p. 61-76.

LOREGIAN-PENKAL, L. **(Re)análise da referência de segunda pessoa na fala da região Sul**. 2005. 260 f. Tese (Doutorado em Letras. Área de concentração: Estudos Lingüísticos) Universidade Federal do Paraná, Curitiba.

LOREGIAN-PENKAL, L. **Alternância tu/você em Santa Catarina: uma abordagem variacionista**. Revista Estudo Linguísticos. v. 4, p. 362-367, 2005. Disponível em: .Acesso em: 12 ago. 2019.

LUCCA, N. N. G. **A variação tu/você na fala brasiliense**. 2005. 126 f. Dissertação (Mestrado em Lingüística) – Universidade de Brasília, Brasília, 2005.

LUCCHESI, D. Norma lingüística e realidade social. In: BAGNO, Marcos (Org.). **Lingüística da norma**. São Paulo: Loyola, 2004b.

MACHADO, A. C. M. **A implementação de "você" no quadro pronominal: as estratégias de referência ao interlocutor em peças teatrais no século XX**. Dissertação de Mestrado em Letras (Letras Vernáculas), Rio de Janeiro, 2006.

MARTINS, G. F. **A alternância Tu/Você/ Senhor no município de Tefé - Estado do Amazonas**. Dissertação (Mestrado em Linguística) – Universidade de Brasília, Brasília, 2010.

MENON, O. P. da S. **Pronome de segunda pessoa no Sul do Brasil: tu /você / o senhor em Vinhas da Ira**. Letras de Hoje. Porto Alegre, 2000, v. 35, n. 1, p.121-164.

MODESTO, A. T. T. **Formas de tratamento no português brasileiro: a alternância “tu/você” na cidade de Santos-SP**. 152f. Dissertação (Mestrado em Letras) – Universidade de São Paulo, São Paulo, 2006.

MOLLICA, M. C.; BRAGA, M. L. (Orgs.). **Introdução à Sociolingüística: o tratamento da variação**. São Paulo: Contexto, 2008.

MONTEIRO, J. L. **Pronomes pessoais: subsídios para uma gramática do português do Brasil**. Fortaleza: Edições UFC, 1994.

MONTEIRO, José L. **Para compreender Labov**. 2ª ed. Petrópolis – RJ: Vozes, 2000.

MOTA, M. A.. **A variação dos pronomes ‘tu’ e ‘você’ no português oral de São João da Ponte (MG)**. 2008. 125 p. Dissertação (Programa de Pós-

Graduação em Estudos Lingüísticos)- Faculdade de Letras, Universidade Federal de Minas Gerais, Minas Gerais, 2008.

MUSSALIM, F.; BENTES, A. C. **Introdução à linguística—domínios e fronteiras**. Editora Cortez, 2009.

NEVES, M. H. M.. **Gramática de Usos do Português**. 6.reimpr. São Paulo: Editora UNESP, 2000.

OTHERO, G. de A.; CYRINO, S.; SCHABBACH, G.; MADRID, L.; ROSITO, R. **Objeto nulo e pronome pleno na retomada anafórica em PB: uma análise em corpora escritos com características de fala**. Revista da Anpoll, Florianópolis, v. 1, n. 45, mai./ago., 2018.

PAREDES SILVA, V. L. **Cartas cariocas: a variação do sujeito na escrita informal**. Tese de Doutorado em Linguística. Rio de Janeiro: Faculdade de Letras da UFRJ, 1988.

PERINI, Mario A. **Gramática descritiva do português**. São Paulo: Ática, 1998.

PINHEIRO, A. F. C. **Das cartas aos chats: a variação do objeto direto de 3ª pessoa e a escrita informal no papel e na web**. Dissertação de Mestrado. Rio de Janeiro: Universidade Federal do Rio de Janeiro, 2021.

PRETI, D. **Sociolinguística: os níveis da fala**. São Paulo. Cia Editora Nacional. 1987.

SANTOS, R. L. de A. **Um estudo variacionista sobre o preenchimento da posição do sujeito em dados de fala**. Entheoria: Cadernos De Letras E Humanas ISSN 2446-6115, 1, 2014.

SAUSSURE, F. **Curso de linguística geral**. São Paulo: Cultrix, 2006.

SCHERRE, M. M. P. Paralelismo linguístico. Revista de Estudos da Linguagem, Belo Horizonte, v. 7, n. 2, p. 29-59, 1998.

SCHERRE, M. M. P. A norma do imperativo e o imperativo da norma: Uma reflexão sociolinguística sobre o conceito de erro. In: **Doa-se lindos filhotes de poodle: variação linguística, mídia e preconceito**. 2. ed. São Paulo : Parábola Editorial, 2005.

SCHERRE, M. M. P.; DIAS, E. P.; ANDRADE, C. Q.; MARTINS, G. F. Variação dos pronomes “tu” e “você”. In: MARTINS, M. A.; ABRAÇADO, J. (org.). **Mapeamento sociolinguístico do português brasileiro**. São Paulo: Contexto, 2015. p. 133-172.

SCHERRE, M. M. P.. et al. **Tu, Você, Cê e Ocê na variedade brasileira**. Papia: Revista Brasileira de Estudos do Contato Linguístico. São Paulo, SP, v.21,p.117-134,2011.Disponível em:. Acesso em: 12 jul. 2018.

SCHERRE, M. M. P.; ANDRADE, C. Q.; CATÃO, R. de C. Redesenhando o mapa dos pronomes tu/você/cê/ocê no português brasileiro falado. In:WITCHS, P. H.; VIEIRA-MACHADO, L. M. da C.; FURLAN, C. K. J.; NOGUEIRA, M. de O. (org.). **Conquistas e desafios dos estudos linguísticos na contemporaneidade: trabalhos do V Congresso Nacional de Estudos Linguísticos – V CONEL**. Porto Alegre: Editora Fi, 2020. p. 270-276.

SCHLOBINSKI, P. **Sprachliche und kommunikative Aspekte in deutschen Tweets**. 2012.

SEDRINS, A. P. **Sobre a classificação tradicional do sujeito da sentença**. Revista Areia, [S. l.], v. 4, n. 5, p. 247–265, 2021.

SOARES, M. E. **As formas de tratamento nas interações comunicativas: uma pesquisa sobre o português falado em Fortaleza**. 1980. 157f. Dissertação (Mestrado em Letras) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 1980.

SOUZA, Janaina Pedreira Fernandes de. (2011). **Mapeando a entrada do você no quadro pronominal: análise de cartas familiares dos séculos XIX-XX /** Rio de Janeiro: UFRJ – FL.

SQUIRES, L. **Enregistering internet language**. *Language in Society*, 2010. 39: 457–492.

SQUIRES, L. Introduction: Variation, representation, and change in English in CMC. In: SQUIRES, L. (ed.). **English in Computer-Mediated Communication: Variation, Representation, and Change**. Berlin, Boston: De Gruyter Mouton, 2016. p. 1-14.

TAGLIAMONTE, S. A.; USCHER, D.; KWOK, L.; *et al.* **So sick or so cool? The language of youth on the internet**. *Language in Society*, v. 45, p. 1-32, 2015 .

TAMANINE, A. **A alternância nós/a gente no interior de Santa Catarina**. (Dissertação de Mestrado). UFPR: Curitiba, 2002 TARALLO, F. A pesquisa Sócio-Linguística. 4ª. Edição. Ática. 1994.

TARALLO, F. (org.). **Fotografias sociolinguísticas..** Campinas: Ed. Pontes, 1989.

R Core Team R: **A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria, 2019.

VIEIRA, V. P. P. **Os efeitos da comunicação digital na dinâmica do ativismo transnacional contemporâneo: Um estudo sobre Al-QAEDA, WIKILEAKS e Primavera ÁRABE**. Trabalho de Conclusão de Curso (Bacharel em Relações Internacionais) – Universidade de Brasília, 2016.

WEINER, E. J. & LABOV, W.. Constraints on the agentless passive. In: **Journal of Linguistics**, 1983.

WEINREICH, U.; LABOV., W.; HERZOG, M. **Fundamentos Empíricos para uma Teoria da Mudança Linguística**. Tradução de Marcos Bagno. São Paulo: Parábola, 2006 [1968].