

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

CONTEXTOS LICENCIADORES DE SUJEITO NULO EM *PODCASTS*

LUIZ FILIPE OLIVEIRA DE MACEDO

PORTO ALEGRE

2023

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS
ESTUDOS DA LINGUAGEM
GRAMÁTICA E SIGNIFICAÇÃO

CONTEXTOS LICENCIADORES DE SUJEITO NULO EM *PODCASTS*

LUIZ FILIPE OLIVEIRA DE MACEDO

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Letras da Universidade Federal do Rio Grande de Sul e à banca examinadora como requisito parcial para a obtenção do título de Mestre em Estudos da Linguagem.

Orientação: Prof. Dr. Gabriel de Ávila Othero.

PORTO ALEGRE

2023

CIP - Catalogação na Publicação

Macedo, Luiz Filipe Oliveira de
Contextos licenciadores de sujeito nulo em podcasts
/ Luiz Filipe Oliveira de Macedo. -- 2023.
92 f.
Orientador: Gabriel de Ávila Othero.

Dissertação (Mestrado) -- Universidade Federal do
Rio Grande do Sul, Instituto de Letras, Programa de
Pós-Graduação em Letras, Porto Alegre, BR-RS, 2023.

1. Sujeito nulo. 2. Análise multifatorial. 3.
Português brasileiro. I. Othero, Gabriel de Ávila,
orient. II. Título.

Luiz Filipe Oliveira de Macedo

CONTEXTOS LICENCIADORES DE SUJEITO NULO EM PODCASTS

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Letras da Universidade Federal do Rio Grande de Sul e à banca examinadora como requisito parcial para a obtenção do título de Mestre em Estudos da Linguagem.

Porto Alegre, 20 de dezembro de 2023.

Resultado: Aprovado.

BANCA EXAMINADORA

Marcos Goldnadel

Universidade Federal do Rio Grande do Sul (UFRGS)

Pablo Nunes Ribeiro

Universidade Federal do Rio Grande do Sul (UFRGS)

Mônica Rigo Ayres Carrion

Instituto Federal do Rio Grande do Sul (IFRS)

AGRADECIMENTOS

Gostaria de agradecer com todo o carinho:

Ao professor Gabriel de Ávila Othero, que aceitou o desafio de me orientar durante o mestrado e que, pacientemente, me ajudou a vencer essa etapa valiosa da minha vida.

Aos professores Marcos Goldnadel, Pablo Ribeiro e à professora Mônica Ayres, pela disposição para a leitura cuidadosa e pelas contribuições ao meu trabalho.

Aos meus professores do PPG, pela participação na minha evolução como estudante e como professor.

Aos meus alunos, pela possibilidade de crescimento diário a partir da prática profissional.

Aos meus amigos da escola, que permanecem ao meu lado em todas as fases da vida.

Aos meus amigos da igreja, que se mostram presentes todos os dias, dos momentos mais difíceis até as mais belas celebrações.

Aos meus nobres colegas profissionais, seja os da área da educação, seja os da área da comunicação, por tornarem o cotidiano mais leve e prazeroso.

Aos meus companheiros do futebol amador, que também são parte fundamental das celebrações da vida.

Aos meus pais, que sonharam comigo com esse momento e que me deram todo o suporte necessário para que eu chegasse até aqui.

À minha esposa, Camila, por sempre acreditar em mim e por ser minha maior inspiração e meu maior exemplo de amor e de dedicação, além de ser suporte para que tudo esteja sempre bem.

A Deus, por se mostrar presente todos os dias mostrando amparo e carinho.

RESUMO

De uma língua caracterizada como *pro-drop*, o português brasileiro tem, com o passar dos anos, demonstrado uma mudança de comportamento preferindo o preenchimento do sujeito. Estudos sobre o tema sugerem que está havendo uma alteração progressiva no padrão de preenchimento do sujeito, mais precisamente e mais evidentemente a partir do século XIX (cf. Duarte, 1993, entre outros). Resta, então, compreender o que favorece que esses casos residuais de sujeito nulo ainda aconteçam na língua. O presente estudo investiga o fenômeno do sujeito nulo em PB a partir de duas questões: i) que fatores gramaticais favorecem a utilização de sujeito nulo pelo falante de PB?; ii) os fatores investigados em Ayres (2021) – a saber: morfologia verbal (Duarte, 1993, 1995), padrão linear *V1 (Kato, 2000), gênero semântico (Creus; Menuzzi, 2004) e conexão discursiva ótima (Paredes Silva, 2003) – são suficientes para justificar todos os dados de sujeitos nulos encontrados em um *corpus* de língua falada do PB?. Portanto, os objetivos desse trabalho são: i) analisar as ocorrências de sujeito nulo a partir dos quatro fatores acima mencionados e ii) verificar se há um fator que possa descrever os casos de sujeito nulo em PB de maneira plena ou se a explicação para esses casos só pode surgir a partir da combinação entre esses fatores. Para isso, foi criado um *corpus* ainda inédito nesse tipo de investigação, a partir da transcrição e da análise de seis episódios de *podcasts*, gênero comum e popular nos dias de hoje, sendo dois episódios de cada canal diferente. A partir de 1.738 contextos de sujeito, verificamos que 70,65% correspondem a posições preenchidas e apenas 29,35% correspondem a posições vazias. Esse resultado vai ao encontro de pesquisas que mostram que a língua vem cada vez menos licenciando contextos nulos. Para os 510 casos de sujeito nulo encontrados no *corpus*, a análise individual dos fatores apontou que o gênero semântico é o fator mais relevante para explicar os casos de sujeitos nulos em PB, chegando a 75,9% dos casos explicados. Ao combinarmos os fatores, seja em duplas, trios ou quarteto, a combinação que explica de maneira mais satisfatória os dados (96,7%) é a união entre todos os quatro fatores. Contudo, a análise de um subgrupo que desconsidera primeira e segunda pessoa devido ao envolvimento do gênero semântico especificamente, em um total de 287 ocorrências, mostra que é possível explicarmos 99,7% dos dados a partir da combinação do trio que envolve padrão linear *V1, conexão discursiva ótima e gênero semântico. Os poucos casos que não são explicados pelos fatores licenciadores nas duas análises (sobre 510 e sobre 287 ocorrências), na verdade, são explicados por fatores externos, podendo ser suprimidos graças ao contexto ativado pelo ambiente comunicativo. Constatamos, com esse estudo, que cerca de 99,7% dos dados analisados podem ser explicados por pelo menos um dos fatores citados por Ayres (2021), atingindo plenamente o objetivo da investigação em um novo *corpus* de gênero oral.

Palavras-chave: sujeito nulo; *podcast*; português brasileiro.

ABSTRACT

From a language characterized as *pro-drop*, Brazilian Portuguese has, over the years, demonstrated a change in behavior preferring to fill in the subject position. Studies on the topic suggest that there is a progressive change in the subject's filling pattern, more precisely and more evidently from the 19th century onwards (cf. Duarte, 1993, among others). It remains, then, to understand what favors these residual cases of null subjects still occurring in the language. The present study investigates the phenomenon of the null subject in BP based on two questions: i) what grammatical factors favor the use of the null subject by the BP speaker?; ii) the factors investigated in Ayres (2021) – namely: verbal morphology (Duarte, 1993, 1995), linear pattern *V1 (Kato, 2000), semantic gender (Creus; Menuzzi, 2004) and optimal discursive connection (Paredes Silva, 2003) – are they sufficient to justify all the data from null subjects found in a BP spoken language *corpus*? Therefore, the objectives of this work are: i) to analyze the occurrences of null subject based on the four factors mentioned above and ii) to verify if there is a factor that can fully describe the cases of null subject in BP or if the explanation for these cases can only arise from the combination of these factors. To this end, an unprecedented *corpus* of this type of investigation was created, based on the transcription and analysis of six *podcast* episodes, a common and popular genre today, two episodes from each different channel. From 1,738 subject contexts, we found that 70.65% correspond to filled positions and only 29.35% correspond to empty positions. This result is in line with research that shows that the language is decreasingly licensing null contexts. For the 510 cases of null subjects found in the *corpus*, the individual analysis of the factors showed that semantic gender is the most relevant factor to explain the cases of null subjects in BP, reaching 75.9% of the cases explained. When combining the factors, whether in pairs, trios or quartet, the combination that most satisfactorily explains the data (96.7%) is the union between all four factors. However, the analysis of a subgroup that disregards first and second person due to the involvement of semantic gender specifically, in a total of 287 occurrences, shows that it is possible to explain 99.7% of the data based on the combination of the trio that involves linear pattern *V1, optimal discursive connection and semantic genre. The few cases that are not explained by the licensing factors in the two analyzes (about 510 and about 287 occurrences), in fact, are explained by external factors, and can be suppressed thanks to the activated context in the communicative environment. We found, with this study, that around 99.7% of the data can be explained by at least one of the factors mentioned by Ayres (2021), fully achieving the objective of the investigation in a new oral genre *corpus*.

Key words: null subject; *podcast*; Brazilian Portuguese.

LISTA DE QUADROS E FIGURAS

Quadro 1	Universo dos sujeitos nulos em PB	29
Quadro 2	Combinações dos quatro fatores licenciadores de sujeitos nulos	32
Quadro 3	Comparação do verbo cantar em latim e em português	34
Quadro 4	Paradigma do verbo cantar em português	35
Quadro 5	Planilha de dados analisados – sujeitos nulos	46

LISTA DE TABELAS

Tabela 1	Número de sujeitos nulos por <i>podcast</i>	52
Tabela 2	Ocorrências de sujeito nulo e morfologia verbal (N=510)	54
Tabela 3	Ocorrências de sujeito nulo por morfologia verbal e pessoa do discurso (N=510)	55
Tabela 4	Ocorrências de sujeito nulo e *V1 (N=510)	56
Tabela 5	Ocorrências de sujeito nulo e gênero semântico (N=287)	57
Tabela 6	Ocorrências de sujeito nulo por gênero semântico e pessoa do discurso (N=287)	57
Tabela 7	Ocorrências de sujeito nulo e conexão discursiva ótima (N=510)	58
Tabela 8	Ocorrências de sujeito nulo por morfologia verbal e *V1 (N=510)	60
Tabela 9	Ocorrências de sujeito nulo por morfologia verbal e conexão discursiva (N=510)	61
Tabela 10	Ocorrências de sujeito nulo por evite V1 e conexão discursiva (N=510)	62
Tabela 11	Ocorrências de sujeito nulo por morfologia verbal e gênero semântico (N=510)	62
Tabela 12	Ocorrências de sujeito nulo por morfologia verbal e gênero semântico sem 1ª e 2ª pessoa (N=287)	64
Tabela 13	Ocorrências de sujeito nulo por evite V1 e gênero semântico (N=510)	64
Tabela 14	Ocorrências de sujeito nulo por evite V1 e gênero semântico sem 1ª e 2ª pessoa (N=287)	65
Tabela 15	Ocorrências de sujeito nulo por gênero semântico e conexão discursiva (N=510)	66
Tabela 16	Ocorrências de sujeito nulo por gênero semântico e conexão discursiva sem 1ª e 2ª pessoa (N=287)	67
Tabela 17	Ocorrências de sujeito nulo por morfologia verbal, conexão discursiva e *V1 (N=510)	69
Tabela 18	Ocorrências de sujeito nulo por morfologia verbal, *V1 e gênero semântico (N=510)	70
Tabela 19	Ocorrências de sujeito nulo por morfologia verbal, evite V1 e gênero	71

- semântico sem 1ª e 2ª pessoa (N=287)
- Tabela 20 Ocorrências de sujeito nulo por morfologia verbal, gênero semântico e conexão discursiva (N=510) 72
- Tabela 21 Ocorrências de sujeito nulo por morfologia verbal, gênero semântico e conexão discursiva sem 1ª e 2ª pessoa (N=287) 73
- Tabela 22 Ocorrências de sujeito nulo por *V1, conexão discursiva e gênero semântico (N=510) 74
- Tabela 23 Ocorrências de sujeito nulo por *V1, conexão discursiva e gênero semântico sem 1ª e 2ª pessoa (N=287) 75
- Tabela 24 Ocorrências de sujeito nulo por morfologia verbal, *V1, gênero semântico e conexão discursiva (N=510) 77
- Tabela 25 Ocorrências de sujeito nulo por morfologia verbal, *V1, gênero semântico e conexão discursiva sem 1ª e 2ª pessoa (N=287) 79
- Tabela 26 Atualização de ocorrências de sujeito nulo por morfologia verbal, *V1, gênero semântico e conexão discursiva sem 1ª e 2ª pessoa (N=287) 82
- Tabela 27 Nova atualização de ocorrências de sujeito nulo por morfologia verbal, *V1, gênero semântico e conexão discursiva sem 1ª e 2ª pessoa (N=287) 84

LISTA DE GRÁFICOS

Gráfico 1	Ocorrências de sujeito nulo em peças teatrais cariocas	23
Gráfico 2	Análise de peças teatrais cariocas por Duarte (1993) e Othero e Spinelli (2019)	24
Gráfico 3	Ocorrências de sujeito preenchido e nulo no <i>corpus</i> analisado	51
Gráfico 4	Ocorrências de sujeito preenchido e nulo por <i>podcast</i> analisado	53
Gráfico 5	Ocorrência de sujeito nulo (%) por variável analisada	59
Gráfico 6	Ocorrência de sujeito nulo (%) por duplas de variáveis analisadas (N=510)	67
Gráfico 7	Ocorrência de sujeito nulo (%) por duplas de variáveis analisadas sem 1ª e 2ª pessoa (N=287)	68
Gráfico 8	Ocorrência de sujeito nulo (%) por trios de variáveis analisadas (N=510)	75
Gráfico 9	Ocorrência de sujeito nulo (%) por trios de variáveis analisadas sem 1ª e 2ª pessoa (N=287)	76
Gráfico 10	Ocorrência de sujeito nulo (%) por todas as variáveis analisadas (N=510)	78
Gráfico 11	Ocorrência de sujeito nulo (%) por todas as variáveis analisadas sem 1ª e 2ª pessoa (N=287)	80

LISTA DE SIGLAS E ABREVIações

CDO	conexão discursiva ótima
ENEM	Exame Nacional do Ensino Médio
GS	gênero semântico
FM	frequência modulada
NURC	Norma Urbana Linguística Culta
SN	sintagma nominal
PB	português brasileiro
V1	verbo na posição 1
V2	verbo na posição 2

SUMÁRIO

1 INTRODUÇÃO	14
2 FUNDAMENTAÇÃO TEÓRICA	18
2.1 A MANIFESTAÇÃO DO SUJEITO E O QUADRO PRONOMINAL DO PB	18
2.2 O SUJEITO NULO EM PB	21
2.3 CONTEXTOS DE OCORRÊNCIA DE SUJEITO NULO	25
2.3.1 Revisão de Ayres (2021)	27
2.3.2 Relação entre sujeito nulo e morfologia verbal	33
2.3.3 Relação entre sujeito nulo e padrão linear V2	35
2.3.4 Relação entre sujeito nulo e gênero semântico	37
2.3.5 Relação entre sujeito nulo e conexão discursiva	39
3 METODOLOGIA	42
3.1 DESCRIÇÃO DA AMOSTRA	42
3.1.1 Gênero <i>podcast</i>	42
3.1.2 Constituição da amostra	43
3.2 VARIÁVEIS ANALISADAS	46
4 RESULTADOS E DISCUSSÃO	51
4.1 ANÁLISE INDIVIDUAL DE VARIÁVEIS	53
4.2 ANÁLISE DUPLA DE VARIÁVEIS	59
4.3 ANÁLISE TRIPLA DE VARIÁVEIS	68
4.4 ANÁLISE QUÁDRUPLA DE VARIÁVEIS	76
4.5 CASOS ESPECIAIS	80
5 CONSIDERAÇÕES FINAIS	87
6 REFERÊNCIAS	89
7 APÊNDICE	92
7.1 EXEMPLO DE <i>PODCAST</i> TRANSCRITO	92

1 INTRODUÇÃO

A possibilidade da omissão de sujeito nas línguas do mundo é um dos assuntos mais intrigantes para o sintaticista. Não causa espanto, mas gera debate. Se considerarmos especificamente o português brasileiro, estudos sobre o tema sugerem que está havendo uma alteração progressiva no padrão de preenchimento do sujeito, mais precisamente e mais evidentemente a partir do século XIX (cf. Duarte, 1993, entre outros).

Na literatura gerativa, as línguas que permitem consistentemente sujeitos nulos em posições marcadas com caso são referidas como línguas *pro-drop* (...) e as línguas que exigem que as posições de sujeito marcadas com caso sejam preenchidas com uma expressão foneticamente realizada como línguas *não pro-drop* (Kato; Martins; Nunes, 2023, p. 132).

De uma língua caracterizada como *pro-drop*, o PB tem, com o passar dos anos, demonstrado uma mudança de comportamento preferindo o preenchimento do sujeito. Então, os sujeitos nulos passam a ser apenas casos residuais, o que atesta um padrão antigo. Antes, os falantes preferiam omitir o sujeito. Atualmente, os falantes preferem expressar o sujeito a partir dos sintagmas nominais ou dos pronomes.

Se entendemos os casos de sujeito nulo no PB como residuais, resta entender o que favorece que ainda hoje esses casos apareçam. Duarte (1993) foi pioneira no estudo que busca essas explicações a partir da utilização de um *corpus* que leva em consideração a fala além da escrita. Seus estudos foram feitos a partir da observação de peças teatrais cariocas entre 1845 e 1992 e revelaram esse comportamento transitório do PB em relação ao fenômeno do sujeito nulo. Vale ressaltar que esse estudo baseado em peças teatrais cariocas teve sua continuação em Othero e Spinelli (2019), que acrescentaram uma análise posterior à investigação de Duarte (1993) e puderam constatar não só que, de fato, a língua está em um processo de aumento nas ocorrências de sujeito preenchido, como também isso acontece em todas as pessoas gramaticais. De certa forma, compreendemos que o sujeito expresso é preferencial em todas as pessoas gramaticais, permanecendo em uma faixa de 70% a 80% das ocorrências.

Existem outros estudos a respeito do fenômeno, que utilizam *corpus* essencialmente escritos, como Gravina (2008), que investiga textos de jornais. Há de se compreender que o texto essencialmente escrito, em especial o jornalístico, é mais

conservador. Esse tipo de texto, apesar de também atestar o funcionamento do fenômeno e apontar indícios de que o sujeito nulo já não é mais o preferencial, não favorece a espontaneidade, o que interfere na naturalidade da comunicação entre interlocutores. Em Macedo (2020), buscamos observar o comportamento desse fenômeno a partir da análise de textos produzidos por estudantes em preparação para os vestibulares da UFRGS e do ENEM. Novamente, apesar de existirem diferenças nas abordagens de cada banca, os dois vestibulares exigem a mesma tipologia textual e, ao produzir textos com o nível de formalidade exigida na preparação para o vestibular, o aluno perde a sua espontaneidade. Mesmo assim, o estudo apontou que os casos de sujeito nulo que eventualmente apareciam nos textos possuíam uma característica comum: a morfologia rica, isto é, a capacidade de apontamento para um único referente a partir da manifestação de uma desinência exclusiva para a pessoa gramatical.

Ayres (2021) também atestou a mudança de comportamento do PB observando o *corpus* de língua falada LínguaPOA, que é composto por entrevistas sociolinguísticas de informantes da cidade de Porto Alegre. A autora, além de confirmar a preferência pelo preenchimento do sujeito em PB, buscou entender que fatores gramaticais estariam envolvidos no licenciamento das ocorrências residuais de sujeito nulo em PB. De acordo com a autora:

de todo modo, estando a mudança em processo ou finalizada, o fato é que o PB mudou, tendo em vista que em estágios anteriores da língua a preferência era por omitir o sujeito. Mesmo que os sujeitos pronominais sejam favorecidos e os nulos não, ainda há ocorrências das formas nulas na posição/função de sujeito em PB contemporâneo; por isso, buscaremos encontrar que fatores licenciam essas ocorrências “pouco esperadas” (Ayres, 2021, p. 23).

O presente estudo, seguindo os passos de Duarte (1993, 2012) e, principalmente, de Ayres (2021), intenciona dar luz e sequência a essas investigações sobre o sujeito nulo em PB. Para isso, devemos observar o comportamento do fenômeno justamente a partir do reconhecimento de que essa já não é mais a forma preferencial na língua.

As questões norteadoras e as hipóteses que guiam a análise estão listadas abaixo.

Questão 1) Que fatores gramaticais favorecem a utilização de sujeito nulo pelo falante de PB?

Hipótese 1) Existem quatro fatores – a saber: morfologia verbal, padrão linear, gênero semântico e conexão discursiva – que explicam a totalidade dos dados de sujeito nulo em PB (cf. Ayres, 2021).

Questão 2) Os fatores de Ayres (2021) são suficientes para justificar todos os dados de sujeitos nulos encontrados em um *corpus* de língua falada do PB?

Hipótese 2) Os quatro fatores não são necessários para explicar cada unidade de ocorrência do fenômeno, mas considerá-los deve explicar de maneira satisfatória todas as oportunidades em que o nulo se torna a escolha do falante.

Dando seguimento à explicação das hipóteses, a partir de Ayres (2021), hipotetizamos que os quatro ambientes gramaticais favorecedores de sujeito nulo analisados pela autora sejam relevantes também em nossa análise:

a. morfologia verbal: quando há uma única forma de apresentar a desinência verbal, o sujeito poderá ser apagado, já que a própria desinência aponta para essa referência de maneira satisfatória;

b. padrão linear *V1 (Evite V1): quando a posição do sujeito na frase já aparece preenchida por outro elemento. Esse fator diz respeito ao fato de que o PB é uma língua de efeito V2, isto é, evita colocar o verbo na primeira posição da oração, tornando necessário o preenchimento dessa primeira posição com um elemento, seja ele um sujeito, seja ele um advérbio ou outra classe que sacie a característica Evite V1 do PB;

c. gênero semântico: referentes sem gênero semântico (sem sexo aparente) favorecem a forma nula, enquanto referentes com gênero semântico aparente favorecem a forma expressa, porque pronomes de 3ª pessoa possuem formas especificadas para gênero na retomada;

d. conexão discursiva ótima: quando o sujeito está relacionado à manutenção tópica do referente, o sujeito poderá ser apagado.

Em concordância com os resultados obtidos em Ayres (2021), hipotetizamos também que o conjunto dos quatro fatores será suficiente para explicar todas as ocorrências residuais do fenômeno do sujeito nulo em PB. É importante, ainda, entender uma diferença: não é esperado que todos os fatores expliquem ocorrência por ocorrência no *corpus*. É possível que tenhamos ocorrências que sejam licenciadas por apenas um desses fatores, da mesma forma que poderemos ter ocorrências

licenciadas por dois, três ou, até mesmo, pelos quatro fatores evidenciados por Ayres (2021). Então, o conjunto de fatores torna-se necessário para a compreensão do fenômeno em si, pois, ao final, esperamos que todos os dados sejam explicados de maneira satisfatória. Como explica Ayres (2021, p. 27),

o que dará conta de explicar os contextos residuais de sujeitos nulos em PB é o conjunto desses quatro fatores. Ainda que o conjunto desses quatro fatores não seja necessário para explicar cada ocorrência de sujeito nulo – por exemplo, pode ocorrer um sujeito nulo que seja licenciado por apenas um desses fatores, enquanto outro seja licenciado graças a dois, três ou todos os fatores. Entretanto, para explicar o conjunto de todos os dados, é preciso que levemos em consideração o conjunto desses quatro fatores (Ayres, 2021, p. 27).

Se a forma nula não é mais a preferencial, logo devem existir fatores ligados à gramática que favoreçam esses resquícios de ocorrências. Esperamos encontrar a relação entre sujeito nulo e os fatores apontados em Ayres (2021) em um *corpus* ainda inédito para que os estudos sobre esse tema tenham cada vez mais evidência e clareza para pesquisas futuras. O *corpus* para atender a essas necessidades foi formado a partir da transcrição de *podcasts*, gênero moderno que consiste em disponibilizar textos, conversas e entrevistas diversas via áudio em diversas plataformas (como rádio, televisão, internet e aplicativos de celular).

A presente dissertação de mestrado se organiza da seguinte forma: no capítulo 2, fundamentamos o tema teoricamente, apresentando estudos mais gerais sobre o fenômeno e explorando cada uma das principais hipóteses que explicam a possibilidade de ocorrência de sujeitos nulos; no capítulo 3, exploramos a metodologia empregada, descrevendo o gênero discursivo *podcast*, a constituição do *corpus* de análise e a classificação dos dados; no capítulo 4, trazemos os principais resultados que apresentam o fenômeno de modo geral e também os resultados para cada uma das variáveis relevantes para análise. Por fim, trazemos as considerações finais, com respostas às questões investigadas e perspectivas futuras de trabalho, além das referências e do apêndice.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo trata de conceitos, definições e explicações que embasam as discussões sobre o objeto de estudo da dissertação. Apresentaremos, aqui, o sistema pronominal do PB, sintetizando as mudanças pelas quais esse sistema vem passando, os contextos de ocorrência de sujeito nulo e, a seguir, as principais hipóteses para seu licenciamento, a saber: (i) morfologia verbal, (ii) padrão linear *V1, (iii) gênero semântico, (iv) conexão discursiva.

2.1 A MANIFESTAÇÃO DO SUJEITO E O QUADRO PRONOMINAL DO PB

Em português brasileiro, o sujeito pode ser preenchido com um sintagma nominal (1a) ou um pronome (1b) ou pode não ser preenchido (1c), como vemos abaixo.

- (1) a. O jogador chutou a bola em direção ao gol.
- b. Ele chutou a bola em direção ao gol.
- c. Chutou a bola em direção ao gol.

O estudo dos pronomes pessoais revela que a sua principal função está em substituir SNs, exercendo, portanto, funções equivalentes às exercidas por SNs, como a função de sujeito. Semanticamente, essa classe gramatical é responsável por apontamentos importantes para as relações de retomadas no texto, sendo elas anafóricas quando buscam elementos anteriores no texto, e catafóricas quando a referência aparecer na sequência da construção. Exatamente por esse motivo, torna-se relevante compreender o funcionamento do quadro pronominal em PB, já que essa classe é crucial para entendermos as motivações para as ocorrências de sujeitos nulos e expressos em PB.

Bechara (2009) aponta que existem duas pessoas determinadas no discurso: a primeira, que representa o emissor (eu), e a segunda, que representa o receptor (tu). A terceira pessoa é considerada indeterminada por apontar para uma referência que não participa ativamente da comunicação. Por fazerem referência justamente a essas pessoas do discurso e, também, a essa terceira pessoa, são classificados como pronomes pessoais:

- i. *eu* (singular) e *nós* (plural) apontando para a primeira pessoa;
- ii. *tu* (singular) e *vós* (plural) apontando para a segunda pessoa;
- iii. *ele, ela* (singular) e *eles, elas* (plural) apontando para a terceira pessoa.

Essa mesma relação é amplamente citada pela gramática tradicional, aparecendo como base também em Cunha e Cintra (2008) entre outras gramáticas. Ainda, vale destacar que Cunha e Cintra acrescentam a importância dos pronomes de tratamento na referência à segunda pessoa (receptor).

Entretanto, a formação do PB sempre foi marcada por diversas alterações sistemáticas durante a sua evolução. O trabalho de Monteiro (1994) prova isso colocando os pronomes como centro do debate ao apontar diferenças cruciais entre o que gramáticas tradicionais apresentam como parte do sistema e o que de fato está em uso considerando dados de fala das capitais brasileiras analisadas no Projeto NURC, a saber: Porto Alegre (RS), São Paulo (SP), Rio de Janeiro (RJ), Salvador (BA) e Recife (PE). Em seu trabalho, o autor menciona algumas grandes divergências nessa relação entre usos pronominais e registros na gramática tradicional:

- i. sobre *você* e *tu*:

Aqui, há uma alternância. Cunha & Cintra (2008) citam que, no PB, o uso de *tu* é restrito ao extremo Sul do país e a alguns pontos da região Norte, mas o pronome *você* o substituiu no restante do país como emprego para tratamento incluindo intimidade com o interlocutor. Peres (2007) aponta a entrada de *você* no português durante o século XX em competição com o pronome *tu* (apesar de ter seus primeiros registros ainda no século XIX, mas de maneira mais rara).

- ii. sobre *vós*:

Aqui, há um desaparecimento. Cunha e Cintra (2008) apontam que o pronome *vós* praticamente desapareceu da linguagem corrente tanto no português brasileiro quanto no europeu. Entretanto, ainda pode aparecer em literatura com tom arcaizante para indicar que o interlocutor tem alta posição hierárquica social.

A partir das variações morfofonológicas, passamos da forma *Vossa Mercê* (que inicialmente possuía alto valor hierárquico e que com o passar do tempo foi mudando sua posição na escala até chegar ao tratamento de camadas sociais mais baixas) para *Vossancê*, *Vós Micê*, até chegar na forma hoje conhecida *você*, como defende Cintra (1972). Essa forma compete até hoje com o pronome *tu* e substituiu no plural a forma *vós*.

Bechara (2009) expressa que o *vós* aparece em orações solenes. Do contrário, o pronome *vocês* entra como o plural válido para a segunda pessoa.

iii. sobre *nós* e *a gente*:

Aqui, há uma competição. Bechara (2009) aponta que o substantivo *gente* precedido do artigo *a* em referência a um grupo de pessoas que inclui a que fala passa a ser pronome e é empregada em contextos coloquiais tendo sua concordância sempre em terceira pessoa do singular. Segundo Lopes (2003), essa competição de *a gente* e *nós* começou ainda no século XIX. Monteiro (1994) percebeu que a forma *nós* ainda predomina no PB, com índices de uso entre 54 e 66% em relação ao uso de *a gente*. Ainda, vale ressaltar que, se for considerado o contexto comunicativo formal, o índice de uso de *nós* sobe para 82% contra 18% de uso de *a gente*, evidenciando a preferência do falante pela forma tradicional nesses contextos. Já no contexto informal, Zilles, Maya e Silva (2000) perceberam o índice de ocorrências em 80% para a forma *a gente* nos dados de fala de entrevistas sociolinguísticas da cidade de Porto Alegre, mostrando que em situações cotidianas já há uma alteração na preferência de referência à primeira pessoa do plural.

Para Monteiro, essas três alterações no quadro pronominal e as perdas de marcas morfológicas estão relacionadas. Além disso, o autor ainda cita como divergências de usos gramaticais o alto uso do pronome *se* e a baixa frequência dos pronomes de tratamento em geral. Entretanto, como essas duas últimas não têm grande influência nas alterações de marcas morfológicas flexionais, optamos por não aprofundá-las neste trabalho.

Ainda, cabe acrescentar que Wink, Finkenauer e Othero (2012) buscaram em gramáticas do PB essas inserções pronominais. Em uma análise de cinco gramáticas publicadas entre 2000 e 2010, apenas duas consideraram tais mudanças: Castilho (2010) e Perini (2010). Para Castilho (2010), há o acréscimo de *você*, *o senhor* e *a senhora* como formas cultas de segunda pessoa iguais à forma *tu*. Para Perini (2010), os pronomes pessoais do caso reto são formados por *eu*, *você (tu)*, *ele*, *ela*, *nós*, *vocês*, *eles*, *elas* e a forma reflexiva, evidenciando já o desaparecimento da forma *vós* e colocando em destaque o pronome *você* em relação ao pronome *tu*.

Tudo isso influencia na manifestação da concordância verbal em PB – hipótese que será explorada em 2.3.2 –, já que o paradigma verbal vem sofrendo simplificações em seu sistema com o passar dos anos.

Depois de exploradas as formas de preenchimento – sintagma nominal ou pronome – da posição de sujeito no PB, passemos à exploração do seu não preenchimento.

2.2 O SUJEITO NULO EM PB

O estudo gramatical tradicional aponta que a constituição da oração depende da relação entre o sintagma nominal e o núcleo verbal. Contudo, este sintagma nominal pode aparecer de maneira explícita ou implícita de acordo com as especificações das desinências desse núcleo verbal. Ainda, Cunha e Cintra (2008) apontam que, apesar de poderem ocorrer muitos sintagmas nominais na oração, apenas um desses será o sujeito e, que, em ordem direta na sentença, seu preenchimento é à esquerda do verbo. Consideremos:

- (2) a. Eu comprei uma casa nova.
- b. Nós compramos uma casa nova.
- c. Eu e minha esposa compramos uma casa nova.

Nas construções acima temos apenas sujeitos expressos em suas relações com o predicado. Em (2a), a desinência verbal aponta para um verbo no pretérito perfeito do indicativo, mas mais do que isso, para encontrarmos o sujeito, entendemos que a desinência aponta para a primeira pessoa do singular. Dessa forma, não há alternativa para o preenchimento da posição de sujeito da oração que não o pronome *eu*, que se refere justamente ao emissor. Já em (2b), a desinência verbal (-mos) aponta para uma referência de primeira pessoa do plural, que inclui o emissor e, também, outras referências que poderiam estar participando do processo. A quem o emissor faz referência quando cita *nós*? Por esse motivo, é comum que o pronome apareça – ou desapareça – de acordo com o contexto comunicativo. Para evitar dúvidas, é comum que o locutor expresse, como em (2c), exatamente quem são os participantes do cenário narrado. O cânone é apresentado a fim de demonstrar que, na verdade, o estudo de sujeitos é mais complexo do que isso. É necessário aprofundar as possibilidades de manifestação de sujeito e ambientes comunicativos para podermos entender o fenômeno do sujeito nulo.

Bechara (2009) aponta que a flexão verbal é capaz de representar o sujeito (se primeira, segunda ou terceira pessoa do singular ou do plural), mas que a expressão desse sujeito depende do próprio texto. Nas palavras do autor,

a rigor, portanto, não se trata da “elipse” do sujeito, mas do “acréscimo” de expressão que identifique ou explicita a que se refere o sujeito gramatical indicado na desinência do verbo finito ou flexionado. Em português, salvo nos casos de ênfase ou contraste, não se explicita o sujeito gramatical mediante os pronomes sujeitos de 1ª e 2ª pessoas do singular e do plural (Bechara, 2009, p. 592).

Dessa forma, fica clara a compreensão do autor de que, na verdade, o PB deveria ser uma língua que tem preferência pela omissão do sujeito, trazendo a obrigatoriedade da explicitação a partir do contexto indicado. Entretanto, estudos apontam uma mudança sistemática nessa preferência por omissão do sujeito no PB. Duarte (1993) sugere que há uma alteração progressiva no padrão de preenchimento do sujeito em PB desde o século XIX.

Basicamente, o português brasileiro está em estágio de mudança com relação ao parâmetro *pro-drop* – parâmetro que se caracteriza pela permissão da categoria vazia na posição de sujeito. Chomsky (1981) formalizou esse parâmetro dentro do modelo de Princípios e Parâmetros, que aponta que existem princípios que atuam em todas as línguas e parâmetros que estabelecem diferenças entre as línguas do mundo. Dessa forma, entende-se que existem línguas marcadas positivamente para o parâmetro do sujeito nulo, ou seja, possuem preferência pela omissão do sujeito, e existem línguas marcadas negativamente para esse parâmetro, demonstrando obrigatoriedade da explicitação do sujeito.

Kato, Martins e Nunes (2023) indicam que os sujeitos nulos em línguas *pro-drop* têm um comportamento de pronomes se os compreendermos como capazes de indicar referência, mas carecem do conteúdo de descrição que aparecem em sintagmas nominais tradicionais.

Ayres (2021) expressa:

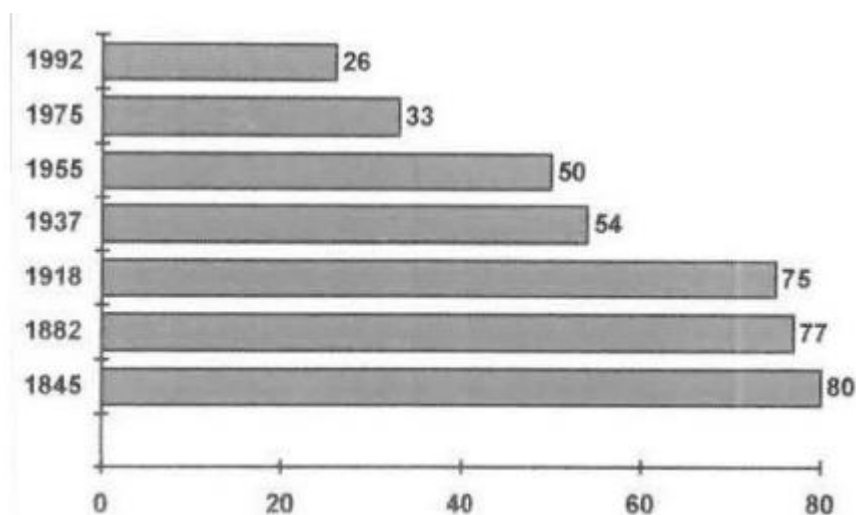
a princípio pode parecer que as línguas se enquadram sempre em um desses dois tipos, -*pro-drop* ou +*pro-drop*; porém, há línguas cujos sujeitos nulos apresentam um comportamento intermediário. Por exemplo, o funcionamento dos sujeitos em PB suscita dúvidas quanto à sua classificação com relação ao parâmetro *pro-drop*, desde Tarallo (1983), que, como vimos na introdução, mostra uma mudança em relação ao preenchimento dos sujeitos do PB. Duarte (1993), ao observar essa mudança, aponta que o PB passa por uma

fase de transição de uma língua +pro-drop para -pro-drop (Ayres, 2021, p. 29 - 30).

Deve-se compreender que, a partir de estudos recentes, a função de sujeito ocorre preferencialmente preenchida no PB atual, diferentemente de como foi apontado por Bechara (2009), que ainda considerava o padrão antigo com a preferência da omissão dessa função. Kato, Martins e Nunes (2023) reforçam que o PB permite o uso de sujeitos nulos, mas há restrições para esse uso.

Duarte (1995) apresenta um gráfico que ilustra, a partir da análise de peças teatrais cariocas em um período de 147 anos (1845 a 1992), essa mudança de comportamento.

Gráfico 1 - Ocorrências de sujeito nulo em peças teatrais cariocas



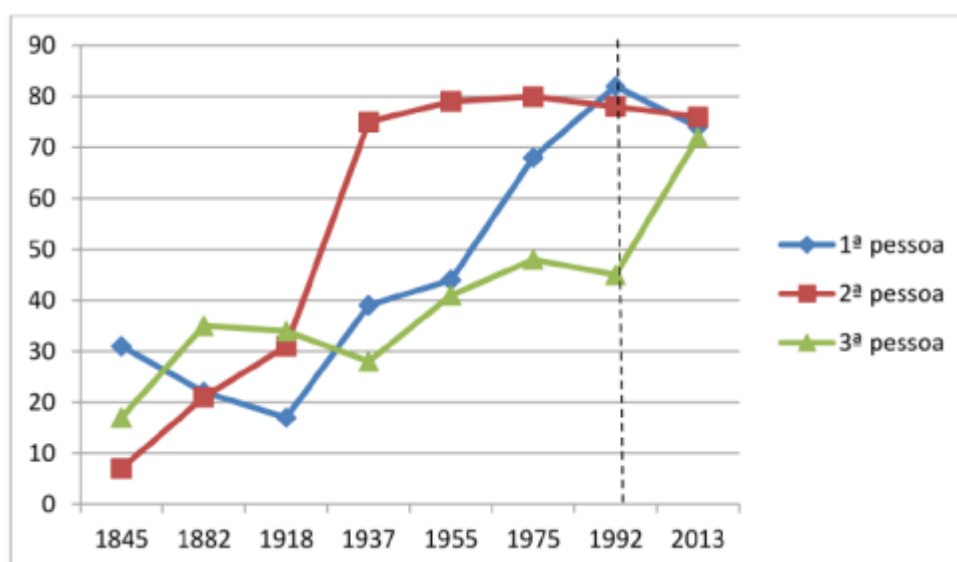
Fonte: Duarte (1995, p. 19)

Percebe-se, com a análise do gráfico 1, que as ocorrências de sujeito nulo caíram de 80% em 1845 para apenas 26% casos em 1992. Ainda, a autora apresenta uma análise para cada pessoa do discurso e aponta que a porcentagem de ocorrência de sujeitos nulos de segunda pessoa em 1918 chegava a 69%, mas caiu para 25% em 1937, coincidindo com a época em que *você* passou a vencer a concorrência com o pronome *tu*, alterando a marca flexional para a referência da segunda pessoa. Esse índice se alterou pouco até o final do período analisado. Além disso, a primeira pessoa também teve alteração significativa: em 1918 chegava a 83% dos casos de nulo, mas caiu para 18% ao final em 1992. A única pessoa que teve pouca alteração foi a terceira. Seus índices mantiveram a média de mais de 50% durante todo o período

analisado, revelando que a morfologia não é o único critério a ser observado no licenciamento de sujeitos nulos. No caso da terceira pessoa em especial, existem outros recursos que auxiliam na recuperação da informação a partir de referências externas. Paredes Silva (1991) aponta que existe uma continuidade tópica que licencia a forma não marcada do sujeito, a qual é atestada por Macedo (2020) em textos escritos e Ayres (2021) em dados de fala.

Na sequência das investigações, Othero e Spinelli (2019) acrescentaram mais duas peças teatrais cariocas à análise, sendo uma de 2011 e outra de 2013. A assimetria revelada nos estudos de Duarte (1995) – que apontava que a terceira pessoa do discurso se mantinha com níveis altos de utilização de sujeito nulo em relação aos pronomes de primeira e segunda pessoas – foi, aos poucos, se desfazendo. Na análise, Othero e Spinelli (2019) perceberam que os índices de sujeito nulo se mantiveram para a primeira e a segunda pessoa, demonstrando a clara preferência pelo sujeito expresso, mas também perceberam que, se antes a tendência era a utilização de nulo na terceira pessoa, agora já se tem a preferência pelo preenchimento. O PB, definitivamente, favorece a utilização do sujeito expresso em todas as pessoas do discurso.

Gráfico 2 – Análise de peças teatrais cariocas por Duarte (1993) e Othero e Spinelli (2019)



Fonte: Othero; Spinelli (2019, p. 16)

No gráfico acima, torna-se visível a alteração mencionada: até 1992, pelos estudos de Duarte (1993, 1995), a terceira pessoa aparecia com mais de 50% das

ocorrências de nulo, mas já em 2013 temos agora a margem de 70% de sujeitos preenchidos nessa pessoa. Dessa forma, todas as pessoas do discurso demonstram favorecer o preenchimento do sujeito se mantendo entre 70% e 80% das ocorrências.

2.3 CONTEXTOS DE OCORRÊNCIA DE SUJEITO NULO

As línguas do mundo variam em relação aos tipos de argumentos que podem ser foneticamente nulos e em relação aos contextos em que isso é permitido. Se o sujeito nulo já não é mais o padrão em PB, então torna-se relevante descobrir e discutir o que de fato licencia a utilização desse recurso anafórico.

Kato, Martins e Nunes (2023) expressam que as línguas naturais variam em relação ao tipo de sujeito nulo que licenciam. Observemos:

- (3) a. Eu comprarei uma casa nova.
b. I am going to buy a new house.

O sujeito expresso na sentença (3a), em PB, pode ser suprimido em uma versão alternativa para a mesma frase, formando “comprarei uma casa nova”. Isso não ocorre na sentença em (3b), em inglês, já que a língua inglesa é uma língua -pro-drop, não licenciando sujeitos nulos, portanto. Mesmo assim, se considerarmos alguns ambientes sintáticos, podemos ver essa construção ser licenciada ainda que o inglês não tenha marca positiva para esse parâmetro. Consideremos o exemplo retirado de Kato, Martins e Nunes (2023).

- (4) Mary began to understand the problem.

No caso em (4), seria agramatical construir a sentença como “Mary began *herself* to understand the problem”, licenciando, então, o sujeito nulo para a subordinada em um ambiente específico.

Em relação ao PB, precisamos compreender o seu comportamento em relação a esse parâmetro, já que, hoje, a língua é classificada como uma língua pro-drop

parcial (cf. Gravina; 2014, Ayres; 2021).¹ Essa classificação se deve ao fato de que, apesar de licenciar o sujeito nulo, possui mais restrições do que as línguas classificadas como pro-drop consistente. Em Holmberg, Nayudu & Sheehan (2009 *apud* Ayres, 2021), há a caracterização de línguas de sujeito nulo parcial a partir de duas grandes propriedades:

i. possuir sujeito nulo indefinido/genérico:

(5) Acorda cedo no verão².

ii. permitir sujeito nulo definido de terceira pessoa se houver uma referência clara em sentença anterior:

(6) Ram disse que comprou uma casa.

O português brasileiro, assim como o marata e o finlandês, de acordo com o trabalho dos autores, se encaixa bem nessas características. Já uma língua de sujeito nulo consistente não possui sujeito nulo indefinido e permite nulo de terceira pessoa, mesmo que não haja referência clara em sentença anterior.

(7) Camila mostrou que ela está satisfeita.

Na sentença em (7), o sujeito da segunda oração pode ser explícito ou nulo. Em caso de nulo, teríamos a frase “Camila mostrou que está satisfeita”. Em uma língua pro-drop consistente, este sujeito seria obrigatoriamente nulo; em uma língua não pro-drop, seria obrigatoriamente expresse; em uma língua pro-drop parcial, há a opcionalidade para a marcação dessa função – caso do PB.

Se o PB ainda licencia o sujeito nulo apesar da mudança de comportamento desse parâmetro, interessa, então, investigar que fatores estão envolvidos nesse licenciamento. Ayres (2021) – trabalho cujas hipóteses servirão de base para nossa

¹ A abordagem considerando o PB como pro-drop parcial não influencia a análise do fenômeno nesse trabalho de cunho descritivo.

² Os exemplos (5) e (6) foram traduzidos por Ayres (2021, p. 32) a partir dos exemplos originais de Holmberg; Nayudu; Sheehan (2009).

análise de *corpus* – apresenta uma proposta interessante que contribui para a investigação e apresenta resultados bastante promissores.

2.3.1 Revisão de Ayres (2021)

Duarte (1993) demonstrou a alteração progressiva no padrão de preenchimento do sujeito em PB desde o século XIX. Apesar de enxergar o acontecimento do fenômeno, Ayres (2021) se propôs a tentar entender o funcionamento dessa alteração a partir dos ambientes licenciadores do sujeito nulo. Se antes o PB tinha um comportamento que favorecia a ocorrência desse tipo de sujeito, hoje o PB demonstra ser uma língua com menos ambientes facilitadores para a não marcação do sujeito expresso.

Quando comparamos o PB com outras línguas românicas +pro-drop (por exemplo, português europeu, espanhol e italiano, cf. Kato, 1999), percebemos que os contextos nos quais os sujeitos nulos ocorrem são mais restritos em PB do que nas outras línguas +pro-drop. É por isso que o PB vem sendo classificado como uma língua pro-drop parcial (Ayres, 2021, p. 31).

Por isso, torna-se crucial determinar em que contextos o PB parece licenciar o uso de sujeitos nulos, visto que hoje há um comportamento diferente neste parâmetro em relação a estudos anteriores. Existem hipóteses para explicar os contextos em que o sujeito nulo se torna preferencial. Ayres (2021) desenvolve o seu trabalho com o objetivo de explicar na totalidade o que condiciona o uso de formas nulas no tratamento do sujeito no PB e apresenta uma hipótese que considera que isso aconteça por razões multifatoriais, isto é, devido à junção de diversos fatores.

Com a finalidade de investigar os ambientes e os contextos licenciadores do sujeito nulo em PB, Ayres (2021) realizou uma análise de transcrições retiradas do banco de dados de fala LínguaPOA³, que é formado por entrevistas sociolinguísticas de moradores da cidade de Porto Alegre. As entrevistas do *corpus* possuem perguntas sobre a cidade e sobre os próprios entrevistados, carregando um caráter mais subjetivo e pessoal para as respostas. A partir disso, a pesquisadora categorizou, para

³ LínguaPoa é um acervo que reúne dados de fala do português brasileiro falado em Porto Alegre entre os anos de 2015 e 2019.

cada ocorrência de sujeito (seja nulo, seja pleno), o tipo de sujeito, a pessoa gramatical e os fatores licenciadores de sujeito nulo envolvidos na ocorrência.

A investigação de Ayres (2021) assume estas quatro hipóteses como definitivas para explicar os ambientes em que ocorrem os sujeitos nulos:

i. Morfologia Verbal: com base nos trabalhos de Chomsky (1981) e Duarte (1993, 1995), essa hipótese trata da influência da desinência verbal para a expressão ou não do sujeito. Em casos de verbos de morfologia rica, ou seja, de verbos que possuam uma terminação que aponte claramente para uma única referência (por exemplo, *jogaríamos*, que tem como única possibilidade um sujeito na primeira pessoa do plural), há favorecimento de sujeito nulo. Já em casos de verbos de morfologia pobre, ou seja, de verbos que possuam uma terminação que aponta para mais de uma referência (por exemplo, *jogava*, que tem como possibilidades de sujeito tanto a primeira pessoa, quanto a terceira pessoa), o sujeito expresso torna-se preferencial (seja pronominal, seja sintagma nominal).

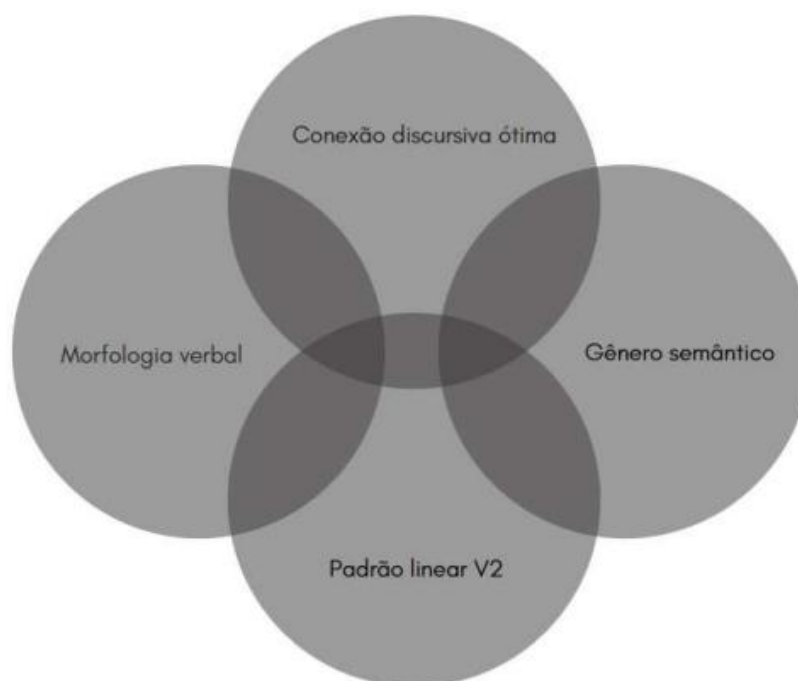
ii. Conexão Discursiva Ótima: com base nos trabalhos de Paredes Silva (1998, 2003) e de Givón (1993, 2012), essa ideia considera a manutenção das referências. Entende-se que o contexto da emissão da mensagem interfere na retomada, pois, se a referência mantém função sintática e aspectos de tempo e modo de uma oração para a outra, torna-se desnecessário marcá-lo novamente com sujeito expresso.

iii. Gênero Semântico: baseada nos trabalhos propostos inicialmente por Creus e Menuzzi (2004) e seguido posteriormente por Othero e Spinelli (2019), essa hipótese aponta que, tanto para objetos quanto para sujeitos, há uma interferência na escolha por formas nulas ou expressas a partir do reconhecimento de uma propriedade do referente em especial: a identificação de seres de sexo natural aparente, ou seja, o gênero semântico. Referentes que possuem gênero semântico (como *menina*, *cavalo*) favorecem a retomada por pronome expresso; referentes que não possuam gênero semântico (como *pessoa*, *guitarra*), favorecem a retomada por forma nula.

iv. Padrão Linear *V1: a partir da ideia inicial de Kato (2000, 2020) e Kato e Duarte (2014), trata-se da posição preferencial dos verbos na sentença, visto que o PB é uma língua que prefere a primeira posição sentencial preenchida por algum elemento que não seja um verbo. Assim, entende-se que há uma necessidade de saciar a posição à esquerda do verbo com o preenchimento de algum elemento que possa ser, de fato, o sujeito, ou ainda, alguma outra categoria a fim de evitar o padrão V1.

Ayres (2021) defende que o conjunto desses quatro fatores explica a totalidade dos casos de sujeito nulo em qualquer *corpus* do PB contemporâneo. Entretanto, é importante ressaltar que os quatro fatores não são necessários para explicar cada caso individualmente. Isso significa dizer que existem casos de sujeitos nulos explicados por apenas um dos fatores apresentados; da mesma forma, podem existir outros casos em que a combinação de dois desses fatores explique o favorecimento da forma não marcada, e assim por diante. Mesmo assim, o reforço é de que a explicação de todos os dados deve considerar a união dos quatro fatores.

Quadro 1 - Universo dos sujeitos nulos em PB



Fonte: Ayres (2021, p. 28)

Na figura acima, ilustração da hipótese de Ayres (2021), há casos explicados apenas pela hipótese da Morfologia Verbal, da Conexão Discursiva Ótima, do Gênero Semântico, ou do Padrão Linear *V1. A intersecção entre morfologia e gênero semântico não aparece na ilustração porque só há morfologia verbal exclusiva em verbos de primeira pessoa (eu, nós) e essas pessoas do discurso sempre aparecem com marcação positiva para gênero semântico. Antecipando as análises do capítulo 4 da presente dissertação, essa falta de intersecção entre os dois fatores motivou uma análise alternativa considerando apenas referentes de terceira pessoa (que sempre

apresentam morfologia pobre) para evidenciar o fator gênero semântico. Nas palavras da autora,

dessa forma, podemos encontrar sujeitos que são licenciados por qualquer um dos fatores isoladamente ou então por conjuntos formados a partir desses fatores, de modo que a soma dos fatores não é necessária para explicar cada ocorrência de sujeito nulo, mas o conjunto dos fatores é necessário para explicar todas as ocorrências de sujeitos nulos (Ayres, 2021, p. 28).

Na pesquisa, para cada ocorrência verbal do *corpus* analisado, foi possível identificar qual hipótese se encaixaria dentro das previamente apontadas por Ayres (2021). Foram encontradas 4.136 ocorrências de sujeitos anafóricos. Destas, 2.884 ocorrências eram de sujeitos expressos por pronome e 1.252 foram de sujeitos nulos. Com os números, foi possível perceber que, da quantidade total de sujeitos, apenas 30% dos dados correspondiam a sujeitos nulos. Anteriormente, o número de sujeitos nulos era mais expressivo, ilustrando a mudança de comportamento deste parâmetro em PB.

Considerando, então, as ocorrências de sujeito nulo, a autora partiu para as análises de que fatores poderiam estar contribuindo para que ainda existam casos desta forma de sujeito. As 1.252 ocorrências de sujeito nulo foram marcadas pela autora com um sinal positivo para quando se encaixassem em algum dos quatro fatores licenciadores, e com um sinal negativo quando não se encaixassem em algum dos fatores. Por exemplo, é possível que uma ocorrência tenha marca positiva para morfologia rica, mas tenha marca negativa para conexão discursiva ótima e negativa para gênero semântico. Também, é possível que uma ocorrência possua marca positiva para padrão linear *V1, para conexão discursiva ótima, mas negativa para morfologia rica, e assim sucessivamente. A partir dessas marcas para cada traço, é possível perceber que fator ou que combinação de fatores pode melhor explicar o fenômeno em sua totalidade.

Em primeira análise, encontram-se dados curiosos ao lidar com apenas um dos fatores. Em relação ao gênero semântico (em que se considera o sexo natural aparente do referente), por exemplo, esperava-se que a marcação negativa para esse traço (ou seja, um termo que não aponta para um masculino ou para um feminino de maneira clara) fosse relevante para a retomada não marcada do sujeito, mas os dados revelaram que 83,8% dos casos possuíam, na verdade, a marcação positiva para esse traço. A autora explicou que alguma inconsistência nesses dados poderia ser causada

pela escolha do *corpus*, que privilegiava discursos subjetivos e, por isso, teria mais ocorrências com gênero semântico aparente – como *menino* ou *menina*, que deixam clara na estrutura o sexo natural do referente. De qualquer forma, observando os demais traços, percebe-se que nenhum deles é capaz de explicar 100% das ocorrências de sujeito nulo em PB, com 54,3% de ocorrências de nulo considerando conexão discursiva ótima, 54,4% considerando padrão linear *V1 e 64,5% considerando morfologia rica.

















































































Na sequência, a autora iniciou a análise das combinações de dois fatores. Aqui, as combinações também não puderam explicar a totalidade dos dados de sujeito nulo, visto que, mesmo com pares sem licenciamento de nulo, ainda assim, foram encontradas ocorrências. Por exemplo, ao combinar marcas positivas para gênero semântico e negativas para conexão discursiva ótima (que considera a capacidade de referenciação a partir de uma sentença anterior), deveríamos ter a ausência de ocorrências de nulo, mas o número real apresentou 37,3% de ocorrências. O par que obteve melhores resultados foi entre padrão linear *V1 (que considera o verbo em qualquer outra posição da sentença que não seja a primeira) e morfologia rica (que trata da capacidade flexional de apontar adequadamente ao referente), apresentando 14,1% de ocorrências de sujeitos preenchidos. Em todos os pares foram encontrados dados, demonstrando que combinar dois fatores não é capaz de explicar na totalidade o fenômeno.

Ainda na busca dessa explicação para a totalidade dos casos, Ayres (2021) passou a combinar três fatores licenciadores. Nessa combinação, aparecem dados mais promissores, mas ainda insuficientes. Ao juntar as marcas negativas para conexão discursiva ótima, padrão linear *V1 e morfologia (o que não licencia o fenômeno), foram encontradas 49 ocorrências dentro das 1.252 possibilidades – um número extremamente baixo, representando apenas 3,9% dos dados, o que ainda não alcança a totalidade das ocorrências.

O que melhor explica os dados de sujeito nulo esteve na combinação dos quatro fatores. Ao investigar a junção de gênero semântico (positivo), conexão discursiva ótima (negativa), padrão linear *V1 (negativo) e morfologia (negativa), foram encontradas 13 ocorrências de sujeito nulo dentro dessa combinação que não favorece o fenômeno, atingindo apenas 1% das ocorrências totais. Isso significa dizer que a junção dos quatro fatores chegou ao expressivo número de 99% das explicações necessárias. Assim, a autora conseguiu chegar o mais próximo possível

do objetivo. Ainda, foi importante considerar, dentro dessas 13 ocorrências, quais eram os ambientes e os contextos em que apareciam. A reconsideração dos casos destoantes fez a autora perceber que, na verdade, se tratavam de conexão discursiva ótima, já que, apesar de haver uma troca de tempos verbais nos casos analisados, era possível manter o referente, o que encaixou 10 dos 13 casos na teoria. Assim, a autora chegou à tabela abaixo.

Quadro 2 – Combinações dos quatro fatores licenciadores de sujeitos nulos

Gênero Semântico	Conexão Ótima	Padrão linear V2	Morfologia	Ocorrências
+ 	+ 	+ 	+ 	 174/1.252 (13,8%)
+ 	+ 	+ 	- 	 75/1.252 (5,9%)
+ 	+ 	- 	+ 	 252/1.252 (20,1%)
+ 	+ 	- 	- 	 92/1.252 (7,3%)
+ 	- 	+ 	+ 	 241/1.252 (19,2%)
+ 	- 	+ 	- 	 72/1.252 (5,7%)
+ 	- 	- 	+ 	 141/1.252 (11,2%)
+ 	- 	- 	- 	 3/1.252 ! (0,2%)
- 	+ 	+ 	+ 	 0/1.252 (0%)
- 	+ 	+ 	- 	 51/1.252 (4%)
- 	+ 	- 	+ 	 0/1.252 (0%)
- 	+ 	- 	- 	 46/1.252 (3,6%)
- 	- 	+ 	+ 	 0/1.252 (0%)
- 	- 	+ 	- 	 69/1.252 (5,5%)
- 	- 	- 	+ 	 0/1.252 (0%)
- 	- 	- 	- 	 36/1.252 (2,8%)
				1.252

Fonte: Ayres (2021, p. 85)

Dessa forma, Ayres (2021) conseguiu explicar 99,8% dos dados, atingindo o objetivo proposto. Cabe ressaltar ainda que apenas três ocorrências não se encaixaram na hipótese esperada pela autora inicialmente. Contudo, em uma análise aprofundada desses casos, a autora concluiu que não há como se ter certeza de todos os seus traços pela composição contextual em que aparecem tornando os casos dúbios. Isso abre margem para que, em caso de dúvida, esses três casos ainda consigam ser encaixados dentro da hipótese também a partir da manutenção de referentes, chegando à explicação de 100% dos casos. Sendo 99,8%, sendo 100%, a

autora conseguiu atingir o seu objetivo de explicar o licenciamento de sujeitos nulos a partir dos quatro fatores.

No capítulo a seguir, inspirados na metodologia de Ayres (2021), utilizaremos o gênero discursivo oral *podcast* a fim de testarmos as quatro possíveis explicações para o fenômeno.

2.3.2 Relação entre sujeito nulo e morfologia verbal

Há inúmeras investigações para explicar as razões para a mudança no preenchimento da posição de sujeito. Uma explicação plausível para esse fato é a de que, conforme Duarte (1993), existe uma redução do paradigma verbal do PB, ilustrada pela inserção de pronomes novos, como *você* e *a gente*, além do apagamento de algumas desinências, como as de 2ª e 3ª pessoas em alguns contextos específicos (ex. *tu comes / tu come*). Nas análises de peças teatrais, Duarte percebeu que a utilização das formas *tu* e *você* e suas respectivas desinências verbais não entravam em competição até o início do século XX, visto que cada um desses pronomes indicava graus de proximidade diferentes em relação ao interlocutor, mantendo, assim, a possibilidade de utilização do sujeito nulo, já que não havia ambiguidade para a compreensão do sujeito.

Entretanto, nos anos posteriores, gradualmente esses pronomes se neutralizaram (passaram a ser utilizados nos mesmos contextos, entrando em competição). Ocorreu um fenômeno semelhante com os pronomes *nós* e *a gente*, que, por muito tempo, foram utilizados em contextos diferentes, mas que, aos poucos, foram se neutralizando. Essas competições, somadas ao apagamento em alguns contextos da marca desinencial do pronome *tu* (ex. *tu comes / tu come*), geraram um enfraquecimento do paradigma verbal, o que prejudicou a ocorrência do sujeito nulo, já que a falta de marcas morfológicas específicas faz com que a informação relacionada ao referente seja expressa de outra forma.

Segundo Vasconcellos (2003), uma causa importante para o enfraquecimento da morfologia verbal foi a perda do /d/ intervocálico em morfemas de segunda pessoa, ainda no português medieval. Assim, *tem-se*, por exemplo, a seguinte evolução, exemplificada pelo autor com o verbo *poder*: *podetis > podedes > podees > podes*. Associado a esse fato, por meio da reorganização dos pronomes em PB e da tendência ao menor esforço, houve uma redução dos traços distintivos dos fonemas de flexão verbal. Isso é ilustrado pelo apagamento de /s/ na fala de sujeitos de localidades onde

ainda se usa o *tu*, como em *tu pode*, em vez de *tu podes*, podendo ser ocasionado pela tendência ao menor esforço (Macedo, 2020, p. 16).

Vasconcellos (2003) mostra que há uma influência fonológica para que exista uma redução nos traços distintivos para a produção da flexão verbal. No exemplo citado em Macedo (2020), o apagamento de /s/ na flexão de segunda pessoa para o pronome *tu* é ocasionado pela tendência ao menor esforço, mas possui uma consequência clara: a neutralização da forma de concordância para *você* e para *tu*. Com isso, o autor defende que o sistema flexional do PB está diminuindo sua quantidade de formas.

Quadro 3 – Comparação do verbo cantar em latim e em português

Latim	Português
cantavi > cantai	cantei
cantavisti	cantaste
cantavit	cantou
cantavimus	cantamos
cantavistis	cantastes
cantaverunt	cantaram

Fonte: Macedo (2020, p. 17), adaptado de Vasconcellos (2003, p. 110)

No quadro acima, podemos perceber que o verbo cantar, quando conjugado no pretérito perfeito do indicativo, apresenta seis flexões verbais distintas se considerarmos o quadro pronominal apresentado nas gramáticas tradicionais com *eu*, *tu*, *ele (ela)*, *nós*, *vós*, *eles (elas)*. Entretanto, se considerarmos as alternâncias indicadas por Monteiro (1994), há um quadro diferente também presente no estudo de Vasconcellos (2003).

Quadro 4 – Paradigma do verbo cantar em português

Pessoa-Número	Forma verbal
1ª pessoa singular	canto/cantei (eu)
2ª pessoa singular	canta/cantou (você)
3ª pessoa singular	canta/cantou (ele/ela)
1ª pessoa plural	cantamos/cantamos (nós)
2ª pessoa plural	cantam/cantaram (vocês)
3ª pessoa plural	cantam/cantaram (eles/elas)

Fonte: Macedo (2020, p. 18), adaptado de Vasconcellos (2003, p. 110)

Neste quadro, podemos perceber que, de seis flexões distintas, passamos agora a ter apenas quatro. A forma *você* sendo preferência para a segunda pessoa possui a mesma flexão para as formas *ele* e *ela*, de terceira pessoa do singular; a forma *vocês* substituindo o pronome *vós* possui a mesma flexão para os pronomes *eles* e *elas*, de terceira pessoa do plural. Ainda, cabe observar que se considerarmos a preferência pela forma *a gente*, que é a preferência em situações não cerimoniais, então esta forma entraria no lugar de *nós* no quadro, retirando a desinência (-*mos*) própria para a primeira pessoa do plural e, mais uma vez, teremos a repetição da forma flexional de terceira pessoa do singular (*a gente canta / a gente cantou*).

Dessa forma, percebe-se um aumento na simplificação do paradigma flexional dos verbos em PB. Acrescentam-se as formas *você*, *vocês* e *a gente* ao quadro pronominal do PB, mas, ao mesmo tempo, por conta da competição e da consequente neutralização das formas, há uma diminuição nas desinências verbais.

2.3.3 Relação entre sujeito nulo e padrão linear *V1

Por outro lado, outra proposta passou a ser discutida como possível explicação também para o aumento das ocorrências de sujeito expresso em PB: o peso prosódico. Kato (2002) defende que o PB falado evita deixar vazia a posição pré-verbal na sentença. “Se nenhum constituinte sintático ocorrer antes do verbo, algum preenchedor discursivo ali ocorrerá [...]. A conclusão foi que o PB não é uma língua V2, mas uma língua de efeito V2, ditado pela prosódia” (Kato; Duarte, 2014, p. 13).

A tendência a evitar sentença iniciada por verbo está ligada a dois fatores, segundo Kato e Duarte (2018):

i. Duarte (1995) mostra que a presença de um elemento à esquerda do verbo (ex. negação, clíticos ou advérbios leves) favorece o apagamento do sujeito;

ii. Kato (2002), ao analisar dados de fala, mostra que o PB prefere não ter verbos no início da sentença e, para cumprir essa preferência, utiliza elementos discursivos (ex. então).

A inversão livre de posição entre verbo e sujeito no PB é restrita a verbos inacusativos, ou seja, verbos que possuem sujeito paciente (ex. Chegaram as cartas). Isso acontece porque, no PB, houve uma redução da utilização de clíticos, em detrimento de formas pronominais menos leves (ex. Comprou-lhe uma joia o Pedro [PB séc. XIX] / *Comprou pra ela uma joia o Pedro [PB séc. XX]). Essa perda da utilização dos clíticos torna o lado dos complementos mais pesado, evitando, assim, a inversão de posição. Dessa forma, percebe-se que não ocorre o padrão VS, pois, com o lado dos complementos mais pesado prosodicamente, o sujeito tende a permanecer na sua posição inicial, favorecendo, assim, o verbo em posição secundária na sentença. Entretanto, se houver algum elemento antes do verbo (não necessariamente o sintagma nominal), o VS torna-se possível.

- | | | |
|-----|--------------------------|---------------------------|
| (8) | a. Vira à esquerda | a'. Cê vira à esquerda |
| | b. Dormem ali os meninos | b'. Ali dormem os meninos |
| | c. Xinguei o cara | c'. Daí xinguei o cara |
| | d. Chove em São Paulo | d'. São Paulo chove |

O grupo de exemplos demonstra a preferência do falante de PB pelo padrão V2; o elemento que impede a ocorrência de verbo em primeira posição não precisa ser necessariamente um sintagma nominal, podendo ser um elemento leve (ex. cê, ali, daí). Para justificar a preferência dos falantes por verbo em segunda posição sentencial, Kato (2002) e Duarte e Kato (2013) introduzem a restrição “Evite V1”. Essa restrição é feita em diálogo com a restrição “Evite Pronome”, proposta por Chomsky (1981) – postulada para línguas de sujeito nulo consistente. A proposta é de que existe uma mudança na prosódia sentencial, trazendo, então, a preferência do PB pelo padrão V2. Duas situações chamam a atenção nos exemplos em (8):

i. segundo Kato (2002) e Duarte e Kato (2013), a sentença em (8b) (Dormem ali os meninos) causa estranhamento e rejeição aos falantes do PB, mas o acréscimo de um advérbio, como em (8b'), já torna a sentença aceitável;

ii. essa preferência pelo padrão V2 faz sentenças como a de (8d') surgirem com a ocorrência de um sujeito, já que, naturalmente, o verbo chover denotativamente não apresentaria um sujeito (como em 8d), justificado pela necessidade de um acréscimo de um elemento pré-verbal.

Dessa forma, percebe-se que, ao evitar deixar a posição antes do verbo vazia, o falante de PB recorrerá a recursos para a ocupação dessa posição. Assim, haverá a preferência por sentenças como em (8a') a sentenças como em (8a), na qual o verbo ocupa a posição inicial, possibilitando o aumento das ocorrências de sujeito expreso.

2.3.4 Relação entre sujeito nulo e gênero semântico

Outra proposta que passou a ser relevante no debate sobre o aumento das ocorrências de sujeito expreso no PB é a do gênero semântico. Esse traço diz respeito à distinção de referências que denotam seres de sexo natural aparente (como autor, menina, cavalo etc) e de referências que não denotam seres de sexo natural aparente (como guitarra, ponte, pessoa etc). Os referentes que são inanimados (como guitarra e ponte) não possuem gênero semântico; já os referentes animados podem possuir esse traço (como autor e menina, que possuem gênero semântico por haver a distinção entre autor/autora e menina/menino) ou não (como pessoa, que não faz distinção de sexo natural). Dessa forma, entende-se que existem substantivos animados que apontam para referentes de ambos os sexos (como gente, habitante) e, por isso, possuem gênero gramatical, mas não possuem gênero semântico. Formulada inicialmente por Creus e Menuzzi (2004), essa proposta considera que o traço de gênero semântico pode ser determinante para o aparecimento de pronomes plenos, atuando como motivador para a retomada anafórica de objetos em terceira pessoa.

Creus e Menuzzi (2004) diferenciam o gênero gramatical do gênero semântico, pois o primeiro se trata apenas de uma classificação para fins de concordância gramatical, já que, em português, todos os substantivos estão divididos nos gêneros masculino e feminino. A noção de gênero semântico, por outro lado, está atribuída a alguns substantivos em português (não a todos).

O traço de gênero semântico difere-se do traço de gênero gramatical: o primeiro é “natural”, é sobre a classificação semântica dos seres denotados por substantivos, que podem ter sexo natural identificável e reconhecível pelo falante ou não; o segundo refere-se à classificação morfossintática dos substantivos, que vai determinar suas relações de concordância gramatical (Ayres, 2016, p. 22).

Da mesma forma em que a teoria se aplica à distribuição entre objetos nulos e objetos pronominais, Othero e Spinelli (2019) apontam que o traço de gênero semântico favorece a retomada de sujeitos por pronomes plenos.

Nos casos em que o referente tem gênero semântico explícito (o marido, uma professora, o cachorro, etc.), sua retomada deve ser preferencialmente feita por um pronome (masculino ou feminino); nos casos em que o referente é marcado como [gênero semântico] (a mesa, a vítima, uma testemunha, etc.), sua retomada deve ocorrer preferencialmente por meio de um elemento vazio (\emptyset) (Othero; Spinelli, 2019, p. 8).

Basicamente, quando o referente possui gênero semântico aparente, há uma tendência de ser retomado por sujeitos pronominalmente expressos; quando o referente não possui o gênero semântico aparente, a tendência é de que seja retomado por sujeito nulo, conforme exemplos abaixo.⁴

(9) É tenso porque o pessoal não tem um transporte de qualidade. Muitas vezes, o motorista não colabora [com os passageiros]_i e acaba apressando \emptyset _i ou também desrespeitando \emptyset _i.⁵

(10) Em dezembro, [essa minha vó]_i que morava com a gente, ela_i era uma excelente cozinheira de doces. Ela fazia [amanteigados]_j pra vender e eu ajudava ela_i a pintar \emptyset _j e decorar \emptyset _j.

Nos exemplos acima, é possível perceber a influência do gênero semântico na retomada dos referentes pelos falantes. Em (9), a referência é motorista, um substantivo que, apesar de possuir gênero gramatical a partir do sistema de concordância indicando que se trata de um nome masculino (no contexto, “o”

⁴ Com a intenção de mostrar apenas a influência do fator nas retomadas, utilizamos exemplos de retomadas de objetos.

⁵ Os exemplos (9) e (10) foram retirados de Othero e Spinelli (2019, p. 10).

motorista), não possui gênero semântico aparente. Dessa forma, a retomada na oração posterior é feita por sujeito nulo nas construções “acaba apressando” e “desrespeitando”. Por outro lado, o exemplo em (10) tem a referência na construção “essa minha vó”, que, além de possuir gênero gramatical, possui também gênero semântico aparente, ocasionando em uma retomada da referência pelo pronome pleno com “ela era” e “ela fazia”. Vale também destacar que o termo é retomado por pronome pleno também em posição de objeto com “eu ajudava ela”. Ainda, vale ressaltar que o pronome pleno na posição de sujeito na construção “eu ajudava ela” pode ser creditado ao traço de gênero semântico, já que o indivíduo que está se manifestando é identificado com algum gênero semântico mesmo que não esclarecido ao receptor.

2.3.5 Relação entre sujeito nulo e conexão discursiva

Existe, ainda, uma última hipótese que coloca em evidência a estrutura informacional da sentença: a conexão discursiva ótima. Segundo Othero, Ayres e Lazzari (2018), a conexão discursiva favorece a utilização do sujeito nulo porque envolve fatores na sua configuração discursiva que dispensam a necessidade de repetição de termos. Os fatores envolvem a manutenção da função sintática do referente anterior e a continuidade do referente no mesmo plano discursivo, a partir da manutenção dos sistemas de tempo, aspecto e modo verbal.

A conexão é algo que se define com base não só em propriedades sintáticas e semânticas das orações em causa, mas também leva em conta o conhecimento pragmático e o contexto discursivo em que se inserem as orações. Além disso, [...] não é variável e absoluta, e, sim, uma questão de grau (Paredes Silva, 1991, p. 26).

Basicamente, entende-se que os estudos dos processos de retomadas anafóricas precisam considerar os contextos em que as orações estão ambientadas. Isso significa dizer que a análise deve transcender as orações em que os casos de sujeito exposto ou de sujeito nulo estão localizadas e deve perceber que essas orações são integrantes de uma unidade textual mais complexa e completa a partir do discurso do emissor. Paredes Silva (1991) aponta que em contextos em que se tem uma maior expectativa de ocorrência de pronomes do que outros deve ter sua variação relacionada a propriedades do discurso que extrapolam a leitura da oração

por si. O modo de organização do discurso passa a ser crucial para a captação de tendências sobre o uso ou não de pronomes de retomada, como vemos em (11).

(11) A falta de cuidados sanitários e de água potável estavam presentes no trabalho escravo do século XIX e ainda estão no século XXI.⁶

No exemplo acima, o verbo *estar* aparece nas duas orações do período, mas na primeira ocorrência possui sujeito pleno e na segunda ocorrência possui sujeito nulo. A segunda ocorrência dispensa a obrigatoriedade de referência plena, pois repete o apontamento da oração anterior: “a falta de cuidados sanitários e de água potável” é a referência das duas orações.

Quanto mais clara é a conexão entre o termo da oração anterior e a sua retomada na oração posterior, maior será a probabilidade dessa retomada ser feita por sujeito nulo, justamente pela falta de necessidade de repetição de termos, já que “a escolha do pronome está fortemente correlacionada à não-manutenção do mesmo referente como sujeito” (Paredes Silva, 2003, p. 104 *apud* Othero; Ayres; Lazzari, 2018, p. 30). Anteriormente, Paredes Silva (1991) já defendia que a mudança de referência é uma variável altamente condicionadora para a presença de sujeito explícito.

⁶ Exemplo retirado de Macedo (2020, p. 42).

Neste capítulo, revisitamos os principais trabalhos que embasam o tema com ênfase especial para os resultados obtidos em Ayres (2021). Foi importante iniciar trazendo um panorama sobre a compreensão a respeito do estudo dos sujeitos de uma maneira mais geral, incluindo o que é amplamente abordado nas instituições de ensino, para, depois, ampliarmos essa visão considerando o comportamento do fenômeno em si.

Logo após, aprofundamos cada uma das hipóteses presentes na literatura, quais sejam:

- (i) morfologia verbal;
- (ii) padrão linear *V1;
- (iii) gênero semântico;
- (iv) conexão discursiva ótima.

Essas hipóteses serão testadas no banco de dados formado para o presente estudo, o qual é explicado no próximo capítulo.

3 METODOLOGIA

A fim de cumprir o objetivo geral do trabalho, que é descrever os contextos de uso de sujeitos nulos em um *corpus* de fala oral espontânea, neste capítulo apresentamos a metodologia adotada para a análise. Iniciamos a exposição pela descrição da amostra, em que caracterizamos o gênero *podcast* e explicamos as etapas de constituição da amostra de dados, seguida das variáveis analisadas, as quais foram inspiradas em Ayres (2021).

3.1 DESCRIÇÃO DA AMOSTRA

Com o objetivo de descrever as características e a constituição da amostra de dados analisada, descrevemos nas subseções abaixo propriedades do gênero discursivo utilizado e etapas de constituição da amostra.

3.1.1 Gênero *podcast*

A palavra *podcast* é uma combinação das palavras *iPod* - marca de um aparelho multimídia - e *broadcast* – transmissão.

Diferentemente de uma conversa informal, face-a-face, o *podcast* é uma gravação em vídeo e/ou em áudio. Embora esteja na modalidade oral, ele apresenta a possibilidade de defasagem temporal entre o ato de elocução e o ato de leitura/escuta, o que costuma ser mais característico da modalidade escrita. Nesse sentido, esse gênero nos permite ampliar a discussão sobre as características dos gêneros orais tradicionalmente elencadas por materiais didáticos que enfocam a modalidade oral (Villarta-Neder; Ferreira, 2020, p. 39).

A partir de conceitos da linguística, do radiojornalismo e da interação do rádio com a *web*, Júnior e Padilha (2020) discutem se *podcasts* se configuram como um novo gênero ou como uma adaptação de um gênero radiofônico já existente. Segundo os autores, “a linguagem das redes digitais (*sites*, redes sociais etc.) traz um novo ambiente inteiro que se utiliza de um gênero discursivo eminentemente híbrido, hipermidiático” (p. 13).

Aplicando as noções de *crossmídia* (veiculação por mais de um meio, sem que haja alteração do conteúdo) e *transmídia* (transmissão de conteúdos complementares

em diferentes mídias) à produção de *podcasts*, os autores imaginam quatro contextos de produção, os quais são apresentados abaixo:

- (i) *podcasts* produzidos por emissoras de rádio para a grade de programação;
- (ii) *podcasts* dissociados das programações de qualquer emissora, mas com os padrões de linguagem de gêneros radiofônicos;
- (iii) *podcasts* como narrativa transmidiática em que há várias mídias necessárias para o entendimento da mensagem;
- (iv) *podcasts* com suportes multimidiáticos, em que mais áudios, textos, fotos e vídeos são complementares à narrativa, não sendo fundamentais para o entendimento do todo.

A escolha do gênero *podcast* como suporte para a constituição do *corpus* de análise é justificada pela ausência de análises sobre sujeito nulo nesse tipo de dado de fala espontânea ou semiespontânea. Além do ineditismo do *corpus*, o gênero *podcast* é, nos dias atuais, cada vez mais popularizado, comum e de fácil acesso à população geral. Há *podcasts* para todos os nichos e classes sociais, e é por esse motivo que os *podcasts* tornam-se representativos do vernáculo.

Em comparação com o gênero entrevista, comumente utilizado na construção de bancos de dados, notamos que algumas diferenças residem no fato de que entrevistas midiáticas eram realizadas apenas para meios midiáticos tradicionais e os convidados eram apenas celebridades, famosos ou personalidades relevantes para a área. Os *podcasts* permitem que, com o mínimo de conhecimento sobre o gênero ou o aporte instrumental necessário, qualquer pessoa entreviste ou seja entrevistada.

3.1.2 Constituição da amostra

Para a constituição da amostra da presente dissertação, utilizamos diferentes episódios de *podcasts* que se encaixam nos padrões (i), (ii) e (iv) da classificação acima⁷, a saber: “Não vai cair no ENEM”, “Podcast Denílson Show com Chico Garcia” e “Tá Vazando”, os quais são caracterizados a seguir.⁸

⁷ Não há seleção de *podcast* do tipo (iii), pois não temos interesse em materiais em que o áudio não seja suficiente para a transmissão da mensagem – o que poderia influenciar no número ou nos contextos de ocorrência de sujeitos nulos.

⁸ A exposição das características de cada *podcast* está ordenada pelo número de participantes em interação.

(i) *Podcast "Não vai cair no ENEM"*

O *podcast* "Não vai cair no ENEM", do canal Buenas Ideias, trata de temas relacionados à história do Brasil. Ele é comandado por Eduardo Bueno, jornalista e "historiador" com formação acadêmica, com idade na faixa de 60 anos. O programa tem formato de conversa descontraída com o interlocutor virtual – ou seja, o jornalista está sozinho, sem outros participantes simultâneos.

Por se tratar de uma aula de história, o *podcast* é roteirizado: há interseções e devaneios do condutor, mas a aula segue tópicos bem estruturados. Também por esse motivo, acaba por abranger um nicho intelectual. Cada programa tem duração média de 20 minutos.

Em relação aos tipos de *podcasts* elencados por Júnior e Padilha (2020), podemos classificá-lo como tipo (iv): *podcasts* com suportes multimidiáticos, em que mais áudios, textos, fotos e vídeos são complementares à narrativa, não sendo fundamentais para o entendimento do todo.

(ii) *Podcast "Denílson Show com Chico Garcia"*

O *podcast* "Denílson Show com Chico Garcia" aborda a temática do futebol e é comandado por dois comunicadores: um ex-jogador de futebol e um jornalista. Há sempre um terceiro participante: alguém relevante no universo do futebol que será entrevistado. Os comunicadores têm entre 40 e 50 anos, enquanto a idade dos entrevistados varia.

O formato do *podcast* é de conversa descontraída entre os comunicadores e o entrevistado; a audiência não tem participação relevante. Existe um roteiro prévio de perguntas, já que a intenção de cada episódio é traçar uma biografia do entrevistado passando por pontos específicos de sua vida. A roteirização, portanto, é no estilo de entrevista semiestruturada.

O *podcast* atende o nicho de esporte, tem duração média de 2 horas por episódio e atende às características, segundo Júnior e Padilha (2020), do tipo (ii): *podcasts* dissociados das programações de qualquer emissora, mas com os padrões de linguagem de gêneros radiofônicos, principalmente devido à presença do radialista Chico Garcia.

(iii) *Podcast “Tá Vazando”*

O “Tá Vazando” é um programa de rádio transformado em *podcast* para circulação e divulgação posterior ao acontecimento do programa e é hospedado em diversas plataformas. Trata de assuntos de cotidiano, repercute notícias e, eventualmente, recebe convidados, em especial do meio musical. É comandado por radialistas experientes e com linguagem jovial para se adequar ao público consumidor, que varia entre 15 e 50 anos de idade essencialmente.

O *podcast* tem formato de conversa descontraída entre os participantes. Os interlocutores participam com o envio de mensagens eletrônicas. Existe roteiro prévio para a divulgação das notícias, mas a condução está a critério do âncora, que foca nas reações e interações não roteirizadas entre os participantes. Além disso, eventualmente recebe convidados para a divulgação de trabalhos, de agendas de shows e para oportunizar a conversa com o ouvinte. Tem como nicho as variedades cotidianas.

Por estar ligado a uma programação de rádio FM, o programa tem duração de uma a duas horas por episódio além das inserções comerciais. Nos termos de Júnior e Padilha (2020), trata-se do tipo (i): *podcasts* produzidos por emissoras de rádio para a grade de programação.

A etapa metodológica iniciou pela seleção dos *podcasts*, com o objetivo de constituir um *corpus* com tipos e modelos variados de programas. Dentro de cada um dos três canais selecionados, os episódios foram escolhidos aleatoriamente.

Na sequência, foi estabelecido um recorte de 20 minutos entre o meio e o fim em cada episódio, a fim de manter uma padronização de tempo entre os diferentes estilos de *podcasts*. Foram selecionados 6 episódios de duração média de 20 minutos – 2 de cada proposta de *podcast*. Os materiais de áudio foram transcritos ortograficamente por meio das ferramentas do Windows (Windows+H). Logo após, foram feitas as correções e os ajustes manuais do texto transcrito.

Depois de criado o *corpus* de análise, realizamos o levantamento de todas as ocorrências de sujeito, a fim de verificarmos a proporção de posições preenchidas ou não. Em seguida, fizemos a tabulação em arquivo Excel de todas as ocorrências de sujeito nulo. Cada ocorrência foi classificada de acordo com as variáveis linguísticas investigadas, que serão apresentadas em 3.2. Após a classificação de 510

ocorrências, descrevemos estatisticamente os principais resultados. Na figura abaixo, está um exemplo de um trecho da planilha de análise.

Quadro 5 – Planilha de dados analisados – sujeitos nulos

	A	B	C	D	E	F	G	H	I	J	K
1	Oco	Podcast	Lexema	Realização	Pessoa	Número	Tempo	Morfologia Rica	Padrão *V1	Gênero Semântico	Con. Disc. Ót.
2	1	BI - O.G.	saber	sabe		2 sing.	Presente	não	não	não	sim
3	2	BI - O.G.	saber	sabe		3 sing.	Presente	não	não	não	sim
4	3	BI - O.G.	evocar	evoca		3 sing.	Presente	não	não	não	sim
5	4	BI - O.G.	fazer	faz		3 sing.	Presente	não	sim	não	sim
6	5	BI - O.G.	saber	sabe		2 sing.	Presente	não	não	não	sim
7	6	BI - O.G.	saber	sabe		2 sing.	Presente	não	não	não	sim
8	7	BI - O.G.	querer	quero		1 sing.	Presente	sim	não	sim	não
9	8	BI - O.G.	estar	estou		1 sing.	Presente	sim	não	sim	sim
10	9	BI - O.G.	ser	é		3 sing.	Presente	não	não	não	não
11	10	BI - O.G.	ser	é		3 sing.	Presente	não	não	não	não
12	11	BI - O.G.	saber	sabe		2 sing.	Presente	não	não	não	sim
13	12	BI - O.G.	falar	falei		1 sing.	Pret. Perf.	sim	sim	sim	não
14	13	BI - O.G.	ser	é		3 sing.	Presente	não	sim	não	não
15	14	BI - O.G.	ser	era		3 sing.	Pret. Imp.	não	não	não	sim
16	15	BI - O.G.	iniciar	inicia		3 sing.	Presente	não	sim	não	não
17	16	BI - O.G.	iniciar	inicia		3 sing.	Presente	não	sim	não	não

Fonte: elaboração própria

3.2 VARIÁVEIS ANALISADAS

Neste trabalho, perseguimos a hipótese de que as ocorrências de sujeito nulo sejam licenciadas por fatores linguísticos, como morfologia verbal, padrão linear, gênero semântico, e conexão discursiva. A fim de verificar o papel de cada fator na explicação da amostra, bem como outras características de cada ocorrência, classificamos as ocorrências de sujeito nulo de acordo com as variáveis listadas abaixo.

i) Lexema verbal:

Preenchemos uma coluna do Excel com o lexema, ou seja, a forma abstrata que compreende todas as formas flexionais do paradigma, para analisar se há algum verbo que apareça mais em relação aos outros, demonstrando assim uma disposição prévia para o apagamento do sujeito.

(12) “Ela era húngara, mas etnicamente *era* considerada alemã, né.”

“Então, pra mim, *foi* o melhor treinador que eu tive.”

ii) Riqueza morfológica:

Duarte (1993, 1995) aponta que o empobrecimento das desinências verbais é fator determinante para que tenhamos maior preenchimento da posição de sujeito. Assim, hipotetizamos que há mais ocorrências de sujeito nulo com verbos de morfologia rica, isto é, verbos que apresentam desinência exclusiva para uma pessoa gramatical específica, e mais preenchimento de sujeito com verbos de morfologia pobre, ou seja, verbos que possuem uma terminação que aponta para mais de uma pessoa gramatical. Na análise, os verbos foram divididos em:

a. Morfologia rica:

(13) “*Estamos* falando de carne humana!”

“*Acho* que o próximo treinador da seleção vai ser um bom nome pelos nomes que a gente tem aí.”

b. Morfologia pobre:

(14) “*Vira* um delator, *vira* aqueles informantes da polícia.”

“Durante o Quarentemo tu falou muito sobre o início lá da Fresno, do emo, até *virou* um podcast depois.”

iii) Padrão Linear:

Segundo Kato (2002), a partir da análise da prosódia, o PB falado apresenta uma preferência por não iniciar a oração com uma forma verbal, seja com sujeito posposto, seja com sujeito nulo. Diante disso, hipotetizamos que, a partir do preenchimento da primeira posição da sentença com elementos que possuem peso prosódico (como advérbios), o sujeito torna-se dispensável para o falante, aumentando, assim, os casos de sujeito nulo na construção. No experimento, os casos foram divididos em:

a. Padrão *V1:

(15) “Não *acho* que ele sairia.”

“Bá, nunca *vou* ter empresário na vida”

b. Padrão V1:

(16) “*Curto* muito.”

“*Tinha* grande interesse pelas artes e *ia* em peças teatrais.”

iv) Gênero Semântico:

A partir dos estudos de Creus e Menuzzi (2004) e Othero e Spinelli (2019a, 2019b), entende-se que, se o referente é marcado positivamente para o traço gênero semântico, há uma maior probabilidade para a retomada ser feita com um pronome pleno. Dessa forma, se a marca for negativa para o traço gênero semântico, compreende-se que pode haver uma preferência para o sujeito nulo. Por esse motivo, na tabela, as ocorrências foram divididas em:

a. Traço negativo para gênero semântico:

(17) “Eu acho que, no final, (o cargo) *vai* cair no colo do Fernando Diniz.”

“(O episódio) não *deu* tanto *view* quanto *deveria* ter dado.”

b. Traço positivo para gênero semântico:

(18) “(Ele) *tinha* aprendido alemão com a Catarina.”

“Como (ele) ainda não *decidiu*, vão esperar até uma decisão final.”

v) Conexão Discursiva Ótima:

Essa variável torna-se relevante a partir da compreensão de que, se há uma manutenção de aspectos formais de uma sentença para a outra, é possível reconhecer o sujeito a partir da repetição estrutural e, por consequência, sua expressão torna-se dispensável. Considerando o contexto em que a oração está ambientada, deve haver a preferência por manter o sujeito implícito. Vale lembrar também que a conexão discursiva ótima exige manutenção de todos os aspectos gramaticais do sujeito de uma sentença para a outra, a fim de facilitar essa identificação. Na análise, as ocorrências de nulo foram divididas em duas categorias:

a. Com conexão discursiva ótima:

(19) “A gente não estava mais falando nem de antes, quando *falava* isso anos atrás assim.”

“Ele foi o maior. *Tem* o maior título, conquista de Champions.”

b. Sem conexão discursiva ótima:

(20) “A gente tocando lá... o playback... *começaram* a encenar...”

“Pensando numa próxima Copa, tendo pouco essa ideia, *fala* assim ó... *estamos* trazendo esse treinador porque ele demonstrou até agora essa característica.”

vi) Pessoa do discurso:

A variável pessoa do discurso tem sua análise intimamente ligada à riqueza morfológica. Verbos com morfologia rica aparecem com desinências exclusivas de primeira pessoa do singular e primeira pessoa do plural. Mesmo assim, existe a possibilidade de termos a primeira pessoa do singular apresentando a marca de morfologia pobre a partir da neutralização com formas de segunda e terceira pessoas do singular e primeira pessoa do plural (ex. eu fazia / ele fazia / você fazia / a gente fazia). Ainda, é válido observar quais pessoas do discurso aparecem com mais frequência nas sentenças de sujeito nulo.

(21) “*Queria* que você tentasse fazer um balanço em relação a isso.”

“*Marcava* show no Garagem Hermética, ali nos bagulho.”

vii) Tempo gramatical:

A variação de tempos gramaticais a partir da narrativa de relatos pessoais oferecida pelo gênero *podcast* pode influenciar na forma como os verbos são realizados e, por consequência, pode também influenciar na capacidade do interlocutor de encontrar a referência de sujeito para a história. Da mesma forma, a manutenção dos tempos verbais pode facilitar o reconhecimento do sujeito. Mesmo estando diretamente ligada à conexão discursiva ótima, a CDO difere-se por exigir que todos os aspectos gramaticais sejam repetidos, não só o tempo verbal. Entretanto, há casos em que apenas a alteração do tempo verbal de uma oração para a outra pode ainda manter a conexão discursiva ótima. Assim, é importante identificar os tempos verbais mais comuns na construção do fenômeno em si.

(22) “Não se *cansava* de ouvir [...] e *partiu* para degolar Paulista, baiano, esses caras que foram lá enfrentar gaúcho, né.”

“*Estava* em um processo de crescimento [...] então mais *fui* com essa consciência e até um pouco com essa escolha mesmo.”

Vale ressaltar que as variáveis (ii), (iii), (iv) e (v) são os fatores que, de fato, podem influenciar no licenciamento do sujeito nulo. Os demais oferecem uma riqueza de compreensão do fenômeno para fins de comparação. Por exemplo, os estudos de Othero e Spinelli (2019) mostram que a alteração no padrão do preenchimento do sujeito em PB está, nesse momento, equilibrada entre as três pessoas do discurso. Todas as pessoas do discurso demonstram preferência por sujeito expreso. Os estudos anteriores de Duarte (1993) mostravam que havia uma assimetria na qual a terceira pessoa ainda mantinha cerca de 50% de ocorrências de sujeito expreso enquanto as demais já atingiam a marca de 70%. Dessa forma, assinalar as ocorrências nas pessoas do discurso, nos tempos verbais e considerando os lexemas enriquece a análise a fim de facilitar reflexões futuras sobre o comportamento das sentenças.

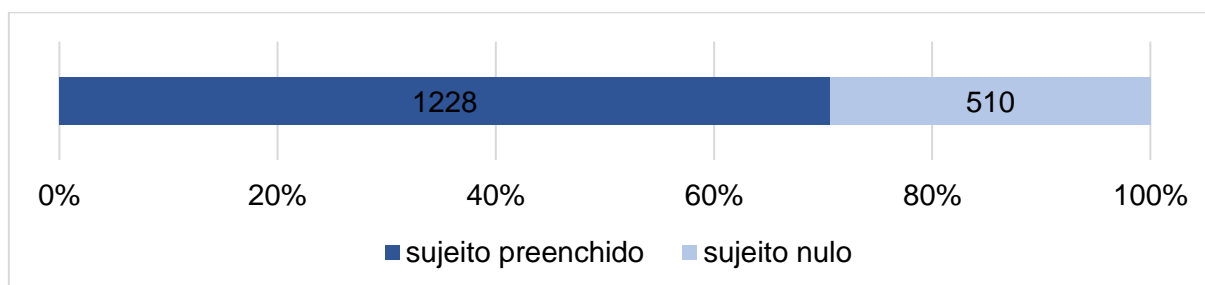
Neste capítulo, apresentamos a formação do *corpus* e as variáveis investigadas no estudo, a saber, o lexema verbal, a riqueza morfológica, o padrão linear, o gênero semântico, a conexão discursiva ótima, a pessoa do discurso e o tempo gramatical. No próximo capítulo, apresentaremos os resultados a partir da análise dos dados obtidos nesse *corpus*.

4 RESULTADOS E DISCUSSÃO

Neste capítulo apresentaremos os resultados da investigação metodológica sobre o licenciamento de sujeitos nulos em *podcasts*. Após comentarmos os resultados gerais, passaremos para a análise individual e em grupos de 2, 3 ou 4 variáveis, inspirados em Ayres (2021).

A partir da análise dos episódios de *podcasts*, contabilizamos o total de 1.738 contextos para análise. O gráfico abaixo ilustra a atual preferência do PB por sujeitos preenchidos em detrimento do antigo padrão, que apresentava preferência por sujeitos nulos.

Gráfico 3 – Ocorrências de sujeito preenchido e nulo no *corpus* analisado



Fonte: elaboração própria.

A ocorrência de sujeitos preenchidos, seja por sintagma nominal, seja por pronome, corresponde a 70,65% (1.228/1.738) dos dados, enquanto a ocorrência de sujeitos nulos corresponde a apenas 29,35% (510/1.738), o que confirma a hipótese geral sobre o português brasileiro de que a língua vem cada vez menos licenciando contextos nulos. Vale ressaltar, também, que as porcentagens são muito semelhantes às encontradas em Ayres (2021), confirmando um padrão. Por esse motivo, a presente dissertação foca nos casos de sujeito nulo, investigando quais variáveis são responsáveis por licenciar contextos desse tipo.

Na tabela a seguir, aparece a contagem total de ocorrências de sujeito nulo em cada episódio de *podcast*.

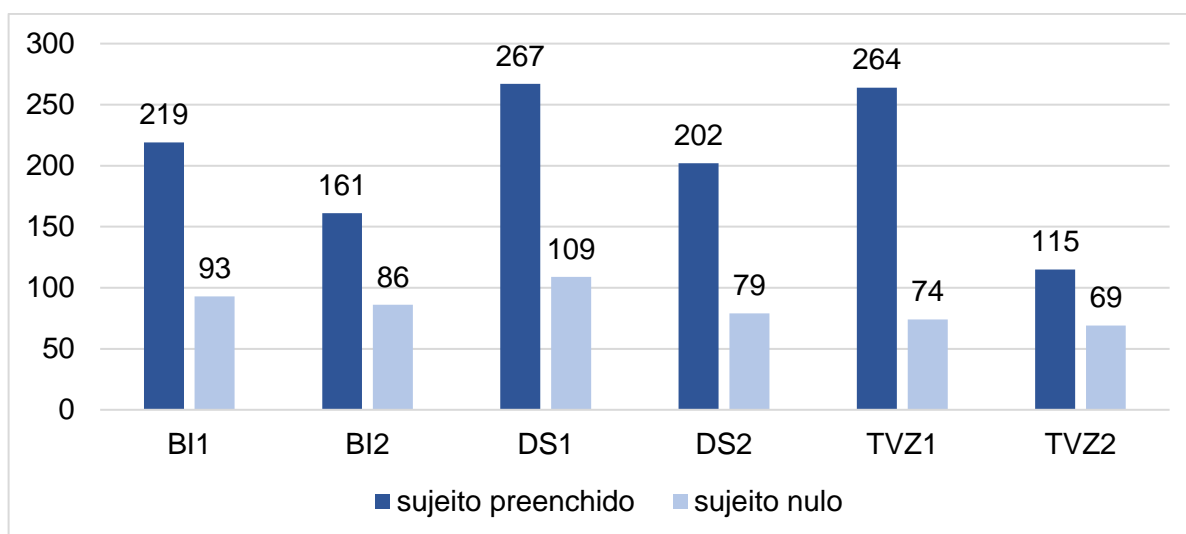
Tabela 1 - Número de sujeitos nulos por *podcast*

Podcast/Episódio	Ocorrências
Buenas Ideias – tema: Linguiceiros de POA	86
Buenas Ideias – tema: Origem do Gaúcho	93
Denílson Show – convidado: Diego Ribas	109
Denílson Show – convidado: Kaká	79
Tá Vazando – convidado: Fresno	69
Tá Vazando – convidado: Victor Kley	74
Total Geral	510

Fonte: elaboração própria.

Os dados mostram que, em um recorte aproximado de duração, os episódios apresentaram um número similar de ocorrências de sujeito nulo. Em todos os episódios foram analisados vinte minutos de conversação. No *podcast* “Não Vai Cair no ENEM”, do canal Buenas Ideias, a quantidade de sujeitos nulos ficou em 86 no episódio “Os linguiceiros de Porto Alegre”. Já no episódio “A origem do Gaúcho”, foram 93 ocorrências. Entre os números apresentados, no *podcast* “Denílson Show com Chico Garcia”, o episódio em que o ex-jogador de futebol Diego Ribas foi o entrevistado chama a atenção por apresentar 109 ocorrências de sujeito nulo, um número um pouco maior do que os demais, enquanto o episódio que conta com a participação do ex-jogador de futebol Kaká apresenta 79 ocorrências. No *podcast* “Tá Vazando”, o episódio que conta com uma entrevista com a banda Fresno apresenta 69 ocorrências no tempo analisado; no episódio em que o artista Victor Kley foi entrevistado, foram 74 ocorrências.

Esses números demonstram que a proporção da quantidade de ocorrências se mantém em todos os episódios.

Gráfico 4 – Ocorrências de sujeito preenchido e nulo por *podcast* analisado

Fonte: elaboração própria.

O gráfico acima ilustra a preferência por sujeitos expressos em relação aos nulos em cada episódio analisado e demonstra a similaridade das ocorrências de nulos em cada *podcast*. Dessa forma, fica evidenciado o caráter atual do PB, que apresenta sujeitos nulos apenas como casos residuais, diferenciando do padrão anterior da língua, que apontava a preferência por esse tipo de sujeito há dois séculos.

A partir disso, voltamos nossa atenção à investigação dos fatores licenciadores de sujeito nulo a fim de buscar uma explicação que contemple a totalidade de casos desse fenômeno no *corpus* em questão.

4.1 Análise individual de variáveis

Inicialmente, é importante observar as ocorrências de sujeito nulo a partir da análise de fatores licenciadores de maneira individual. Devemos considerar os fatores em Ayres (2021), sendo eles (a) morfologia verbal, (b) padrão linear *V1, (c) gênero semântico do referente e (d) conexão discursiva ótima, a fim de atestar a frequência com que aparecem no *corpus* e, também, se há algum fator que possa explicar de maneira mais ampla os casos residuais do fenômeno.

a) Morfologia verbal

Dos 510 casos de sujeito nulo encontrados no *corpus*, apenas 32,5%, o equivalente a 166 casos, apresentava morfologia verbal rica, isto é, uma forma

desinencial exclusiva para o verbo (ex. acho / sei). Os outros 344 casos apresentavam morfologia verbal pobre, sem a exclusividade da forma verbal, o que, em tese, dificultaria o reconhecimento do referente pelo interlocutor.

Tabela 2 - Ocorrências de sujeito nulo e morfologia verbal (N=510)

Morfologia Verbal	Ocorrências	%
Pobre	344	67,5
Rica	166	32,5
	510	100

Fonte: elaboração própria.

É importante salientar que apenas os referentes de primeira pessoa apresentavam morfologia verbal rica nos contextos analisados. Na fala espontânea, a transcrição revelou que a segunda pessoa, que poderia apresentar também desinência exclusiva, não foi utilizada dessa forma em nenhuma ocorrência. Os exemplos abaixo ilustram casos de primeira, segunda e terceira pessoa respectivamente.

(23) “**Estamos** trazendo esse treinador porque ele demonstrou até agora essa característica.” (Denílson Show – Kaka)

“Eu sei que tu vai fazer piadinha de gaúcho, aliás, eu queria te convidar pra fazer piadinha de gaúcho lá no Alegrete, em Bagé, em Quaraí. **Sabe** o Casseta e Planeta, aquele grupo patético que existia anos atrás?” (Buenas Ideias – Origem do Gaúcho)

“Aí, do nada, ele está lá na frente, baita talento que saiu de ti para o Brasil inteiro, para o mundo. **Deve** ser muito legal.” (Tá Vazando – Victor K.)

A tabela abaixo ilustra a quantidade de ocorrências por pessoa e número do discurso.

Tabela 3 - Ocorrências de sujeito nulo por morfologia verbal e pessoa do discurso (N=510)

Pessoa e Número	Morfologia		Total Geral
	Pobre	Rica	
1	22	166	188
Plural	4	30	34
sing.	18	136	154
2	35	0	35
sing.	35	0	35
3	287	0	287
plural	38	0	38
sing.	249	0	249
	344	166	510

Fonte: elaboração própria.

Foram 188 ocorrências de sujeito nulo em primeira pessoa, sendo 166 em morfologia rica e 22 em morfologia pobre. Na segunda pessoa, foram 35 ocorrências de sujeito nulo, todas sem a marcação de exclusividade desinencial. Na terceira pessoa, foram 287 ocorrências e, também, todas apresentam morfologia pobre, já que a desinência de terceira pessoa poderia apontar a sua referência para os pronomes *ele, ela, você* e *a gente* no singular e *eles, elas e vocês* no plural, o que torna obrigatório que a referência seja alcançada por outros recursos além da morfologia.

Dessa forma, fica claro que esse fator não é suficiente de forma isolada para explicar todos os casos de sujeitos nulos em PB.

b) Padrão linear *V1

Analisando o padrão linear, as ocorrências que consideravam o verbo fora da primeira posição na oração estiveram na faixa dos 30%. Foram 154 ocorrências de sujeito nulo em *V1 e 356 ocorrências de sujeito nulo V1, conforme tabela abaixo.

Tabela 4 - Ocorrências de sujeito nulo e *V1 (N=510)

Padrão linear	Ocorrências	%
V1	356	69,8
*V1	154	30,2
	510	100

Fonte: elaboração própria.

Foram consideradas as construções que apresentavam algum elemento com peso prosódico (como advérbios) à esquerda do verbo:

(24) “Aquele ali foi uma molecagem mesmo, né. Não **faria** de novo.” (Denílson Show – Diego Ribas)

“Já **foi** embora por outro portão.” (Denílson Show – Diego Ribas)

Fica evidente que o fator *V1 também não é capaz de explicar isoladamente de maneira satisfatória os casos de sujeito nulo ainda existentes em PB, já que suas ocorrências permanecem na casa dos 30% dos dados analisados no *corpus*.

c) Gênero semântico

A literatura indica que referentes que não possuem gênero semântico tendem a ser retomados por nulo. Por exemplo:

(25) (Sobre o futebol nos Estados Unidos) **Está** crescendo, tem Copa aí de 2026. (Denílson Show – Kaká)

É importante observar que, no *corpus* analisado, em algumas oportunidades eram feitas perguntas subjetivas ao entrevistado, oportunizando que a resposta fosse de caráter pessoal. Essa subjetividade gera um alto número de usos de primeira pessoa do singular. A primeira pessoa do singular sempre possui gênero semântico, visto que o próprio falante é capaz de reconhecer o seu sexo biológico, o que define a existência de gênero semântico no ser. Além disso, quando há um direcionamento de conversa ao interlocutor utilizando, assim, a segunda pessoa, a referência sempre apontará também para um gênero semântico aparente. Portanto, há uma peculiaridade envolvendo esse fator licenciador: apenas são contabilizados os

referentes de terceira pessoa visto que a primeira e a segunda pessoa sempre terão gênero semântico aparente. Dessa forma, temos a seguinte tabela com o número de ocorrências de sujeito nulo desconsiderando as ocorrências que envolvem a primeira pessoa e, também, a segunda pessoa do discurso.

Tabela 5 - Ocorrências de sujeito nulo e gênero semântico (N=287)

Gênero Semântico	Ocorrências	%
-	218	75,9
+	69	24,1
	287	100

Fonte: elaboração própria.

Na tabela abaixo, aparecem os números de ocorrências de sujeito nulo em um comparativo entre gênero semântico e terceira pessoa do singular e do plural.

Tabela 6 - Ocorrências de sujeito nulo por gênero semântico e pessoa do discurso (N=287)

3ª pessoa	Gênero Semântico		Total
	Sem GS	Com GS	
plural	30	8	38
sing.	188	61	249
	218	69	287

Fonte: elaboração própria.

Na terceira pessoa, há 188 casos de terceira pessoa do singular sem gênero semântico e 30 casos de terceira pessoa do plural sem gênero semântico, totalizando 218 casos que se encaixam no esperado pela teoria. Outros 69 casos somando singular e plural na terceira pessoa apresentam sujeito nulo com gênero semântico aparente. A contrariedade dos dados em relação às expectativas pode ser justificada pela ativação de outros fatores no contexto em que aparecem.

Em todos os casos, fica claro que, também, o gênero semântico não é suficiente para a explicação satisfatória dos casos de sujeito nulo em PB, apesar de nos apresentar boas pistas em relação ao comportamento dos referentes a partir do sexo natural.

d) Conexão discursiva ótima

Em 35,9% dos sujeitos nulos há papel da conexão discursiva ótima. Outros 327 casos (64,1%) não apontavam relação com esse fator licenciador. A tabela abaixo ilustra esses números.

Tabela 7 - Ocorrências de sujeito nulo e conexão discursiva ótima (N=510)

Conexão Discursiva Ótima	Ocorrências	%
Não	327	64,1
Sim	183	35,9
	510	100

Fonte: elaboração própria.

Foram considerados os casos em que o sujeito é identificado a partir da manutenção dos aspectos do referente de orações anteriores:

(25) “A gente meio que fazia esse relacionamento com os outros colégios para de fato tentar tocar. **Marcava** show no Garagem Hermética.” (Tá Vazando – Fresno)

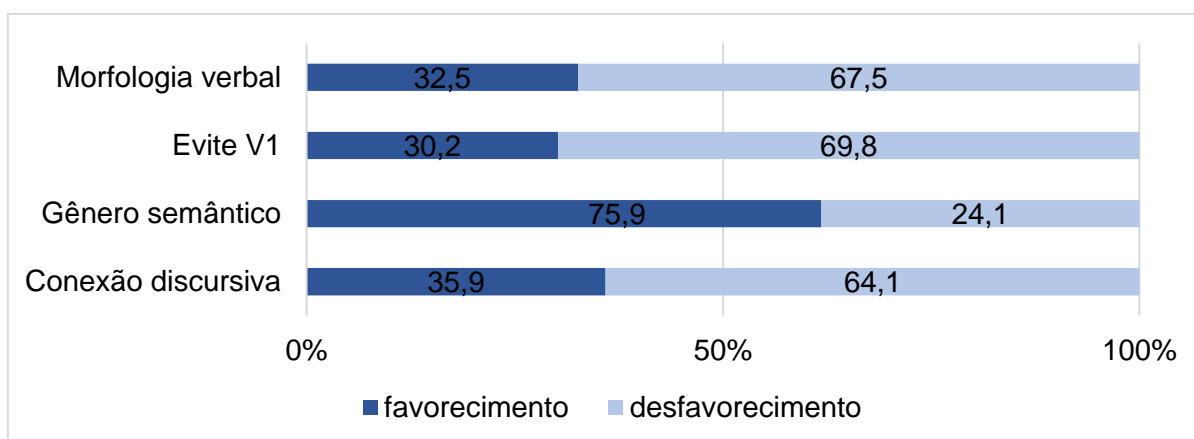
“Preciso te dizer que esse Ramos era tido como um cara de olhos enormes com a voz sinistra, uma barba rala tipo a minha. (...) **Vivia** sempre exageradamente perfumado.” (Buenas Ideias – Linguiceiros de POA)

Nos casos em que esse fator licenciador não aparece como central para o fenômeno, espera-se que outros fatores estejam ativados. Mesmo assim, também é possível perceber que a conexão discursiva ótima não explica a totalidade dos casos de sujeito nulo em PB quando analisada de maneira isolada.

O gráfico abaixo ilustra a porcentagem de ocorrências de sujeito nulo explicadas por cada fator isolado.⁹

⁹ Os fatores morfologia verbal, padrão linear *V1 e conexão discursiva ótima foram calculados sobre 510 ocorrências, enquanto o fator gênero semântico foi calculado sobre 287 ocorrências.

Gráfico 5 – Ocorrência de sujeito nulo (%) por variável analisada



Fonte: elaboração própria.

É possível perceber que o fator que mais aparece como influência direta nos casos de sujeito nulo é o gênero semântico, ultrapassando a metade dos dados. Mesmo assim, nenhum dos traços é capaz de explicar o licenciamento dos sujeitos nulos em PB isoladamente de maneira satisfatória a fim de contemplar a totalidade dos dados. Dessa forma, o próximo passo é analisar os dados a partir do conjunto de dois fatores, com o objetivo de investigar se dois fatores combinados são capazes de explicar os licenciamentos de sujeito nulo de maneira plena.

4.2 Análise dupla de variáveis

A investigação, agora, será concentrada na união de dois fatores licenciadores de sujeito nulo. Também, é importante observar se essa união será suficiente para explicar os casos de sujeito nulo em PB na sua totalidade.

Para facilitar a leitura e a análise, iremos adotar o seguinte sistema: quando o fator for licenciador de sujeitos nulos em PB, colocaremos o sinal ✓ ao lado; quando o fator não for licenciador de sujeitos nulos em PB, colocaremos o sinal x ao lado.

Nossa análise leva em conta os 510 casos de sujeitos nulos a partir dos quatro fatores de Ayres (2021). Nas análises envolvendo o fator gênero semântico, apresentaremos, ainda, uma nova alternativa de análise considerando apenas os 287 casos de terceira pessoa, já que os casos de primeira e segunda pessoa sempre apresentam gênero semântico aparente. Por esse motivo, teremos as seguintes análises:

Análise sobre 510 ocorrências considerando todas as pessoas do discurso:

- a) morfologia verbal e padrão linear *V1;
- b) morfologia verbal e conexão discursiva ótima;
- c) padrão linear *V1 e conexão discursiva ótima;

Análise sobre 510 ocorrências considerando todas as pessoas do discurso e, também, análise de 287 ocorrências considerando apenas a terceira pessoa do discurso:

- d) morfologia verbal e gênero semântico;
- e) padrão linear *V1 e gênero semântico;
- f) gênero semântico e conexão discursiva ótima.

- a) morfologia verbal e padrão linear *V1

Na tabela a seguir, estão os dados referentes à união entre morfologia verbal e padrão linear *V1.

Tabela 8 - Ocorrências de sujeito nulo por morfologia verbal e *V1 (N=510)

Morfologia	*V1	Ocorrências	%
Pobre x	V1 x	241	47,2%
Pobre x	*V1 ✓	103	20,3%
Rica ✓	V1 x	115	22,5%
Rica ✓	*V1 ✓	51	10%
		510	100%

Fonte: elaboração própria.

A combinação que une dois fatores licenciadores de sujeito nulo apresenta morfologia verbal rica e a marcação positiva para *V1. A junção dos dois fatores atuando no licenciamento do fenômeno acontece em 51 casos, o que equivale a 10%. Se considerarmos a atuação de pelo menos um dos fatores, teremos, então, 269 ocorrências, o que representa 52,8% dos casos.

Na contramão, 241 casos de sujeito nulo não são influenciados por nenhum desses dois fatores, já que estão marcados negativamente para o padrão linear *V1 e apresentam morfologia verbal pobre. Esse número representa 47,2% das ocorrências

de sujeito nulo, deixando claro que esses dois traços apenas não são capazes de explicar o fenômeno em sua totalidade.

b) morfologia verbal e conexão discursiva ótima

A combinação entre morfologia verbal e conexão discursiva ótima está ilustrada na tabela a seguir.

Tabela 9 - Ocorrências de sujeito nulo por morfologia verbal e conexão discursiva (N=510)

Morfologia Verbal	Conexão Discursiva Ótima	Ocorrências	%
Pobre x	Não x	186	36,5%
Pobre x	Sim ✓	158	31%
Rica ✓	Não x	141	27,6%
Rica ✓	Sim ✓	25	4,9%
		510	100%

Fonte: elaboração própria.

Nesse conjunto de dois fatores, os contextos licenciadores de sujeito nulo são os de morfologia verbal rica e marcação positiva para conexão discursiva ótima. Apenas 25 ocorrências estão nessa intersecção. Observando pelo menos um desses dois fatores como atuante para a ocorrência do fenômeno, teremos 324 ocorrências, o que significa que 63,5% dos casos de sujeito nulo em PB são explicados por essa dupla.

Por outro lado, existem 186 casos em que estão envolvidos fatores que, na verdade, não licenciam o fenômeno, já que não há conexão discursiva ótima e a morfologia verbal é pobre. Esses casos chegam a 36,5% das ocorrências, revelando que essa dupla de fatores também não é suficiente para a explicação da totalidade de casos residuais do fenômeno.

c) padrão linear *V1 e conexão discursiva ótima

Na tabela a seguir, temos a combinação dos fatores que envolvem padrão linear *V1 e conexão discursiva ótima. A distribuição de sujeitos nulos ficou da seguinte forma.

Tabela 10 - Ocorrências de sujeito nulo por evite V1 e conexão discursiva (N=510)

*V1	Conexão Discursiva Ótima	Ocorrências	%
V1 x	Não x	227	44,5%
V1 x	Sim ✓	129	25,3%
*V1 ✓	Não x	100	19,6%
*V1 ✓	Sim ✓	54	10,6%
		510	100%

Fonte: elaboração própria.

Essa combinação de fatores licencia o sujeito nulo a partir do momento em que há marcação positiva para o padrão linear *V1 e, também, positiva para a conexão discursiva ótima. Assim, os dados revelam que a união dos dois fatores agindo ao mesmo tempo ocorre em 54 casos (10,6%). Além disso, é possível perceber que em 283 casos há pelo menos um desses fatores atuando. Dessa forma, tem-se 55,5% das ocorrências residuais de sujeitos nulos sendo explicadas por essa combinação.

Mesmo assim, ainda restam 227 casos em que os fatores que operam não são licenciadores de sujeito nulo. Isso representa 44,5% dos casos. Dessa forma, a combinação entre *V1 e conexão discursiva ótima não foi capaz de explicar as ocorrências de sujeito nulo em PB mais uma vez.

d) morfologia verbal e gênero semântico

A seguinte tabela representa a união entre morfologia verbal e gênero semântico sobre 510 ocorrências de sujeito nulo no *corpus*.

Tabela 11 - Ocorrências de sujeito nulo por morfologia verbal e gênero semântico (N=510)

Morfologia Verbal	Gênero Semântico	Ocorrências	%
Pobre x	- ✓	253	49,6%
Pobre x	+ x	91	17,8%
Rica ✓	- ✓	17	3,3%
Rica ✓	+ x	149	29,3%
		510	100%

Fonte: elaboração própria.

A união entre morfologia rica e referentes sem gênero semântico é o alvo dessa análise. Com os dados, é possível perceber que apenas 17 casos apresentam a intersecção entre os dois fatores licenciadores, representando 3,3% das ocorrências. Anteriormente explicados, os casos envolvendo morfologia rica são de primeira pessoa do discurso. Se considerarmos a primeira pessoa do singular, então o referente sempre terá gênero semântico aparente. Se considerarmos primeira pessoa do plural, o referente poderá aparecer sem ou com o gênero semântico aparente.

Compreende-se, então, que os fatores licenciadores de sujeito nulo envolvem a morfologia rica e os referentes sem gênero semântico aparente. Se considerarmos pelo menos um desses fatores licenciadores na atuação do sujeito nulo, teremos o equivalente a 419 ocorrências, totalizando 82,2% de casos explicados por esse par.

Analisando a combinação de fatores que não licenciam o sujeito nulo em PB, morfologia pobre e antecedentes com gênero semântico, percebe-se que os dados apresentam 91 ocorrências em que esses fatores estão envolvidos, o que representa 17,8% dos casos. Esse percentual já é suficiente para demonstrar que esses dois fatores combinados não são suficientes para explicar as ocorrências de sujeito nulo, apesar de termos um bom número de casos sendo explicados por essa combinação.

Por outro lado, a análise pode levar em conta que a primeira e a segunda pessoa do discurso sempre apresentam gênero semântico aparente. Dessa forma, podemos voltar a observação aos dados que apontam apenas a marcação de gênero semântico para referentes de terceira pessoa (ex. (a gadaria) se ***espalha*** por essas campinas sem fim)¹⁰. Assim, temos uma nova tabela, dessa vez sobre 287 casos.

¹⁰ Exemplo retirado do *podcast* Buenas Ideias, episódio A origem do Gaúcho.

Tabela 12 - Ocorrências de sujeito nulo por morfologia verbal e gênero semântico sem 1ª e 2ª pessoa
(N=287)

Morfologia Verbal	Gênero Semântico	Ocorrências	%
Pobre x	- ✓	69	24,1%
Pobre x	+ x	218	75,9%
Rica ✓	- ✓	0	0
Rica ✓	+ x	0	0
		287	100%

Fonte: elaboração própria.

Nesse caso, ao considerarmos apenas as ocorrências de terceira pessoa, naturalmente teremos apenas verbos de morfologia pobre. A divisão torna-se simples, pois teremos apenas a diferença entre referentes com marcação positiva para gênero semântico e referentes com marcação negativa. Apenas 69 casos dentre os 287 são licenciados por esse fator, representando 24,1%. Outros 75,9% dos casos acontecem sem que riqueza morfológica ou gênero semântico estejam envolvidos em seu licenciamento. Assim, essa combinação demonstra ser insuficiente para a explicação dos casos de sujeito nulo em PB.

e) padrão linear *V1 e gênero semântico

Na tabela abaixo, a união entre padrão linear *V1 e gênero semântico aparece representada.

Tabela 13 - Ocorrências de sujeito nulo por evite V1 e gênero semântico (N=510)

*V1	Gênero Semântico	Ocorrências	%
V1 x	- ✓	187	36,7%
V1 x	+ x	169	33,1%
*V1 ✓	- ✓	83	16,3%
*V1 ✓	+ x	71	13,9%
		510	100%

Fonte: elaboração própria.

Nessa dupla de fatores, os sujeitos nulos são licenciados quando há marcação positiva para *V1 e quando o referente não possui gênero semântico aparente. Nesse caso, a intersecção dos dois fatores revela 83 ocorrências (16,3% dos casos). Além disso, outros 187 casos são licenciados pelo antecedente sem gênero semântico aparente e 71 casos são licenciados pela marca positiva para *V1. A soma dos casos, levando em conta a intersecção dos fatores, totaliza 341 ocorrências, o que representa 66,9%.

Contudo, restam 169 ocorrências em que há fatores não licenciadores envolvidos, já que nesses casos não há *V1 e o referente possui gênero semântico aparente. Esses dados representam 33,1% de sujeitos nulos sem o envolvimento desses fatores. Assim, fica evidenciado que esses dois fatores combinados são insuficientes para a explicação da totalidade de dados residuais do fenômeno.

Além disso, podemos novamente analisar esses casos envolvendo gênero semântico a partir da observação das ocorrências apenas de terceira pessoa. Assim, teremos a seguinte tabela.

Tabela 14 - Ocorrências de sujeito nulo por evite V1 e gênero semântico sem 1ª e 2ª pessoa (N=287)

*V1	Gênero Semântico	Ocorrências	%
V1 x	- ✓	54	18,8%
V1 x	+ x	141	49,1%
*V1 ✓	- ✓	15	5,3%
*V1 ✓	+ x	77	26,8%
		287	100%

Fonte: elaboração própria.

Na nova configuração, chegamos a apenas 15 casos em que os dois fatores ocorrem simultaneamente e a 146 casos (50,9%) no total em que pelo menos um desses dois fatores está envolvido no licenciamento do fenômeno. Ademais, 141 casos (49,1%) não tem nenhum desses fatores licenciadores envolvidos, o que evidencia que essa combinação não é suficiente para a explicação dos sujeitos nulos sob nenhuma análise.

f) gênero semântico e conexão discursiva ótima

As últimas tabelas dessa seção apresentam os dados a respeito da união entre gênero semântico e conexão discursiva ótima.

Tabela 15 - Ocorrências de sujeito nulo por gênero semântico e conexão discursiva (N=510)

Gênero Semântico	Conexão Discursiva Ótima	Ocorrências	%
- ✓	Não x	182	35,7%
- ✓	Sim ✓	88	17,3%
+ x	Não x	145	28,4%
+ x	Sim ✓	95	18,6%
		510	100%

Fonte: elaboração própria.

Nessas últimas tentativas de combinar dois fatores, o sujeito nulo é licenciado quando o referente não possui gênero semântico aparente e quando há marcação positiva para conexão discursiva ótima na tabela. Com base nisso, é possível identificar pela tabela que em 365 casos há pelo menos um desses dois fatores atuando e em 88 casos os dois fatores juntos atuam para o licenciamento do sujeito nulo. Esses dados representam 71,6% dos casos.

Mesmo com esse alto índice, ainda há 28,4% de casos em que há atuação de fatores não licenciadores (nesse caso, referente com gênero semântico aparente e marcação negativa para conexão discursiva ótima), totalizando 145 ocorrências. Desse modo, novamente, a combinação dos fatores gênero semântico e conexão discursiva ótima não é suficiente para explicar a totalidade dos casos residuais de sujeito nulo em PB.

Ainda, se retirarmos os referentes de primeira e segunda pessoa do discurso dessa análise para que possamos observar os casos de terceira pessoa relacionados ao gênero semântico, teremos a seguinte tabela.

Tabela 16 - Ocorrências de sujeito nulo por gênero semântico e conexão discursiva sem 1ª e 2ª pessoa (N=287)

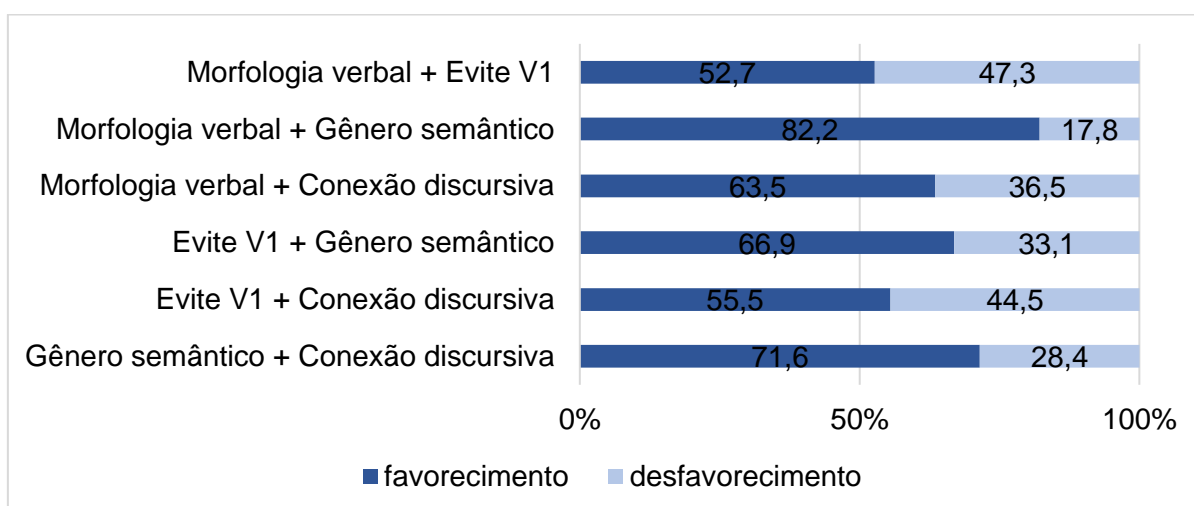
Gênero Semântico	Conexão Discursiva Ótima	Ocorrências	%
- ✓	Não x	146	50,9%
- ✓	Sim ✓	72	25,1%
+ x	Não x	7	2,4%
+ x	Sim ✓	62	21,6%
		287	100%

Fonte: elaboração própria.

Na nova tabela, entre os 287 dados que consideram as ocorrências de terceira pessoa, 280 casos são explicados por pelo menos um dos fatores analisados, o que representa o total de 97,6% dos casos de sujeitos nulos em PB. Apenas 7 casos (2,4%) não são explicados por nenhum desses dois fatores. Aqui temos dados realmente promissores em busca do objetivo de explicar os casos do fenômeno em PB.

Os gráficos abaixo ilustram o percentual de casos explicados e não explicados em cada uma das duplas analisadas.

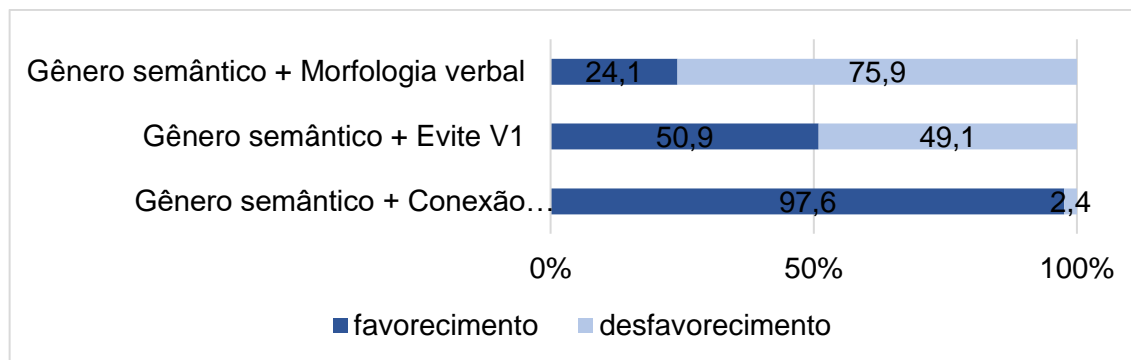
Gráfico 6 – Ocorrência de sujeito nulo (%) por duplas de variáveis analisadas (N=510)



Fonte: elaboração própria.

No gráfico acima, temos os percentuais relacionados a 510 ocorrências de sujeito nulo. No gráfico abaixo, apenas as análises envolvendo gênero semântico com ocorrências de terceira pessoa do discurso considerando 287 casos.

Gráfico 7 – Ocorrência de sujeito nulo (%) por duplas de variáveis analisadas sem 1ª e 2ª pessoa (N=287)



Fonte: elaboração própria.

A análise de dois fatores se mostrou bastante promissora para a explicação dos casos de sujeito nulo em PB, alcançando a importante marca de 97,6% de casos explicados pela combinação entre gênero semântico e conexão discursiva ótima nas ocorrências de terceira pessoa. Entretanto, como buscamos a explicação de 100% dos casos, torna-se relevante, então, investigar outras combinações. Por esse motivo, é crucial partir para a investigação da combinação entre três fatores licenciadores de sujeito nulo para que essa explicação seja ainda mais satisfatória.

4.3 Análise tripla de variáveis

A análise com apenas dois fatores não foi suficiente para as explicações de todos os casos de sujeito nulo em PB, então, agora, é necessário partir para a investigação que envolve a combinação de três fatores licenciadores a fim de buscar explicações para 100% dos casos residuais do fenômeno. Seguimos considerando os 510 casos de sujeitos nulos a partir dos quatro fatores de Ayres (2021). Entretanto, assim como na análise dupla de variáveis, apresentaremos uma análise alternativa considerando apenas os 287 casos de terceira pessoa quando o fator gênero semântico estiver envolvido na combinação, visto que os casos de primeira e segunda pessoa sempre apresentam gênero semântico aparente. Dessa forma, teremos as seguintes análises:

Análise sobre 510 ocorrências considerando todas as pessoas do discurso:

- a) morfologia verbal, conexão discursiva ótima e padrão linear *V1;

Análise sobre 510 ocorrências considerando todas as pessoas do discurso e, também, análise de 287 ocorrências considerando apenas a terceira pessoa do discurso:

- b) morfologia verbal, gênero semântico e padrão linear *V1;
- c) morfologia verbal, conexão discursiva ótima e gênero semântico;
- d) padrão linear *V1, conexão discursiva ótima e gênero semântico.

- a) Morfologia verbal, conexão discursiva ótima e padrão *V1

A tabela abaixo apresenta os dados referentes à combinação de morfologia verbal, conexão discursiva ótima e padrão linear *V1. Observemos os resultados.

Tabela 17 - Ocorrências de sujeito nulo por morfologia verbal, conexão discursiva e *V1 (N=510)

MORFOLOGIA	CDO	*V1	Ocorrências	%
Pobre x	Sim ✓	*V1 ✓	42	8,2%
Pobre x	Sim ✓	V1 x	116	22,7%
Pobre x	Não x	*V1 ✓	61	12%
Pobre x	Não x	V1 x	125	24,5%
Rica ✓	Sim ✓	*V1 ✓	12	2,4%
Rica ✓	Sim ✓	V1 x	13	2,5%
Rica ✓	Não x	*V1 ✓	39	7,7%
Rica ✓	Não x	V1 x	102	20%
			510	100%

Fonte: elaboração própria.

Na união desses três fatores, os sujeitos nulos são licenciados sob os seguintes contextos: morfologia verbal rica, marcação positiva para conexão discursiva ótima e marcação positiva para padrão linear *V1. A intersecção desses critérios aponta para apenas 12 casos no corpus. Também, é importante ressaltar que em 75,5% das ocorrências pelo menos um desses fatores está envolvido.

Entretanto, se observarmos os fatores que não licenciam sujeitos nulos em PB, encontraremos 125 ocorrências do fenômeno com a intersecção de morfologia pobre, marca negativa para conexão discursiva ótima e marca negativa para padrão linear *V1. Esse número representa a marca de 24,5% de sujeitos nulos sob fatores não

licenciadores do fenômeno. O número é mais expressivo do que o da combinação anterior. Dessa forma, fica evidente que essa união também não dá conta de todos os dados de maneira satisfatória.

b) Morfologia verbal, gênero semântico e padrão linear *V1.

Na tabela a seguir, estão os dados referentes à união do trio entre morfologia verbal, gênero semântico e padrão linear *V1.

Tabela 18 - Ocorrências de sujeito nulo por morfologia verbal, *V1 e gênero semântico (N=510)

MORFOLOGIA	GS	*V1	Ocorrências	%
Pobre x	- ✓	*V1 ✓	81	15,9%
Pobre x	- ✓	V1 x	172	33,7%
Pobre x	+ x	*V1 ✓	22	4,3%
Pobre x	+ x	V1 x	69	13,5%
Rica ✓	- ✓	*V1 ✓	2	0,4%
Rica ✓	- ✓	V1 x	15	2,9%
Rica ✓	+ x	*V1 ✓	49	9,7%
Rica ✓	+ x	V1 x	100	19,6%
			510	100%

Fonte: elaboração própria.

Nessa combinação, os contextos licenciadores de sujeito nulo apresentam morfologia rica, referente sem gênero semântico e marcação positiva para padrão linear *V1. Esses critérios são atendidos em apenas 2 casos de maneira concomitante. Ademais, se considerar os casos em que pelo menos um desses critérios é atendido, chegamos ao total de 86,5% dos casos de sujeito nulo em PB explicados pelos fatores descritos em Ayres (2021).

Por outro lado, os contextos que não licenciam o sujeito nulo apontam para morfologia pobre, referente com gênero semântico e marcação negativa para padrão linear *V1. Isso acontece em 69 casos de acordo com a tabela acima, totalizando 13,5% das ocorrências.

Entretanto, como há gênero semântico envolvido na combinação, podemos observar apenas os casos de terceira pessoa, chegando a uma nova tabela ao eliminarmos os casos de primeira e de segunda pessoa.

Tabela 19 - Ocorrências de sujeito nulo por morfologia verbal, evite V1 e gênero semântico sem 1ª e 2ª pessoa (N=287)

MORFOLOGIA	GS	*V1	Ocorrências	%
Pobre x	- ✓	*V1 ✓	77	26,8 %
Pobre x	- ✓	V1 x	141	49,1%
Pobre x	+ x	*V1 ✓	15	5,2%
Pobre x	+ x	V1 x	54	18,9%
Rica ✓	- ✓	*V1 ✓	0	0
Rica ✓	- ✓	V1 x	0	0
Rica ✓	+ x	*V1 ✓	0	0
Rica ✓	+ x	V1 x	0	0
			287	100%

Fonte: elaboração própria.

Na nova análise, não há morfologia rica envolvida já que todos os casos apontam para a terceira pessoa do discurso. Dessa forma, é possível observar que há 233 casos (81,1%) em que há pelo menos um desses fatores envolvidos. Apenas 54 casos (18,9%) não tem nenhum desses três fatores atuando no licenciamento do fenômeno, o que demonstra que essa combinação não explica de forma satisfatória os casos de sujeito nulo em PB.

c) Morfologia verbal, conexão discursiva ótima e gênero semântico

A combinação de fatores seguinte une morfologia verbal, conexão discursiva ótima e gênero semântico. Na tabela abaixo, estão expressos os dados dessa união.

Tabela 20 - Ocorrências de sujeito nulo por morfologia verbal, gênero semântico e conexão discursiva
(N=510)

MORFOLOGIA	CDO	GS	Ocorrências	%
Pobre x	Sim ✓	- ✓	88	17,3%
Pobre x	Sim ✓	+ x	70	13,7%
Pobre x	Não x	- ✓	165	32,4%
Pobre x	Não x	+ x	21	4,1%
Rica ✓	Sim ✓	- ✓	0	0
Rica ✓	Sim ✓	+ x	25	4,9%
Rica ✓	Não x	- ✓	17	3,3%
Rica ✓	Não x	+ x	124	24,3%
			510	100%

Fonte: elaboração própria.

O licenciamento de sujeitos nulos a partir desses três fatores acontece quando há morfologia rica, marcação positiva para conexão discursiva ótima e referente sem gênero semântico. Nenhum caso foi constatado a partir da intersecção desses três critérios. Mesmo assim, se observarmos pelo menos um desses fatores como atuante para a ocorrência do fenômeno, chegaremos à marca de 95,9% dos sujeitos nulos explicados pela teoria.

Inversamente, são fatores não licenciadores de sujeito nulo: morfologia verbal pobre, marcação negativa para conexão discursiva ótima e referente com gênero semântico. Foram encontrados 21 casos de sujeito nulo em que acontece a intersecção desses fatores não licenciadores, o que representa apenas 4,1% das ocorrências. Ainda que o número seja baixo, os três fatores acima não se mostram suficientes para a explicação satisfatória da totalidade dos casos de sujeito nulo, mesmo que os dados pareçam bastante promissores.

Por outro lado, novamente temos a participação do gênero semântico na combinação, o que nos permite analisar apenas os casos de terceira pessoa do discurso. Dessa forma, a nova tabela eliminando os casos de primeira e de segunda pessoa fica da seguinte maneira.

Tabela 21 - Ocorrências de sujeito nulo por morfologia verbal, gênero semântico e conexão discursiva sem 1ª e 2ª pessoa (N=287)

MORFOLOGIA	CDO	GS	Ocorrências	%
Pobre x	Sim ✓	- ✓	72	25,1%
Pobre x	Sim ✓	+ x	62	21,6%
Pobre x	Não x	- ✓	146	50,9%
Pobre x	Não x	+ x	7	2,4%
Rica ✓	Sim ✓	- ✓	0	0
Rica ✓	Sim ✓	+ x	0	0
Rica ✓	Não x	- ✓	0	0
Rica ✓	Não x	+ x	0	0
			287	100%

Fonte: elaboração própria.

Novamente, por estarmos analisando apenas casos de terceira pessoa, não há nenhum dado para morfologia rica. Dessa forma, desconsiderando a riqueza morfológica já que todos os casos passam a envolver morfologia pobre, temos uma tabela idêntica à investigação de dois fatores envolvendo conexão discursiva ótima e gênero semântico, o que apresentou os dados mais promissores até então: 97,6% dos casos são explicados por pelo menos um desses fatores enquanto apenas 2,4% dos casos não possuem nenhum desses fatores agindo no licenciamento do fenômeno.

d) Padrão linear *V1, conexão discursiva ótima e gênero semântico

Por fim, as últimas combinações de três fatores a serem analisadas unem padrão linear *V1, conexão discursiva ótima e gênero semântico. A tabela abaixo apresenta os dados dessa união sobre as 510 ocorrências do fenômeno no *corpus*.

Tabela 22 - Ocorrências de sujeito nulo por *V1, conexão discursiva e gênero semântico (N=510)

*V1	CDO	GS	Ocorrências	%
*V1 ✓	Sim ✓	- ✓	24	4,7%
*V1 ✓	Sim ✓	+ x	30	5,9%
*V1 ✓	Não x	- ✓	59	11,7%
*V1 ✓	Não x	+ x	41	8%
V1 x	Sim ✓	- ✓	64	12,5%
V1 x	Sim ✓	+ x	65	12,7%
V1 x	Não x	- ✓	123	24,1%
V1 x	Não x	+ x	104	20,4%
			510	100%

Fonte: elaboração própria.

Essa última combinação apresenta o seguinte contexto para o licenciamento de sujeitos nulos: marcação positiva para padrão linear *V1, marcação positiva para conexão discursiva ótima e referente sem gênero semântico. A intersecção desses fatores é apresentada na primeira linha da tabela e expõe a marca de 24 casos (4,7%). Se considerarmos pelo menos um desses três fatores como atuante para o licenciamento de nulos no corpus, atingiremos, então, 406 ocorrências, representando 79,6% dos casos.

Contudo, os fatores não licenciadores de sujeito nulo são: marcação negativa para padrão linear *V1, marcação negativa para conexão discursiva ótima e referente com gênero semântico. A última linha da tabela mostra que em 104 casos, a intersecção dos fatores não licenciadores apareceu, o que representa a marca de 20,4% dos casos. Dessa forma, mais uma vez, a missão de explicar a totalidade de dados de sujeito nulo no corpus a partir da união de três fatores não foi atingida.

Mais uma vez, estamos lidando aqui com uma combinação que envolve o gênero semântico. Por esse motivo, podemos observar apenas os casos de terceira pessoa do discurso, chegando a uma nova tabela de dados.

Tabela 23 - Ocorrências de sujeito nulo por *V1, conexão discursiva e gênero semântico sem 1ª e 2ª pessoa (N=287)

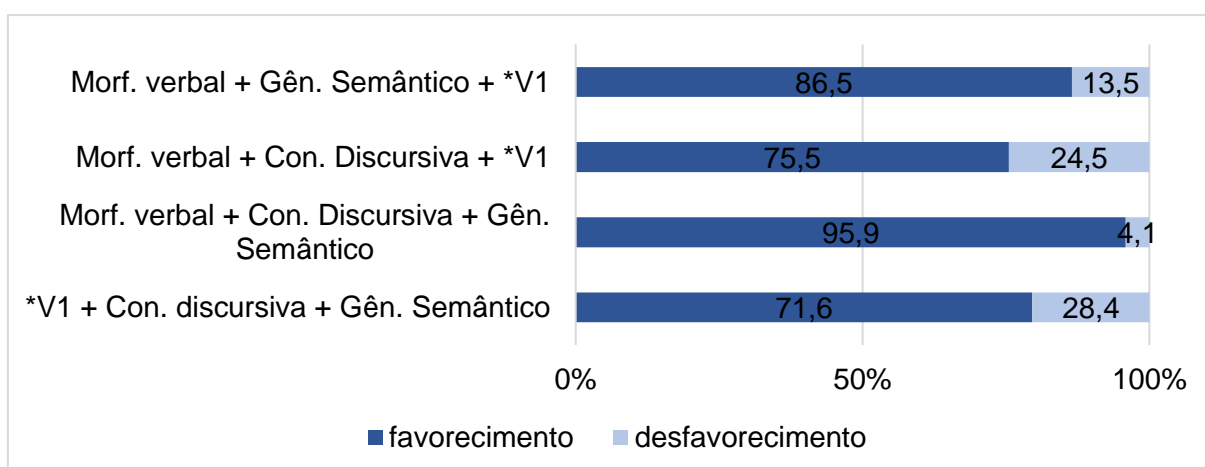
*V1	CDO	GS	Ocorrências	%
*V1 ✓	Sim ✓	- ✓	24	8,4%
*V1 ✓	Sim ✓	+ x	14	4,9%
*V1 ✓	Não x	- ✓	53	18,5%
*V1 ✓	Não x	+ x	1	0,3%
V1 x	Sim ✓	- ✓	48	16,7%
V1 x	Sim ✓	+ x	48	16,7%
V1 x	Não x	- ✓	93	32,4%
V1 x	Não x	+ x	6	2,1%
			287	100%

Fonte: elaboração própria.

A nova análise permite observar que em 281 casos há pelo menos um dos três fatores envolvidos, o que representa a importante marca de 97,9% das ocorrências sendo explicadas por essa combinação. Apenas 6 casos (2,1%) não têm nenhum desses fatores atuando no licenciamento do fenômeno.

Os gráficos abaixo ilustram o percentual de casos explicados e não explicados em cada um dos trios analisados.

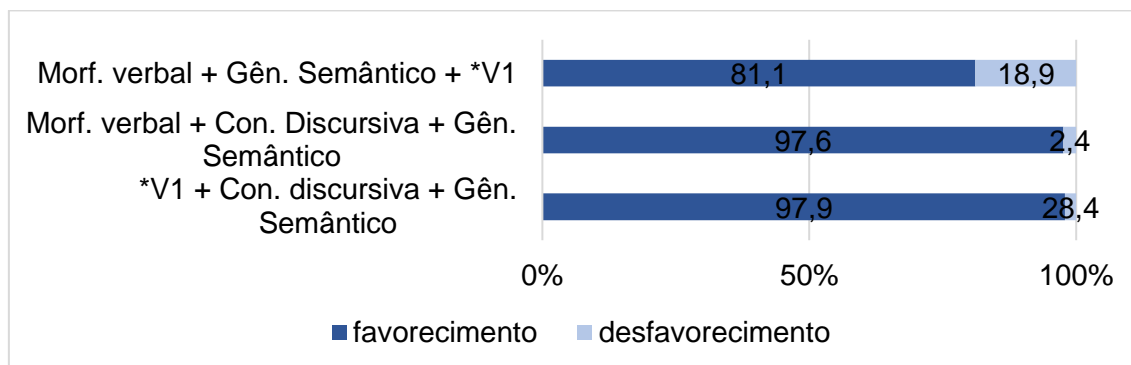
Gráfico 8 – Ocorrência de sujeito nulo (%) por trios de variáveis analisadas (N=510)



Fonte: elaboração própria.

No gráfico acima, temos os percentuais relacionados a 510 ocorrências de sujeito nulo; no gráfico abaixo, apenas as análises envolvendo gênero semântico com ocorrências de terceira pessoa do discurso sobre, então, 287 ocorrências.

Gráfico 9 – Ocorrência de sujeito nulo (%) por trios de variáveis analisadas sem 1ª e 2ª pessoa (N=287)



Fonte: elaboração própria.

A análise de três fatores se mostrou bastante eficiente para a explicação dos casos remanescentes de sujeito nulo em PB. Foi possível encontrar a relevante marca de 97,9% de ocorrências explicadas pela combinação de padrão linear *V1, conexão discursiva ótima e gênero semântico. Temos apenas 2,1% de casos não explicados por essa combinação é um resultado bastante promissor. Entretanto, o objetivo do trabalho é alcançar a explicação de 100% dos casos. Por isso, é válido, ainda, avançar para a investigação da combinação dos quatro fatores para entendermos se juntos, podemos encontrar marcas ainda mais significativas do que as encontradas até então.

4.4 Análise quádrupla de variáveis

A busca por explicações para a totalidade de sujeitos nulos passou por algumas etapas: analisamos apenas um fator, evoluímos para a combinação de dois fatores, passamos por três fatores e, agora, é necessário olhar com atenção para a combinação de todos os quatro fatores na expectativa de atingir o objetivo. Ainda que na combinação de três fatores tenhamos encontrado uma configuração que explica 97,9% dos casos, há a esperança de verificarmos se o acréscimo de um fator auxilia na explicação de 100%.

A tabela abaixo apresenta os dados referentes à combinação de morfologia verbal, padrão linear *V1, gênero semântico e conexão discursiva ótima considerando todos as 510 ocorrências do corpus.

Tabela 24 – Ocorrências de sujeito nulo por morfologia verbal, *V1, gênero semântico e conexão discursiva (N=510)

MORFOLOGIA	*V1	GS	CDO	Ocorrências	%
Pobre x	*V1 ✓	- ✓	Sim ✓	24	4,7%
Pobre x	*V1 ✓	- ✓	Não x	57	11,2%
Pobre x	*V1 ✓	+ x	Sim ✓	18	3,5%
Pobre x	*V1 ✓	+ x	Não x	4	0,8%
Pobre x	V1 x	- ✓	Sim ✓	64	12,6%
Pobre x	V1 x	- ✓	Não x	108	21,2%
Pobre x	V1 x	+ x	Sim ✓	52	10,2%
Pobre x	V1 x	+ x	Não x	17	3,3%
Rica ✓	*V1 ✓	- ✓	Sim ✓	0	0
Rica ✓	*V1 ✓	- ✓	Não x	2	0,4%
Rica ✓	*V1 ✓	+ x	Sim ✓	12	2,3%
Rica ✓	*V1 ✓	+ x	Não x	37	7,3%
Rica ✓	V1 x	- ✓	Sim ✓	0	0
Rica ✓	V1 x	- ✓	Não x	15	2,9%
Rica ✓	V1 x	+ x	Sim ✓	13	2,6%
Rica ✓	V1 x	+ x	Não x	87	17%
				510	100%

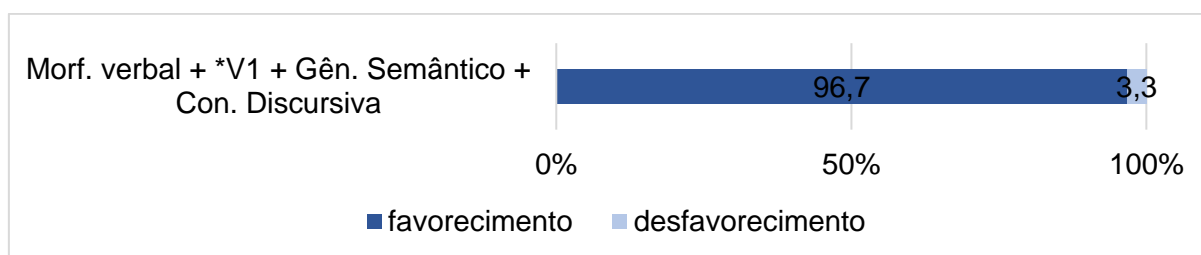
Fonte: elaboração própria.

O licenciamento de sujeitos nulos é feito a partir dos seguintes critérios na tabela: morfologia verbal rica, marcação positiva para padrão linear *V1, referente sem gênero semântico aparente e marcação positiva para conexão discursiva ótima. Com base nisso, os dados mostram que nenhum caso foi observado a partir da concomitância desses fatores. Se analisarmos os casos em que pelo menos um

desses fatores foi ativado, perceberemos o total de 493 ocorrências, o que representa 96,7% de casos de sujeitos nulos explicados pelos quatro fatores expostos por Ayres (2021).

Por outro lado, morfologia verbal pobre, marcação negativa para padrão linear *V1, referente com gênero semântico aparente e marcação negativa para conexão discursiva ótima são fatores não licenciadores de sujeitos nulos. Foram encontrados 17 casos na intersecção dos critérios não licenciadores, representando 3,3% dos casos, o que contraria fortemente a expectativa gerada pela teoria.

Gráfico 10 – Ocorrência de sujeito nulo (%) por todas as variáveis analisadas (N=510)



Fonte: elaboração própria.

Um dos objetivos desse trabalho é oferecer pontos de análise e dados suficientes para que as investigações a respeito do sujeito nulo em PB possam ter avanços significativos. Nesse sentido, apresentar dados sobre os quatro fatores incluindo os casos de gênero semântico para primeira e segunda pessoa enriquece o trabalho, mesmo que, nesses casos, sempre haja marcação positiva para gênero semântico, o que não é um fator licenciador do fenômeno. Entretanto, para que tenhamos a análise aperfeiçoada, é necessário, agora, observar os casos envolvendo apenas a terceira pessoa do discurso, pois esses, sim, podem apresentar referentes sem gênero semântico, o que é um fator licenciador de sujeito nulo em PB. Nessa perspectiva, é possível observar uma nova tabela de dados, dispensando, então, as ocorrências de primeira e segunda pessoa.

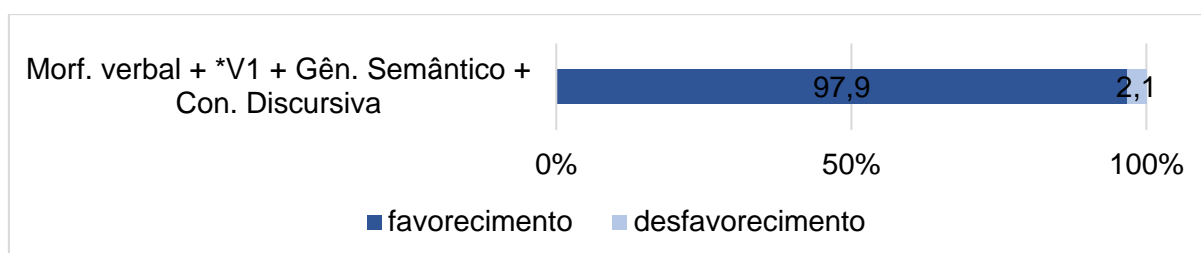
Tabela 25 – Ocorrências de sujeito nulo por morfologia verbal, *V1, gênero semântico e conexão discursiva sem 1ª e 2ª pessoa (N=287)

MORFOLOGIA	*V1	GS	CDO	Ocorrências	%
Pobre x	*V1 ✓	- ✓	Sim ✓	24	8,4%
Pobre x	*V1 ✓	- ✓	Não x	53	18,5%
Pobre x	*V1 ✓	+ x	Sim ✓	14	4,9%
Pobre x	*V1 ✓	+ x	Não x	1	0,3%
Pobre x	V1 x	- ✓	Sim ✓	48	16,7%
Pobre x	V1 x	- ✓	Não x	93	32,4%
Pobre x	V1 x	+ x	Sim ✓	48	16,7%
Pobre x	V1 x	+ x	Não x	6	2,1%
Rica ✓	*V1 ✓	- ✓	Sim ✓	0	0
Rica ✓	*V1 ✓	- ✓	Não x	0	0
Rica ✓	*V1 ✓	+ x	Sim ✓	0	0
Rica ✓	*V1 ✓	+ x	Não x	0	0
Rica ✓	V1 x	- ✓	Sim ✓	0	0
Rica ✓	V1 x	- ✓	Não x	0	0
Rica ✓	V1 x	+ x	Sim ✓	0	0
Rica ✓	V1 x	+ x	Não x	0	0
				287	100%

Fonte: elaboração própria.

Nessa nova tabela, por estarmos observando casos de terceira pessoa, não temos nenhuma ocorrência de morfologia rica. Logo, se há apenas casos de morfologia pobre, podemos perceber que esse fator se torna irrelevante para a análise, restando a combinação entre padrão linear *V1, gênero semântico e conexão discursiva ótima. Nesse caso, temos uma tabela idêntica à da combinação desses três fatores, aparecendo, então, nosso dado mais relevante até aqui: 281 ocorrências são explicadas por pelo menos um dos fatores citados, o que representa 97,9% dos casos do *corpus*; apenas 6 casos (2,1%) não têm nenhum fator licenciador atuando no fenômeno do sujeito nulo. O gráfico abaixo ajuda a ilustrar os dados da tabela.

Gráfico 11 – Ocorrência de sujeito nulo (%) por todas as variáveis analisadas sem 1ª e 2ª pessoa (N=287)



Fonte: elaboração própria.

Apesar de termos uma ótima projeção com a explicação de 97,9% dos casos, devemos dar atenção aos 6 casos que contrariam as expectativas e que não são adequados a nenhum fator licenciador de sujeito nulo.

4.5 Casos Especiais

Apenas 6 (entre 287) ocorrências de sujeito nulo não apresentaram nenhum dos fatores licenciadores expressos em Ayres (2021). Nesta seção, falaremos sobre cada uma dessas ocorrências na busca de explicações para a quebra de expectativa.

Conexão Discursiva Ótima

Os processos de retomadas anafóricas devem considerar os contextos que ambientam as orações. A conexão discursiva ótima evidencia a importância da estrutura informacional da sentença e, segundo Paredes Silva (2003), deve privilegiar a retomada a partir da manutenção da função de sujeito do referente e, também, da manutenção do sistema verbal, a partir da repetição dos aspectos verbais da sentença.

Entretanto, neste trabalho, adotamos o mesmo critério utilizado por Ayres (2021), que considera que a conexão discursiva ótima é mantida quando há permanência na função do sujeito, mesmo que haja mudança de tempo verbal, desde que não existam outros elementos que atrapalhem a identificação do sujeito no contexto comunicativo. Dessa forma, é possível considerar que os casos especiais em que há essa mudança de tempo verbal (e que foram marcados negativamente para conexão discursiva ótima inicialmente), na verdade, preservam a conexão ótima. São eles:

(CE1) O Mané respondeu: “Deixa de frescura, Chinoca! O guri é normal, um filho de gaúcho”. **Vivia** contando para o piá as façanhas dele ou supostas façanhas durante a tal Revolução Farroupilha. (Buenas Ideias – Linguiceiros de POA)

(CE2) No porão da casa do Ramos e da Catarina, lá na rua do Arvoredo, estavam enterrados pedaços de um corpo humano e a vítima foi logo identificada: **era** o açougueiro Carlos Clausner! (Buenas Ideias – Linguiceiros de POA)

(CE3) É esse Manoel Lobo, Dom Manuel Lobo, que é o cara que tinha sido mandado criar a Colônia do Sacramento em 1680. **Criou** e os espanhóis invadiram e destruíram a Colônia e prenderam o Manuel Lobo. (Buenas Ideias – Origem do Gaúcho)

Além disso, quando há manutenção da referência e manutenção do verbo (o que naturalmente facilita para o apontamento da referência), mesmo que o pronome seja alterado para fins de adequação ao diálogo, podemos, também, encarar como conexão discursiva ótima:

(CE4) Entrevistado: Eu tive que sair no carro da polícia, né, nesse jogo. A torcida queria me matar, cara.

Entrevistador: **Teve**? (Denílson Show – Diego Ribas)

Dos seis casos especiais, quatro podem ser reclassificados positivamente para conexão discursiva ótima, passando a ser, então, exemplos adequados à teoria inicial. Assim, temos uma nova tabela a partir dos quatro fatores.

Tabela 26 – Atualização de ocorrências de sujeito nulo por morfologia verbal, *V1, gênero semântico e conexão discursiva sem 1ª e 2ª pessoa (N=287)

MORFOLOGIA	*V1	GS	CDO	Ocorrências	%
Pobre x	*V1 ✓	- ✓	Sim ✓	24	8,4%
Pobre x	*V1 ✓	- ✓	Não x	53	18,5%
Pobre x	*V1 ✓	+ x	Sim ✓	14	4,9%
Pobre x	*V1 ✓	+ x	Não x	1	0,3%
Pobre x	V1 x	- ✓	Sim ✓	48	16,7%
Pobre x	V1 x	- ✓	Não x	93	32,4%
Pobre x	V1 x	+ x	Sim ✓	52	18,1%
Pobre x	V1 x	+ x	Não x	2	0,7%
Rica ✓	*V1 ✓	- ✓	Sim ✓	0	0
Rica ✓	*V1 ✓	- ✓	Não x	0	0
Rica ✓	*V1 ✓	+ x	Sim ✓	0	0
Rica ✓	*V1 ✓	+ x	Não x	0	0
Rica ✓	V1 x	- ✓	Sim ✓	0	0
Rica ✓	V1 x	- ✓	Não x	0	0
Rica ✓	V1 x	+ x	Sim ✓	0	0
Rica ✓	V1 x	+ x	Não x	0	0
				287	100%

Fonte: elaboração própria.

Com a nova tabela, chegamos ao total de 285 ocorrências de sujeito nulo explicadas pela teoria, atingindo a marca de 99,3% dos casos. Ainda restam 2 ocorrências desencaixadas com o esperado pela teoria, totalizando 0,7%. Seguimos na investigação de casos especiais.

Ambiguidade

Observemos o exemplo a seguir, retirado do *corpus*:

(CE5) Com o Luca, hoje, ele está muito na fase do futebol... Nessa fase, é muito direcionar ele, né? Mostrar para ele como é que é esse mundo. [pausa] **Faz** 15 anos [pausa]. Para ele ter entendimento da dificuldade que é e do sacrifício que tem que fazer para tentar ser jogador profissional. (Denílson Show – Kaká)

No exemplo, há a construção “Faz 15 anos”. No *podcast* analisado, essa construção vem seguida de uma pausa do entrevistado e tem uma nova pausa na sequência. Essas pausas e o contexto em que a frase está inserida geram a dúvida: o verbo “fazer” nesse caso é impessoal, referindo-se ao tempo decorrido, ou tem a referência em “Luca” e posteriormente no sujeito pleonástico “ele”? É importante contextualizar que Luca, filho do entrevistado, tem 14 anos e está prestes a completar 15, algo previamente explicado pelo entrevistado durante o *podcast*.

A ambiguidade gerada pela construção permite que o dado seja dispensado da análise caso seja considerado impessoal, ou seja assumido com a referência apontando para o pronome “ele”, que se refere a Luca e, nesse caso, teremos novamente conexão discursiva ótima como fator licenciador.

A nova tabela com os dados, então, fica da seguinte maneira considerando a segunda interpretação.

Tabela 27 – Nova atualização de ocorrências de sujeito nulo por morfologia verbal, *V1, gênero semântico e conexão discursiva sem 1ª e 2ª pessoa (N=287)

MORFOLOGIA	*V1	GS	CDO	Ocorrências	%
Pobre X	*V1 ✓	- ✓	Sim ✓	24	8,4%
Pobre X	*V1 ✓	- ✓	Não X	53	18,5%
Pobre X	*V1 ✓	+ X	Sim ✓	14	4,9%
Pobre X	*V1 ✓	+ X	Não X	1	0,3%
Pobre X	V1 X	- ✓	Sim ✓	48	16,7%
Pobre X	V1 X	- ✓	Não X	93	32,4%
Pobre X	V1 X	+ X	Sim ✓	53	18,5%
Pobre X	V1 X	+ X	Não X	1	0,3%
Rica ✓	*V1 ✓	- ✓	Sim ✓	0	0
Rica ✓	*V1 ✓	- ✓	Não X	0	0
Rica ✓	*V1 ✓	+ X	Sim ✓	0	0
Rica ✓	*V1 ✓	+ X	Não X	0	0
Rica ✓	V1 X	- ✓	Sim ✓	0	0
Rica ✓	V1 X	- ✓	Não X	0	0
Rica ✓	V1 X	+ X	Sim ✓	0	0
Rica ✓	V1 X	+ X	Não X	0	0
				287	100%

Fonte: elaboração própria.

Com a nova tabela, chegamos ao total de 286 ocorrências de sujeito nulo explicadas pela teoria, atingindo a marca de 99,7% dos casos. Ainda resta 1 caso de sujeito nulo operando com fatores não licenciadores, totalizando 0,3%. A investigação de casos especiais segue.

Sujeito nulo ativado pelo contexto

Observemos o último exemplo a seguir de um caso ainda não explicado pelos quatro fatores expostos por Ayres (2021).

(CE6) Entrevistado: Bom, Seleção Brasileira, o que eu tenho é só orgulho, o respeito. Foi o sonho que eu realizei, o sonho de ser campeão do mundo. Aquela taça ali, que é incrível, que você teve esse privilégio e venceu. Não alcancei, era um sonho que eu tinha também. Mas não fica frustração nenhuma, cara. Em alguns momentos, sim, foi difícil de eu lidar. Foi: “pô, como é que é possível, cara”.

Entrevistador: **Teve** duas oportunidades.

Entrevistado: É, sim, três no caso, né. Eu comecei lá com Parreira. (Denílson Show – Diego Ribas)

No exemplo, o entrevistado é Diego Ribas, ex-jogador de futebol, conversando com o também ex-jogador de futebol Denílson, âncora do *podcast*. No diálogo, o entrevistado expressa seus sentimentos em relação aos momentos em que teve a oportunidade de ser convocado para a seleção brasileira e o entrevistador interage perguntando a ele sobre a quantidade de oportunidades que ele teve. O verbo “teve” não possui morfologia exclusiva, possui gênero semântico aparente porque a pergunta é direcionada a um ser humano específico que deve responder de maneira subjetiva ao comando e, portanto, reconhece o seu próprio gênero semântico, não respeita ao Evite V1 já que é iniciado por verbo e não se encaixa em conexão discursiva ótima porque a pergunta é direcionada ao interlocutor, trazendo “tu” ou “você” como sujeito de “teve” sem que esse sujeito tenha sido mantido de alguma das orações anteriores para a analisada, exigindo uma resposta subjetiva. Dessa forma, o sujeito nulo é ativado por fatores externos que consideram o contexto comunicativo, tornando o sujeito óbvio.

A interação oportunizada pelo gênero *podcast* torna possível que, em alguns casos, haja a compreensão do sujeito a partir do assunto e dos elementos externos como apontamentos e olhares para que os interlocutores reconheçam do que se trata a sentença e a quem se refere o sujeito da oração em si. Essas interações a transcrição dos dados de fala não é capaz de captar, trazendo um estranhamento em um primeiro momento para a análise dos dados.

Uma segunda interpretação para esse caso específico é de que entendamos o verbo “teve” como impessoal e, nesse caso, não teria sujeito algum e seria dispensado da nossa análise. Nesse caso, teríamos 100% das ocorrências de sujeito nulo sendo explicados pelos quatro fatores. Na pior das hipóteses (em que não consideraríamos

esse verbo como impessoal), manteríamos a margem de 99,7% dos casos sendo explicados pelos fatores licenciadores, o que já é um resultado bastante satisfatório aos objetivos da pesquisa.

A teoria prevê que os quatro fatores licenciadores descritos em Ayres (2021) estejam envolvidos diretamente na totalidade de ocorrências de sujeitos nulos remanescentes em PB. Os resultados obtidos com a análise demonstram que as hipóteses, assim como em Ayres (2021), foram confirmadas em um *corpus* de gênero oral moderno e popular atingindo o objetivo do trabalho.

Neste capítulo, apresentamos a discussão a respeito dos resultados obtidos a partir da análise do *corpus*. Foi possível constatar que em pelo menos 99,7% das ocorrências do fenômeno havia pelo menos um dos quatro fatores licenciadores envolvido. Esse é um número absolutamente relevante para atestar o objetivo do trabalho, que é confirmar a hipótese de Ayres (2021) sobre o licenciamento de sujeitos nulos no PB a partir dos quatro fatores: morfologia verbal, padrão linear *V1, gênero semântico e conexão discursiva ótima.

5 CONSIDERAÇÕES FINAIS

A investigação sobre sujeitos nulos em PB permite que algumas questões sejam revistas e revisitadas. O comportamento do fenômeno na língua demonstra que é possível tentar identificar fatores que licenciem as ocorrências desse tipo de sujeito, já que, na atualidade, essa forma não é mais a preferencial dos falantes de PB. Por esse motivo, esse trabalho se propôs a verificar se os fatores descritos em Ayres (2021) para explicar a totalidade de ocorrências de sujeitos nulos em PB – morfologia verbal, padrão linear *V1, gênero semântico e conexão discursiva ótima – funciona também em um *corpus* diversificado a partir da análise de episódios de *podcasts*.

Dessa forma, podemos chegar às respostas das questões que nortearam essa pesquisa. A primeira questão buscava a reflexão sobre que fatores favorecem a utilização de sujeito nulo pelo falante em PB; a segunda buscava entender se os fatores de Ayres (2021) são suficientes para justificar todos os casos de sujeitos nulos encontrados em um *corpus* de língua falado do PB.

Para isso, observamos as análises sobre os fatores isolados e, também, sobre as suas combinações em duplas e trios, além da análise do conjunto com todos os quatro fatores unidos. A análise dos fatores isolados apontou que o gênero semântico é o fator mais relevante para explicar os casos remanescentes de sujeitos nulos em PB, chegando a 75,9% dos casos explicados. Entretanto, esses dados não se mostraram satisfatórios para que o fenômeno fosse plenamente explicado e, por esse motivo, foi crucial que a análise partisse para as combinações entre os fatores. Essas combinações mostraram resultados realmente promissores. Contudo, há duas análises concomitantes: a primeira considera 510 casos de sujeitos nulos obtidos a partir do *corpus* em que todas as pessoas do discurso são analisadas, e a segunda considera 287 casos do fenômeno quando há o envolvimento do fator gênero semântico, pois esse fator oportuniza variação em relação ao seu comportamento apenas na terceira pessoa do discurso. Então, a primeira e a segunda pessoa (por sempre apresentarem gênero semântico aparente) são dispensadas da análise.

Quando olhamos para os 510 casos de sujeitos nulos considerando todas as pessoas do discurso, a combinação que explica de maneira mais satisfatória os dados é a partir da união entre todos os quatro fatores, pois 96,7% das ocorrências são justificadas. Contudo, se observarmos os 287 casos que desconsideram primeira e segunda pessoa por causa do envolvimento do gênero semântico especificamente,

chegamos à explicação de 99,7% dos dados a partir da combinação do trio que envolve padrão linear *V1, conexão discursiva ótima e gênero semântico. Os poucos casos que não são explicados pelos fatores licenciadores nas duas análises (sobre 510 ocorrências e sobre 287 ocorrências), na verdade, podem ser explicados por fatores externos, podendo ser suprimidos graças ao contexto ativado pelo ambiente comunicativo.

É válido apontar que, ao analisarmos apenas casos de terceira pessoa pelo envolvimento de gênero semântico, naturalmente o fator morfologia torna-se irrelevante, já que só há morfologia rica em referentes de primeira pessoa do discurso. Os referentes de terceira pessoa analisados sempre apresentarão morfologia pobre, o que não é um fator licenciador de sujeitos nulos. Significa dizer que, no fim das contas, há uma economia na explicação dos casos do fenômeno: os quatro fatores descritos em Ayres (2021) podem, de fato, licenciar os sujeitos nulos em PB, mas no *corpus* analisado nesse trabalho, três fatores já são suficientes para dar conta dessas ocorrências.

Ayres (2021) analisou o fenômeno em um *corpus* de entrevistas semiestruturadas da região de Porto Alegre (RS) e chegou à marca de 30% de ocorrências de sujeito nulo; esse trabalho analisou o fenômeno em um *corpus* de gênero oral (*podcasts*) a nível nacional e também chegou à marca de 30% de ocorrências, mantendo a proporção da pesquisa anterior. Novos trabalhos, no futuro, poderão investigar o fenômeno em *corpus* de gêneros orais diversificados e poderão atestar se estamos diante de uma variação estável ou se as ocorrências de sujeitos nulos em PB serão ainda mais escassas.

A partir desse trabalho, esperamos contribuir de forma analítica e descritiva para as pesquisas sobre o tema em PB para que cada vez mais possamos entender o funcionamento da mudança da língua em direção ao preenchimento da posição de sujeito.

6 REFERÊNCIAS

- AYRES, M. R. *Aspectos condicionadores do objeto nulo e do pronome pleno em português brasileiro: uma análise da fala infantil*. Dissertação de mestrado: PUCRS, 2016.
- AYRES, M. R. *Contextos licenciadores de sujeitos nulos em português brasileiro*. Tese de doutorado: UFRGS, 2021.
- BECHARA, E. *Moderna gramática portuguesa*. 37.ed. revista, ampliada e atualizada conforme o novo Acordo Ortográfico. Rio de Janeiro: Nova Fronteira, 2009.
- CASTILHO, A. T. *Nova gramática do português brasileiro*. São Paulo: Contexto, 2010.
- CHOMSKY, N. *Lectures on Government and Binding*. Dordrecht: Foris, 1981.
- CINTRA, L. F. L. *Sobre Formas de Tratamento na língua portuguesa: ensaios*. Lisboa: Horizonte, 1972.
- CREUS, S; MENUZZI, S. O papel do gênero na alternância entre objeto nulo e pronome pleno em português brasileiro. *Revista da ABRALIN*, Florianópolis, v. 3, n. 1-2, 2004.
- CUNHA, C.; CINTRA, L. *Nova gramática do português contemporâneo*. 5.ed. Rio de Janeiro: Lexikon, 2008.
- DUARTE, M. E. L. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In: ROBERTS, I.; KATO, M. (Coords.). *Português brasileiro: uma viagem diacrônica*. Campinas: UNICAMP, 1993.
- DUARTE, M. E. L. *A perda do princípio “evite pronome” no português brasileiro*. Tese de doutoramento. Instituto de Estudos da Linguagem, Universidade Estadual de Campinas. Campinas, 1995.
- DUARTE, M. E. L. *O sujeito em peças de teatro (1833-1992): estudos diacrônicos*. São Paulo: Parábola Editorial, 2012.
- DUARTE, M. E. L.; KATO, M. A. The lack of expletives in Brazilian Portuguese. Comunicação apresentada no ENCONTRO DO GT DE TEORIA DA GRAMÁTICA, USP, 2013.
- GIVÓN, T. *English grammar: A function-based introduction*. John Benjamins Publishing, 1993.
- GIVÓN, T. *A compreensão da gramática*. São Paulo: Cortez \ Natal: EDUFRRN, 2012.
- GRAVINA, A. P. *A Natureza do Sujeito Nulo na Diacronia do PB: estudo de um corpus mineiro (1845 a 1950)*. Dissertação de mestrado, UNICAMP, 2008.

- GRAVINA, A. P. Diacronia e sujeito nulo no português brasileiro: um estudo comparativo. *Filol. Linguíst. Port.*, São Paulo, v. 16, n. spe, p. 199-231, dez. 2014.
- HOLMBERG, A., NAYUDU, A. & SHEEHAN, M. Three partial null-subject languages: a comparison of Brazilian Portuguese, Finnish, and Marathi, *Studia Linguistica*, pp. 59– 97, 2009.
- JÚNIOR, A. B.; PADILHA, L. D. O radiojornalismo como gênero discursivo aplicado ao conceito de podcast. *Anais do 18º Encontro Nacional de Pesquisadores em Jornalismo*, 2020. Disponível em <https://sbpjour.org.br/congresso/index.php/sbpjour/sbpjour2020/paper/viewFile/2722/1410>. Acesso em 10 de set. de 2023.
- KATO, M. A. Strong and weak pronominals in the null subject Parameter. *Probus* (Dordrecht), Berlin, v. 11, n.1, p. 1-37, 1999.
- KATO, M. A. The partial pro-drop nature and the restricted VS order in Brazilian Portuguese, in M. E. N. Kato, ed., *Brazilian Portuguese and the Null Subject Parameter*, Vervuert- Iberoamericana, Frankfurt, Germany, pp. 223–258, 2000.
- KATO, M. A evolução da noção de parâmetros. *Revista DELTA*, vol.18, no.2, 2002.
- KATO, M. A. Determinantes prosódicos em mudança sintática. *Abralin ao vivo*, 2020. [<https://www.youtube.com/watch?v=t3BLRPloZJI>]
- KATO, M. A.; DUARTE, M. E. L. Restrição na distribuição de sujeitos nulos no Português Brasileiro. *Veredas* (UFJF. Online), v. 18, p. 1-21, 2014.
- KATO, M. A.; DUARTE, M. E. L. Restrições na distribuição de sujeitos nulos no Português Brasileiro. *Revista Veredas*, vol. 18/1, 2014.
- KATO, M. A.; DUARTE, M. E. L. Pre-verbal position in BP: a reinterpretation of the “Avoid Pronoun Principle”. *Revista Diadorim*, Rio de Janeiro, vol. 20, 2018, p. 610-626.
- KATO, M. A; MARTINS, A. M. L; NUNES, J. Português brasileiro e português europeu: sintaxe comparada. São Paulo: Contexto, 2023. 384p.
- LOPES, C. R. dos S. “Vossa Mercê > você e Vuestra Merced > usted: o percurso evolutivo ibérico”. *Linguística - publicação da ALFAL*, vol. 14, 2003
- MACEDO, L. F. O. de. *Relação entre morfologia verbal e distribuição de sujeitos nulos e expressos em textos dissertativo-argumentativos*. Trabalho de Conclusão de Curso. Porto Alegre, 2020.
- MONTEIRO, J. L. Pronomes pessoais: subsídios para uma gramática do português do Brasil. Fortaleza: Edições UFC, 1994.
- OTHERO, G. A.; AYRES, M. R.; LAZZARI, M. G. A conexão discursiva e a manifestação de sujeito pronominal e nulo em português brasileiro. *Caderno de*

squibs: temas em estudos formais da linguagem, v. 4, p. 28-34, 2020. Disponível em <https://periodicos.unb.br/index.php/>

[cs/article/view/30470](https://periodicos.unb.br/index.php/cs/article/view/30470). Acesso em 5 de nov. de 2020.

OTHERO, G. A.; SPINELLI, A. C. Um tratamento unificado da omissão e da expressão de sujeitos e objetos diretos pronominais de 3ª pessoa em português brasileiro. *Caderno de Estudos Linguísticos*, vol. 61, n. 1, 2019.

PAREDES SILVA, V. L. Por trás das frequências. *Organon*, Porto Alegre, n. 18, p. 23-36, 1991.

PAREDES SILVA, V. L. Variação e funcionalidade no uso de pronomes de 2a. pessoa do singular no português carioca. *Revista de Estudos da Linguagem*, Belo Horizonte, v. 7, n.2, p. 121-138, 1998.

PAREDES SILVA, V. L. Motivações funcionais no uso do sujeito pronominal: uma análise em tempo real. In: PAIVA, M. da C.; DUARTE, M. E. L. (org.). *Mudança lingüística em tempo real*. 1ed. Rio de Janeiro: Contracapa/FAPERJ, v. 1, p. 97-114, 2003.

PERES, E. P. De “vossa mercê” a “cê”: os processos de uma mudança em curso. *Revista (Con)Textos Linguísticos*, vol. 1 n. 1, 2007 Disponível em <https://periodicos.ufes.br/contextos-linguisticos/issue/view/368>. Acesso em 10 de out. de 2023.

PERINI, M. A. Gramática do português brasileiro. São Paulo: Parábola, 2010

VASCONCELLOS, J. A. de. A redução dos paradigmas flexionais dos verbos e a perda do sujeito nulo no português brasileiro. *Revista Argumento*, v. 5 n. 9, 2003.

VILLARTA-NEDER, M. A.; FERREIRA, H. M. O podcast como gênero discursivo: oralidade e multisssemiose aquém e além da sala de aula. *Letras*, Santa Maria, Especial 2020, n. 01, p. 35-55.

WINK, C. O.; FINKENAUER, L.; OTHERO, G. de A. Quadro pronominal e colocação dos pronomes à luz de cinco gramáticas do português brasileiro. *Domínio de Lingu@gem*, vol. 6, nº 1, 2012.

ZILLES, A. M. S; MAYA, L. Z.; SILVA, K. Q. da. A concordância verbal com a primeira pessoa do plural em Panambi e Porto Alegre. *Organon*, 14, n. 28-29, p. 195-219, 2000.

7 APÊNDICE

Este capítulo pretende mostrar, de forma não exaustiva, um exemplo de transcrição de um dos podcasts escolhidos para análise e a classificação realizada com base nas hipóteses aqui investigadas.

7.1 EXEMPLO DE *PODCAST* TRANSCRITO

Segue abaixo exemplo de um pequeno trecho do *podcast* Denílson Show com Chico Garcia, em que o entrevistado era o jogador Diego Ribas.

CG: Denilson, só mais uma porque agora eu começo, né.

DE: Só até às 9.

CG: é, eu eu eu sempre brinco que eu ganho a vida fazendo perguntas, né. Vocês ganham a vida fazendo gol, eu ganho fazendo pergunta. Mudamos agora.

DR: Feliz aquele que sabe perguntar, cara.

CG: Aquele evento né que, se fosse nos dias de hoje, ele seria muito mais viralizado do que ele já viralizou, que é o jogo no Morumbi, gol do Santos, e vocês vão vibrar naquele escudo gigante lá - que agora, em jogos mais problemáticos, o São Paulo bota até segurança lá em volta. Como é que se deu aquilo, Diego? E quando você viu, você já tava pulando lá em cima ou vocês pensaram? Como é que é aquele evento?

DR: Aquele ali foi uma molecagem mesmo, né. Eu não faria de novo. Eu não faria de novo. Foi uma molecagem. A torcida do São Paulo maravilhosa decidiu, fui fazer um jogo lá. Porra, carinho, respeito... mas aquilo ali eu já programei, já na categoria de base, foi jogar lá semifinal do Campeonato Paulista, eu falei para meus amigos "vou fazer um gol, vou subir naquele símbolo" que nós jogamos no Morumbi, né. Mas eu não fiz gol esse jogo. Nós ganhamos, mas eu não fiz gol. Acabei que fui profissional e fiquei com aquela história "eu vou fazer um gol, vou lá zoar os cara, vou subir no símbolo" e aí sai o pênalti, eu vou, Rogério Ceni no gol, e aí eu falo "Robinho, ó: prepara que é agora, vamos para o símbolo, hein? Para com isso aí, para isso aí, ó, a torcida do Santos aí. Que bonito". Os cara gritando "para com isso" e eu na meia lua para bater o pênalti. Aí vou e "pum", bati, só que o Rogério Ceni com aquela mania que veio quase nos dois metros, o carrinho, o carrinho, pegou, aí o juiz volta, aí o Robinho fala: "falei! fala que vai subir lá, vai errar de novo! Vamos para a torcida do Santos!" Falei "não, vamos lá". Bati, ele ainda pegou. Gol, tocou na mão dele igual, eu fui direto. Então foi programado mesmo que lá subi e deu aquela confusão. Só que depois da confusão falei "caramba, que besteira que eu fiz, cara!"... Eu tive que sair no carro de polícia, né, nesse jogo. A torcida queria me matar, cara.