



UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
Faculdade de Ciências e Letras - Campus de Araraquara - SP

CAROLINE LAURIANO SILVA

**REQUISITOS DE APOIO DISCURSIVO EM PEÇAS DE
TEATRO DO SÉCULO XX - 1950 A 1970: um estudo
sociolinguístico**



ARARAQUARA – S.P.
2023

CAROLINE LAURIANO SILVA

**REQUISITOS DE APOIO DISCURSIVO EM PEÇAS DE
TEATRO DO SÉCULO XX - 1950 A 1970: um estudo
sociolinguístico**

Dissertação apresentada ao Programa de Pós-Graduação em Linguística e Língua Portuguesa, da Faculdade de Ciências e Letras – UNESP/Araraquara, como requisito para obtenção do título de Mestre em Linguística e Língua Portuguesa.

Linha de pesquisa: Análise Fonológica, Morfossintática, Semântica e Pragmática

Orientador: Profa. Dra. Rosane de Andrade Berlinck

Bolsa: CNPq

ARARAQUARA – S.P.
2023

S586r Silva, Caroline Lauriano
Requisitos de apoio discursivo em peças de teatro do século XX -
1950 a 1970: : um estudo sociolinguístico / Caroline Lauriano Silva. --
Araraquara, 2023
151 p.

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp),
Faculdade de Ciências e Letras, Araraquara
Orientadora: Rosane de Andrade Berlinck

1. Marcadores discursivos. 2. Requisitos de apoio discursivo. 3.
Peças teatrais. 4. Variação linguística. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de
Ciências e Letras, Araraquara. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

CAROLINE LAURIANO SILVA

REQUISITOS DE APOIO DISCURSIVO EM PEÇAS DE TEATRO DO SÉCULO XX - 1950 A 1970: um estudo sociolinguístico

Dissertação apresentada ao Programa de Pós-Graduação em Linguística e Língua Portuguesa, da Faculdade de Ciências e Letras – UNESP/Araraquara, como requisito para obtenção do título de Mestre em Linguística e Língua Portuguesa.

Linha de pesquisa: Análise Fonológica, Morfossintática, Semântica e Pragmática

Orientador: Profa. Dra. Rosane de Andrade Berlinck

Bolsa: CNPq

Data da defesa: 26/05/2023

MEMBROS COMPONENTES DA BANCA EXAMINADORA:

Presidente e Orientadora: Profa. Dra. Rosane de Andrade Berlinck
Universidade Estadual Paulista – UNESP/FCLAr

Membra Titular: Profa. Dra. Carla Regina Martins Paza
Universidade Federal de Santa Catarina – UFSC

Membro Titular: Prof. Dr. Sebastião Carlos Leite Gonçalves
Universidade Estadual Paulista – UNESP/Ibilce

Local: Universidade Estadual Paulista
Faculdade de Ciências e Letras
UNESP – Campus de Araraquara

Àqueles que me incentivaram a nunca desistir dos meus sonhos...

AGRADECIMENTOS

Nem todas as palavras do mundo seriam suficientes para agradecer a todos aqueles que, de algum modo, fizeram parte desse período de aprendizado e de desenvolvimento; entretanto, eu não poderia deixar de expressar a minha gratidão.

Primeiramente, agradeço a Deus por essa oportunidade e por ter me conduzido, me guardado e me fortalecido ao longo de todo o processo.

Além disso, deixo os meus agradecimentos:

À minha mãezinha, essa mulher batalhadora e sábia que carrega uma alegria contagiante. Obrigada por me inspirar a ser forte e corajosa, por me ensinar a agir com sabedoria e gentileza e por me incentivar a sonhar e a dançar em meio às tempestades da vida; também agradeço pelas palavras de encorajamento e pelo “colinho” nos momentos difíceis.

À minha irmã, Gabizinha, que, além de me encorajar e de me colocar para cima, sempre está ao meu lado. Obrigada pelo apoio, pela força e por todas aquelas conversas, que, mesmo à distância, fizeram toda diferença.

À minha querida orientadora, Rosane, que, com toda certeza, é uma grande inspiração para mim. Por todo o apoio e disponibilidade, por todos os direcionamentos e conhecimentos compartilhados, os quais transcendem as páginas desta dissertação, por toda sua gentileza e parceria, deixo o meu mais sincero agradecimento!

Aos meus familiares e amigos, que, constantemente, torcem por mim.

Aos colegas do *Núcleo de Pesquisas em Sociolinguística de Araraquara* (SoLAr), os quais, direta ou indiretamente, contribuíram para o desenvolvimento desta dissertação. Deixo um agradecimento especial a Milena, que, mesmo em meio ao doutorado e à correria do dia a dia, arranjou um tempinho para me auxiliar com o programa R.

Ao professor Sebastião e à professora Carla, por terem, bondosamente, aceitado fazer parte das bancas de qualificação e de defesa. Acrescento, ainda, um agradecimento ao professor Sebastião, por ter debatido o meu trabalho no SELin - 2022, contribuindo, imensamente, para o desenvolvimento deste estudo.

Ao Programa de Linguística e Língua Portuguesa da UNESP de Araraquara e a todos os docentes que contribuíram para a minha formação ao longo desse período.

O presente trabalho foi realizado com apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico - Brasil (CNPq).

JUIZ JOVEM	Antes... Não sei Excelência. É tão difícil colocar Esse tempo que pedis Antes, depois, agora... Cone do passado Cone da memória E a hora em que tudo se faz. Vede, Excelência... O tempo o que é? É o que demora!
JUIZ VELHO	É o que nos escapa.
JUIZ JOVEM	Isso é poesia.
JUIZ VELHO	O tempo... depende Se é visto pela física Ou pela metafísica.
JUIZ JOVEM	Na física tudo é coerente.
JUIZ VELHO	Mas alguém me disse Que o relógio da poesia Anda mais depressa E com mais maestria do que aquele da física.

Hilda Hilst (2000[1968], p. 218-219)

RESUMO

Este estudo verifica a ocorrência de *marcadores discursivos* (MD), mais especificamente daqueles que integram o subconjunto dos *requisitos de apoio discursivo* (RAD) – *compreendeu?, entendeu?, percebeu?, sabe?, viu? etc.* –, em peças de teatro brasileiras produzidas entre as décadas de 1950 e 1970 (século XX). O objetivo desta pesquisa é investigar o uso variável dos RAD originados dos verbos de cognição *compreender, entender, perceber, saber e ver*, verificando quais forças linguísticas e sociais motivam o uso desses elementos. Para tal fim, parte-se dos pressupostos da Sociolinguística Variacionista (WEINREICH; LABOV; HERZOG, 2006[1968]; LABOV, 2008[1972]), da Sociolinguística Histórica (ROMAINE, 1982; CONDE SILVESTRE, 2007) e do Sociofuncionalismo (BYBEE, 2010, 2016; TAVARES, 2013). Parte-se, também, dos estudos desenvolvidos por Schiffrin (1987, 2001, 2006) e da Gramática Textual-Interativa (RISSO; SILVA; URBANO, 2019), abordagens de estudo de MD adotadas neste trabalho. Selecionaram-se 39 peças teatrais para compor o *corpus* do trabalho; essas obras foram extraídas de acervos disponíveis virtualmente e de um acervo destinado a pesquisas (socio)linguísticas, organizado pelo *Núcleo de Pesquisas em Sociolinguística de Araraquara* (SoLAr), no âmbito do projeto *Língua EmCena*. Sendo assim, controlaram-se variáveis de natureza linguística (apresentação formal, posição e presença/ausência de respostas) e variáveis de natureza social (gênero/sexo, fase da vida, instrução, classe social, etnia, relação entre os personagens/interlocutores, ambientação e autoria das peças). Após a coleta e o processamento estatístico dos dados, constataram-se 319 ocorrências de RAD no *corpus* analisado, distribuídas da seguinte forma: *entender* (104 – 33%), *ver* (91 – 28%), *saber* (72 – 23%), *compreender* (42 – 13%) e *perceber* (10 – 3%). Utilizando o programa R (CORE TEAM, 2022), a análise foi realizada por meio de testes de qui-quadrado e do modelo de árvores de inferências condicionais. Ao controlar as variáveis de natureza linguística – apresentação formal, posição e presença/ausência de respostas –, observou-se que os RAD em análise, em sua maioria, possuem forma simples (262 – 82%), atuam em posição medial de turno (240 – 75%) e são caracterizados pela ausência de respostas após seu uso (300 – 94%), particularidades que correspondem aos traços típicos dos RAD, sinalizando que as formas analisadas já não operam como verbos plenos, mas encontram-se em diferentes estágios do processo de gramaticalização. Quanto às variáveis de natureza social, verificou-se que, apesar de a produção dos RAD ser mais influenciada pela percepção dos dramaturgos, usuários reais da língua, as macrocategorias sociais (gênero/sexo, fase da vida, instrução, classe social e etnia) e a ambientação demonstraram ser aspectos relevantes para a construção e representação dos personagens de peças teatrais.

Palavras-chave: marcadores discursivos; requisitos de apoio discursivo; peças teatrais; variação linguística.

ABSTRACT

This study verifies the occurrence of *discursive markers* (DM), more specifically those that integrate the subset of *discourse markers of support* (RAD) – *compreendeu?*, *entendeu?*, *percebeu?*, *sabe?*, *viu?* etc. –, in Brazilian theater plays produced between the 1950s and 1970s (20th century). The objective of this research is to investigate the variable use of RAD originated from the cognition verbs *compreender*, *entender*, *perceber*, *saber*, and *ver*, verifying which linguistic and social forces motivate the use of these elements. To this end, this research is based on the assumptions of Variationist Sociolinguistics (WEINREICH; LABOV; HERZOG, 2006[1968]; LABOV, 2008[1972]), Historical Sociolinguistics (ROMAINE, 1982; CONDE SILVESTRE, 2007), and Sociofunctionalism (BYBEE, 2010, 2016; TAVARES, 2013). It is also based on studies developed by Schiffrin (1987, 2001, 2006) and Textual-Interactive Grammar (RISSO; SILVA; URBANO, 2019), DM study approaches adopted in this work. We selected thirty-nine theater plays to compose our *corpus*; these plays were extracted from collections available virtually and from a collection intended for (socio)linguistic researches, organized by *Núcleo de Pesquisas em Sociolinguística de Araraquara (SoLAr)*, within the scope of *Língua EmCena* project. Thus, we controlled variables of linguistic nature (formal presentation, position, and presence/absence of answers) and variables of social nature (gender/sex, stage of life, education, social class, ethnicity, relationship between characters/interlocutors, setting, and authorship of the plays). After data collection and statistical processing, we verified 319 occurrences of RAD in the analyzed *corpus*, distributed as follows: *entender* (104 – 33%), *ver* (91 – 28%), *saber* (72 – 23%), *compreender* (42 – 13%), and *perceber* (10 – 3%). Analysis was performed in R programming (CORE TEAM, 2022), using chi-squared tests and the conditional inference trees model. When controlling variables of linguistic nature – formal presentation, position, and presence/absence of answers –, we observed that the RAD under analysis, for the most part, have a simple form (262 – 82%), act in a medial position of turn (240 – 75%), and are characterized by the absence of answers after their use (300 – 94%). These characteristics correspond to typical traits of RAD, signaling that the analyzed forms no longer operate as verbs, but they are at different stages of the grammaticalization process. As for variables of social nature, although the production of RAD is more influenced by the perception of the playwrights, real language users, social macrocategories (gender/sex, stage of life, education, social class, and ethnicity) and setting proved to be relevant aspects for the construction and representation of characters in theatrical plays.

Keywords: discourse markers; discourse markers of support; theater plays; linguistic variation.

LISTA DE QUADROS

Quadro 1	Os RAD considerados na análise	20
Quadro 2	Matrizes básicas propostas pela GTI	45
Quadro 3	Traços identificadores dos MD	62
Quadro 4	Núcleo-piloto de MD proposto pela GTI	62
Quadro 5	<i>O corpus</i>	66
Quadro 6	Variáveis e critérios propostos pelo projeto <i>Língua EmCena</i>	68
Quadro 7	Os RAD considerados na análise	70
Quadro 8	Variáveis independentes	72
Quadro 9	Critérios da variável fase da vida	78
Quadro 10	Sistematização das variáveis independentes	84
Quadro 11	Valores dos testes de qui-quadrado: variáveis linguísticas	88
Quadro 12	Formas compostas dos RAD	90
Quadro 13	Núcleo-piloto de MD proposto pela GTI	103
Quadro 14	Valores dos testes de qui-quadrado: variáveis extralinguísticas	104
Quadro 15	Critérios da variável fase da vida	107
Quadro 16	Critérios da variável instrução	110
Quadro 17	Critérios da variável classe social	112
Quadro 18	Categorização de personagens de classe 4	118
Quadro 19	Categorização de personagens não brancos	120
Quadro 20	Local de nascimento dos autores	124

LISTA DE TABELAS

Tabela 1	Os RAD mais frequentes no <i>corpus</i>	95
Tabela 2	Distribuição dos RAD de acordo com a apresentação formal	96
Tabela 3	Posição e apresentação formal dos RAD	98
Tabela 4	Distribuição dos RAD de acordo com a posição no turno	98
Tabela 5	Distribuição dos RAD de acordo com a presença/ausência de respostas	100
Tabela 6	Distribuição dos RAD de acordo com a fase da vida dos personagens	108
Tabela 7	Distribuição dos RAD de acordo com o nível de instrução dos personagens	111
Tabela 8	Distribuição dos RAD de acordo com a classe social dos personagens	113
Tabela 9	Distribuição dos RAD de acordo com a etnia dos personagens	115
Tabela 10	Distribuição dos RAD produzidos por personagens de classe 4	119
Tabela 11	Distribuição dos RAD de acordo com as relações entre os personagens/interlocutores: íntimas e não íntimas	122
Tabela 12	Distribuição dos RAD de acordo com as relações entre os personagens/interlocutores: assimétricas e simétricas	122
Tabela 13	Ambientação das peças de teatro	123
Tabela 14	Distribuição dos RAD de acordo com a ambientação das peças	123
Tabela 15	Distribuição dos RAD de acordo com o local de nascimento dos autores	124
Tabela 16	Distribuição dos RAD de acordo com a autoria	126
Tabela 17	Distribuição de personagens de acordo com a autoria	127

LISTA DE FIGURAS

Figura 1	Distribuição dos RAD no <i>corpus</i>	86
Figura 2	Apresentação formal dos RAD	95
Figura 3	Posição dos RAD no turno conversacional	97
Figura 4	Presença/ausência de respostas após o uso dos RAD	99
Figura 5	Diagrama arbóreo: variáveis linguísticas	101
Figura 6	Diagrama arbóreo: posição e presença/ ausência de respostas	102
Figura 7	Distribuição dos personagens de acordo com o gênero/sexo	105
Figura 8	Distribuição dos personagens de acordo com a fase da vida (1950 – 1963)	107
Figura 9	Distribuição dos personagens de acordo com a fase da vida (1964 – 1970)	108
Figura 10	Distribuição dos personagens de acordo com o nível de instrução	110
Figura 11	Distribuição dos personagens de acordo com a classe social	113
Figura 12	Distribuição dos personagens de acordo com a etnia	115
Figura 13	Diagrama arbóreo: macrocategorias sociais – gênero/sexo, fase da vida, instrução, classe social e etnia	117
Figura 14	Diagrama arbóreo: etnia e instrução	120
Figura 15	Diagrama arbóreo: ambientação e local de nascimento dos autores	125
Figura 16	Diagrama arbóreo: ambientação e autoria	132
Figura 17	Diagrama arbóreo: ambientação e variáveis de natureza linguística	133
Figura 18	Diagrama arbóreo: ambientação e perfil social dos personagens	134
Figura 19	Diagrama arbóreo: correlação entre os RAD e as variáveis independentes	136

LISTA DE ABREVIATURAS E SIGLAS

BAD	Busca de Aprovação Discursiva
GTI	Gramática Textual-Interativa
MD	Marcador Discursivo
NURC	Projeto da Norma Urbana Linguística Culta
RAD	Requisito de Apoio Discursivo
SoLAr	Núcleo de Pesquisas em Sociolinguística de Araraquara
TG	Teoria da Gramaticalização
TVML	Teoria da Variação e Mudança Linguística

SUMÁRIO

1 INTRODUÇÃO	16
2 FUNDAMENTAÇÃO TEÓRICA: SUBSÍDIOS PARA O ESTUDO DOS RAD	23
2.1 A Sociolinguística Variacionista: pressupostos teórico-metodológicos	23
2.2 Sociofuncionalismo	25
2.3 A Sociolinguística Histórica: breves considerações	27
2.4 O teatro brasileiro nas décadas de 1950, 1960 e 1970	29
2.4.1 O texto teatral como fonte de análise para investigações linguísticas	31
2.5 Marcadores discursivos: diferentes rótulos, diferentes abordagens	35
2.5.1 A abordagem de Schifffrin	37
2.5.2 A Gramática Textual-Interativa	41
2.5.3 Um diálogo entre abordagens: a multifuncionalidade dos MD em foco	46
2.6 Delimitação do objeto	48
2.6.1 Os requisitos de apoio discursivo	51
2.6.2 Gramaticalização dos RAD	54
2.7 Sintetizando...	64
3 METODOLOGIA	65
3.1 As etapas da pesquisa	65
3.2 O <i>corpus</i> : peças de teatro produzidas entre as décadas de 1950 e 1970	66
3.3 Coleta e tratamento dos dados	69
3.4 Variáveis dependente e independentes	69
3.4.1 Apresentação formal	72
3.4.2 Posição	74
3.4.3 Presença/ausência de respostas	75
3.4.4 Perfil social dos personagens	76
3.4.4.1 Gênero/sexo	77
3.4.4.2 Fase da vida	77
3.4.4.3 Instrução	79
3.4.4.4 Classe social	79
3.4.4.5 Etnia	79
3.4.5 Relação entre os personagens/interlocutores	81
3.4.6 Ambientação	82
3.4.7 Autoria	82
3.5 Sintetizando...	83

4 ANÁLISE E DISCUSSÃO DOS DADOS: UM OLHAR PARA OS RAD NAS PEÇAS TEATRAIS	85
4.1 Visão geral dos resultados	85
4.2 Variáveis de natureza linguística	88
4.2.1 Apresentação formal	89
4.2.2 Posição	96
4.2.3 Presença/ausência de respostas	98
4.2.4 Cruzamentos: variáveis de natureza linguística	100
4.3 Variáveis de natureza extralinguística	103
4.3.1 As macrocategorias sociais: o perfil dos personagens	104
4.3.1.1 Gênero/sexo	104
4.3.1.2 Fase da vida	106
4.3.1.3 Instrução	110
4.3.1.4 Classe social	112
4.3.1.5 Etnia	115
4.3.1.6 Correlacionando as macrocategorias sociais	117
4.3.2 Relação entre os personagens/interlocutores	121
4.3.3 Ambientação	122
4.3.4 Autoria	126
4.3.5 Cruzamentos	131
4.4 Sintetizando...	139
5 CONCLUSÃO	140
REFERÊNCIAS	145

1 INTRODUÇÃO

A espontaneidade é uma das principais características da conversação, visto que sua formulação ocorre concomitantemente à sua produção; por essa razão, é comum nos depararmos com hesitações, reformulações, interrupções e ajustes na fala espontânea. Entretanto, além dessas particularidades, os textos orais contam com certos recursos que auxiliam em sua estruturação e facilitam a comunicação, dentre os quais destacamos os marcadores discursivos (MD), objeto de análise desta pesquisa.

Os MD, de maneira geral, operam na organização do texto falado, amarrando as porções textuais e orientando os interlocutores no decorrer da interação (RISSO; SILVA; URBANO, 2019), e, assim como outros elementos linguísticos/discursivos, “são formas variáveis, sensíveis aos contextos sociocultural e regional” (FREITAG, 2008, p. 4). Posto isso, são diversos os elementos que compõem o grupo dos MD (*né?, olha, uhn uhn, bem, veja, é?, hein?, entendeu?, então, aí, daí, mas*), os quais provêm das mais variadas classes gramaticais, constituindo, assim, um conjunto bastante amplo e heterogêneo. Dada a relevância do papel executado por esses elementos no discurso, os estudos referentes aos MD têm se multiplicado ao longo do tempo.

Devido à amplitude e à heterogeneidade desse conjunto, decidimos investigar um dos subconjuntos de MD: o de *requisitos de apoio discursivo* (RAD). De acordo com Valle (2014), que se fundamenta em estudos diversos acerca dos MD, os RAD atuam tanto no plano textual quanto no plano interacional, desempenhando diversas funções, que, por vezes, se sobrepõem, porque ambos os planos estão inter-relacionados. Silva e Macedo (1996), por seu turno, destacam que essas unidades são usadas para testar a atenção do interlocutor ou para obter a sua aquiescência. Os trechos abaixo, extraídos das peças teatrais selecionadas para a análise deste estudo, exemplificam o uso de alguns RAD:

(1)

CAFUNÉ - Pode se boba pra você! Cê num precisa de mandinga, **né?** Já é o melhor do campo desde hoje, **né?** (*Chapetuba Futebol Clube* – Oduvaldo Vianna Filho, 1959).

(2)

VITOR - Eu tenho só que acabar de arrumar as coisas e quero uma companhia pra ficar comigo até a última hora. Só isso, te pago mais 20, **tá?** Daí fica arredondado pra cem. (*O assalto* – José Vicente de Paula, 1969).

(3)

RICO – Pô, cara, você também quer saber de tudo, **hein?** (*Cordélia Brasil* - Antônio Bivar, 1967).

Por meio desses excertos, percebe-se que, não obstante as unidades que constituem o subconjunto dos RAD possuam certas semelhanças, há, ainda, discrepâncias quanto a outros critérios, como a classe gramatical de origem. A partir disso, objetivando reunir um grupo coerente para nossa análise, selecionamos os RAD derivados dos verbos de cognição *compreender*, *entender*, *perceber*, *saber* e *ver*, os quais são tratados como variantes, uma vez que se assemelham quanto à natureza conceptual, à classe gramatical de origem e partilham contextos de uso, atuando no domínio da requisição de apoio discursivo (VALLE, 2014).

Apesar de *ver* ser um verbo de percepção visual, seu sentido desliza para a cognição, tornando-o semelhante aos verbos cognitivos. Rost Snichelotto (2009), que estuda MD derivados dos verbos de percepção *ver* e *olhar*, defende que, além de terem um sentido mais concreto, relacionado à percepção visual, esses verbos adquirem sentidos mais abstratos, ligados à cognição. No caso de *ver*, especificamente, a autora conclui que,

Em sua trajetória de mudança semântica, *ver* perde parte do significado relativo à percepção visual (*perceber com os olhos*) e passa a veicular a acepção *de saber*; *perceber com a razão*, e depois, *compreender*. Observa-se a perda parcial do sentido de base da forma sob transformação, desenvolvendo-se novos sentidos relacionados uns com os outros, num movimento que vai de um significado-fonte, mais concreto, para sentidos derivados, mais abstratos, preservado o sentido original de *captar algo*. (ROST SNICHELOTTO, 2009, p. 54, grifos da autora)

Os trechos abaixo ilustram o uso de *ver* como verbo pleno, cujo significado está relacionado à percepção visual, e como RAD, que é usado de modo abstratizado, associado à cognição.

(4)

CAFUNÉ - [...] Ninguém qué que eu tire a barba, né, Bila? Engraçado. Riram de mim, **cê viu?** Te abraçaram também? O que é que diz aí, hein? Fiquei com vergonha de perguntá... (*Chapetuba Futebol Clube* – Oduvaldo Vianna Filho, 1959).

(5)

DURVAL – [...] Olha... não gosto de escrevê carta, **viu?** Confesso mesmo... (*Chapetuba Futebol Clube* – Oduvaldo Vianna Filho, 1959).

No exemplo (4), percebemos que a expressão *[vo]cê viu?* faz referência à ação mencionada anteriormente (“riram de mim”); logo, o personagem deseja saber se o amigo viu, em sentido concreto, certas pessoas rindo dele. Por outro lado, no exemplo (5), notamos que o verbo desempenha papel de RAD e assume um sentido mais abstrato, podendo ser substituído, sem prejuízo de sentido proposicional, por *entendeu?*, *sabe?*, *compreendeu?* ou *percebeu?*, que são formas intercambiáveis nesse contexto:

- i) Olha... não gosto de escrevê carta, **entendeu?** Confesso mesmo...
- ii) Olha... não gosto de escrevê carta, **sabe?** Confesso mesmo...
- iii) Olha... não gosto de escrevê carta, **compreendeu?** Confesso mesmo...
- iv) Olha... não gosto de escrevê carta, **percebeu?** Confesso mesmo...

Por serem fenômenos típicos da fala, presentes em situações comunicativas e dialógicas, o estudo dos MD, e, particularmente, dos RAD, em sincronias passadas torna-se desafiador. A invenção de ferramentas que permitem gravar a língua falada e registrá-la é, ainda, recente, o que significa que, para investigar os usos linguísticos de épocas mais distantes, o pesquisador deve recorrer a registros escritos, os quais, nem sempre, são fidedignos quanto à realidade oral de determinado período. A probabilidade de encontrarmos MD em fontes escritas que seguem, rigidamente, os padrões normativos da época é quase inexistente; daí a necessidade de selecionar fontes escritas que sejam representativas da fala, como é caso das peças teatrais.

Os dados analisados provêm, então, de peças de teatro produzidas por autores brasileiros ao longo do século XX, mais especificamente entre as décadas de 1950 e 1970. A escolha de peças teatrais como fonte de análise e de extração de dados se justifica pelo fato de esse gênero textual se estruturar a partir do diálogo, que pode ser tomado como uma representação do texto falado. Ademais, o diálogo exerce um papel importante no processo de criação dos personagens, pois, por meio dele, é possível identificar as características linguísticas e sociais desses personagens. As peças constituem uma rica fonte de dados (socio)linguísticos, sobretudo de períodos para os quais não se tem acesso a dados de fala (BERLINCK; BARBOSA; MARINE, 2008; BERLINCK, 2020; ALMEIDA; PINTO, 2020; BERLINCK; BRANDÃO, 2021). Somado a isso, os MD são elementos que se manifestam, majoritariamente, no discurso oral, o que torna ainda mais justificável o uso de textos que são representativos da fala e da interação entre interlocutores/personagens. Quanto ao recorte temporal proposto, trata-se de um período historicamente relevante por dois motivos: primeiramente, devido às modificações no contexto sócio-histórico brasileiro e à influência do

meio social sobre os usos que os membros da sociedade fazem da língua – cumpre destacar que língua e sociedade são indissociáveis e que a língua muda porque a sociedade muda (FARACO, 2006) – e, em segundo lugar, por se tratar de um período profícuo no que tange à produção de textos dramáticos.

Este trabalho, portanto, tem como objetivo geral investigar o uso variável dos RAD originados dos verbos *compreender*, *entender*, *perceber*, *saber* e *ver*, tendo como fonte de análise peças de teatro brasileiras produzidas entre as décadas de 1950 e 1970. À guisa de exemplo, nos trechos abaixo, destacamos as ocorrências de cada um desses RAD:

(6)

TUNINHO - E não saiu coisa nenhuma. Está em casa, **compreendeu?** Vai avisar a teu patrão que eu vou me plantar aqui e que não saio nem a tiro! (*A Falecida* - Nelson Rodrigues, 1953).

(7)

HOMEM - [...] Toda quarta-feira eu tiro feriado e faço uns passeios “digestivos” só pra “espairecer” o espírito sem “finalidades” mais “profundas”, **tá entendendo?** Eu sou um cara de classe, Mariazinha. [...] (*Fala baixo, senão eu grito* - Leilah Assumpção, 1969).

(8)

AURORA - Não há drama. Eu sou assim, de veneta, **percebeu?** Quando cismo com um camarada, já sabe: topo qualquer parada. [...] (*Os sete gatinhos* - Nelson Rodrigues, 1958).

(9)

GERVÁSIO - Isso é uma injustiça. Pouca gente pode atirar a primeira pedra. Até me revolta, **sabe?** Mas vocês não vão entrar com o mandado de segurança? Vocês querem ir no mole? [...] (*Em moeda corrente do país* - Abílio Pereira de Almeida, 1960).

(10)

BENEDITO - Tratem de calar a boquinha, **viu?** Se essa história se espalha! Os dois valentões aqui, ajoelhados, pedindo misericórdia! (*A pena e a lei* - Ariano Suassuna, 1959).

Visto que os RAD podem apresentar variação em sua forma, os dados são controlados a partir da fonte verbal desses elementos, de modo que todas as variantes sejam incluídas na análise, conforme exposto no quadro 1.

Quadro 1 - Os RAD considerados na análise

RAD	Fontes verbais	Variantes
	Compreender	Exemplos: <i>compreendeu? compreende? compreendestes?</i>
	Entender	Exemplos: <i>tá entendendo? entendeu? entende?</i>
	Perceber	Exemplos: <i>percebeu? percebeste? tá percebendo?</i>
	Saber	Exemplos: <i>sabe? sabia? tá sabendo?</i>
	Ver	Exemplos: <i>viu?</i> ¹

Fonte: própria

Para atingir o propósito geral, estabelecemos os seguintes objetivos específicos: i) verificar quais forças linguísticas e sociais estão correlacionadas ao uso dos RAD em estudo; e ii) discorrer sobre o processo de gramaticalização desses elementos para que possamos compreender de que modo essas formas passaram de verbos plenos a RAD. Nosso intuito, então, não é recuperar a trajetória de mudança de cada um desses RAD, mas propor um panorama geral do processo de gramaticalização das formas em estudo. Portanto, considera-se que este trabalho contribuirá com os estudos sócio-históricos, que destacam a relevância de materiais escritos para as investigações linguísticas, bem como com discussões acerca dos MD no âmbito da Teoria da Variação e Mudança Linguística (TVML). Nossa hipótese é que o uso dos RAD esteja relacionado a fatores linguísticos, como a posição dos RAD no discurso, e extralinguísticos, tais como idade, gênero/sexo, classe social etc. Apesar de os resultados de trabalhos anteriores (SILVA; MACEDO, 1996; VALLE, 2001, 2014) terem revelado que as macrocategorias sociais exercem pouca influência no uso dos RAD, supomos que – por estarmos lidando com textos dramáticos, diferentemente dos referidos estudos, que analisaram dados de fala – há relação entre os RAD e fatores como gênero/sexo, etnia, escolaridade, idade e classe social. Consideramos que personagens com determinados perfis sociais usarão RAD específicos.

Em vista disso, a temática deste trabalho se mostra relevante pelas seguintes razões: i) todos os falantes, com maior ou menor frequência, fazem uso dos MD, que são essenciais para o planejamento e para a organização dos textos falados; ii) ainda que as peças teatrais tragam uma representação da língua em uso, seus autores, ao criarem os personagens, atribuem a eles determinados “modos de falar”, dando aos diálogos características de uma conversação natural; iii) os MD são, frequentemente, rotulados como vícios de linguagem,

¹ Não encontramos variação dos RAD provenientes da fonte verbal *ver* em nosso *corpus*.

tornando-se alvo de estigma social e de preconceito linguístico, posições que desconsideram a relevância desses elementos para a estruturação dos textos falados (FREITAG, 2007); e iv) o tratamento variável a fenômenos discursivos, tais como os RAD, é uma prática pouco recorrente em trabalhos que versam sobre variação linguística. Somado a isso, Valle (2014), que também investiga RAD derivados de verbos de cognição², declara que são poucos os trabalhos que se dedicam à investigação de marcadores dessa origem, o que torna o tema ainda mais relevante e pertinente para os estudos dos MD e, especialmente, dos RAD. Em síntese, os MD são elementos passíveis de múltiplas e variadas análises, constituindo um amplo campo ainda pouco explorado, particularmente sob a perspectiva da TVML; logo, a presente pesquisa se dedica à investigação desse fenômeno, de modo a contribuir com estudos prévios e futuros.

Além desta introdução, na qual apresentamos a proposta, o tema, o problema da pesquisa, a justificativa, os objetivos geral e específicos e a hipótese levantada, esta dissertação é estruturada em alguns capítulos. O segundo capítulo é destinado à discussão dos subsídios teóricos; dessa forma, dissertamos sobre os pressupostos teórico-metodológicos da Sociolinguística Variacionista, do Sociofuncionalismo e da Sociolinguística Histórica, apresentamos as principais características do teatro brasileiro entre as décadas de 1950 e 1970 e, com base em estudos prévios, destacamos a relevância do gênero dramático para as investigações (socio)linguísticas. Em seguida, direcionamos o foco das discussões para os MD, fazendo uma breve contextualização do fenômeno. Feito isso, nos debruçamos sobre as duas perspectivas que norteiam este trabalho, a saber, Schifffrin (1987, 2001, 2006) e Risso, Silva e Urbano (2019), e estabelecemos um diálogo entre ambas as abordagens. Adiante, explicamos como se deu a delimitação do objeto, justificando a escolha dos RAD sob análise e esmiuçando a definição desses elementos. Para encerrar o capítulo, discorremos sobre o processo de gramaticalização dos RAD com o intuito de compreender o percurso de mudança desses elementos. No terceiro capítulo, detalhamos a metodologia da pesquisa: apresentamos o *corpus* do trabalho, indicando quais foram as peças teatrais selecionadas, bem como seus autores e datas de produção/publicação; elucidamos os procedimentos adotados para a realização da coleta e do tratamento dos dados; e especificamos a variável dependente e as variáveis independentes a serem controladas. No quarto capítulo, realizamos a análise e a discussão dos resultados obtidos. Por fim, sintetizamos as conclusões, as contribuições da

² Valle (2014) investiga as seguintes formas: *sabe?*, *sabes?*, *entende?*, *entendeu?*, *entendesse?*, *tá entendendo?* e *tás entendendo?*.

dissertação e os encaminhamentos futuros. Ao final de cada capítulo, apresentamos uma síntese, de modo a condensar os aspectos centrais das discussões desenvolvidas.

2 FUNDAMENTAÇÃO TEÓRICA: SUBSÍDIOS PARA O ESTUDO DOS RAD

Neste capítulo, nos dedicamos à revisão do arcabouço teórico. Primeiramente, discorreremos acerca da Sociolinguística Variacionista (WEINREICH; LABOV; HERZOG, 2006[1968]; LABOV, 2008[1972]; TARALLO, 1997; MOLLICA; BRAGA, 2004; BORTONI-RICARDO, 2014), assumindo a língua como heterogênea, variável e mutável. Nos atemos, também, ao Sociofuncionalismo (BYBEE, 2010, 2016; TAVARES, 2013), que articula as perspectivas da Sociolinguística Variacionista e do Funcionalismo Linguístico, tomando como objeto de estudo a língua em situações reais de uso e de interação entre os falantes. Somado a isso, nos debruçamos sobre os pressupostos teórico-metodológicos da Sociolinguística Histórica (ROMAINE, 1982; CONDE SILVESTRE, 2007; ROSA, 2015), cujo objetivo é investigar e descrever fenômenos linguísticos de épocas anteriores por meio de fontes escritas, explicando os processos de variação e mudança linguísticas no contexto sócio-histórico. Finalmente, dissertamos a respeito do teatro brasileiro (PRADO, 1988; CEBULSKI, 2013), dos principais fatos históricos que marcaram o período de 1950 a 1970 (GOMES, 2013; KLEIN; LUNA, 2014) e da relevância do texto dramático para os estudos linguísticos (BERLINCK; BARBOSA; MARINE, 2008; BERLINCK, 2020; ALMEIDA; PINTO, 2020; BERLINCK; BRANDÃO, 2021).

Mais adiante, enfocamos o objeto de estudo deste trabalho, os MD (MARCUSCHI, 1986; SCHIFFRIN, 1987, 2001, 2006; FRASER, 1996, 1999, 2006, 2009; SILVA; MACEDO, 1996; URBANO, 1999b, 2019; VALLE, 2001, 2014; BLAKEMORE, 2002; PENHAVAL, 2005, 2010, 2012; FISCHER, 2006; FREITAG, 2007; GUERRA, 2007; FREITAG; SILVA; EVANGELISTA, 2017; RISSO, 2019; RISSO; SILVA; URBANO, 2019). Realizamos uma breve sistematização das abordagens de Schiffrin (1987, 2001, 2006) e de Risso, Silva e Urbano (2019), nas quais este estudo se fundamenta, e estabelecemos um diálogo entre ambas as perspectivas. Na sequência, justificamos a delimitação do objeto de estudo e focalizamos os RAD, expondo suas propriedades e particularidades à luz de estudos prévios. Por fim, abordamos o processo de gramaticalização desses elementos para melhor compreender de que modo eles passaram de verbos a MD.

2.1 A Sociolinguística Variacionista: pressupostos teórico-metodológicos

Opondo-se às teorias dos neogramáticos, dos estruturalistas e dos gerativistas e refutando a concepção da homogeneidade da língua, Weinreich, Labov e Herzog (2006[1968]) propõem alguns princípios para o estudo da mudança linguística. Os linguistas

admitem que, antes de elaborar teorias a respeito da mudança, é preciso enxergar a língua – sincrônica e diacronicamente – como “um objeto constituído de heterogeneidade ordenada” (WEINREICH; LABOV; HERZOG, 2006[1968], p. 35). Nessa direção, os autores postulam cinco problemas empíricos que devem ser considerados em uma pesquisa acerca da mudança linguística: i) **problema da restrição**: faz referência ao conjunto de mudanças possíveis e às condições para que elas ocorram, isto é, buscam-se princípios universais que possibilitem prever os rumos da mudança; ii) **problema do encaixamento**: se refere a como um fenômeno linguístico em processo de variação ou mudança é encaixado nas estruturas linguística e social; iii) **problema da transição**: diz respeito à transmissão e à incrementação de novas formas, buscando compreender de que modo os elementos em processo de variação e mudança se propagam e mudam de estágio; iv) **problema da avaliação**: está associado à atitude do falante em relação às formas em variação e mudança linguísticas; e v) **problema da implementação**: corresponde à investigação dos fatores responsáveis pela implementação da mudança em determinados períodos, lugares e contextos linguísticos (WEINREICH; LABOV; HERZOG, 2006[1968]; COELHO *et al.*, 2019).

Dos cinco problemas postulados pelos linguistas, o presente trabalho lida apenas com dois: o *problema da restrição*, pois avaliamos os determinantes, linguísticos e extralinguísticos, da escolha de um RAD em detrimento de outros, e o *problema do encaixamento*, porque propomos explicar de que modo os RAD estão inseridos na matriz social do período histórico em investigação; todavia, não iremos tratar desse problema na matriz linguística, tarefa que demandaria um olhar à influência dos RAD sobre outros fenômenos gramaticais, distanciando-se do propósito deste trabalho.

Haja vista tais considerações, a Sociolinguística Variacionista, também conhecida como Teoria da Variação e Mudança Linguística (TVML), Sociolinguística Quantitativa ou Laboviana, dedica-se à investigação da variação e da mudança linguísticas, correlacionando esses processos a fatores internos e externos à língua. Para essa abordagem, a língua é social e não pode ser estudada de maneira autônoma, à parte dos contextos situacionais, pois é impossível negar a relação entre língua e sociedade. A língua, então, é heterogênea, variável e sujeita a mudanças ao longo do tempo. Essas características são inerentes à língua e não a transformam em um sistema desorganizado; pelo contrário, “aparentemente caótica e aleatória, a face heterogênea imanente da língua é regular, sistemática e previsível, porque os usos são controlados por variáveis estruturais e sociais” (MOLLICA; BRAGA, 2004, p. 27). Em síntese, a heterogeneidade da língua não compromete a sua sistematicidade (TARALLO, 1997; ALKMIM, 2001; CAMACHO, 2001).

Em um primeiro momento, os estudos sociolinguísticos (LABOV 2008[1972]; entre outros) se dedicaram apenas a fenômenos fonético-fonológicos; entretanto, a consolidação da Sociolinguística abriu espaço para que outros níveis (sintático, morfológico, discursivo etc.) passassem a ser considerados nas análises variacionistas (LABOV, 1978; LAVANDERA, 1978). Dessa maneira, é possível investigar unidades discursivas à luz da Sociolinguística.

A esse respeito, Görski e Valle (2016) explicam que nem sempre as correlações entre variáveis discursivas e fatores sociais serão estatisticamente relevantes. Em conformidade com as autoras, isso ocorre porque, quanto mais alto o nível linguístico de análise, mais influentes se tornam os aspectos semântico-pragmáticos sobre as variantes. Consequentemente, lidar com variáveis discursivas é uma tarefa bastante complexa; por essa razão, as autoras sugerem que, “em vez de se partir de uma certa forma, é preciso partir dos contextos discursivos em que a forma é usada, identificando suas funções discursivas e, então, as formas que estão disponíveis para desempenhar tais funções” (GÖRSKI; VALLE, 2016, p. 83). Por conseguinte, as pesquisadoras enfatizam alguns passos para lidar com unidades discursivas, quais sejam: i) a escolha do fenômeno discursivo; ii) as abordagens teóricas que irão nortear a pesquisa; iii) a elaboração do envelope de variação; e iv) o estabelecimento de critérios unificadores. Esses passos, segundo elas, podem ser utilizados para orientar e auxiliar os pesquisadores no tratamento de variáveis discursivas.

2.2 Sociofuncionalismo

Diversos estudos, que conciliam teorias sociolinguísticas e funcionalistas, foram desenvolvidos no âmbito do Programa de Estudos sobre o Uso da Língua (PEUL/RJ), de onde surgiu o termo *sociofuncionalismo* (NEVES, 1999; TAVARES, 2013; GONÇALVES, 2021). De acordo com Bybee (2016, p. 303), “a teoria baseada no uso se desenvolveu diretamente do funcionalismo norte-americano, praticado por muitas décadas (Noonan, 1998), e, em certo sentido, é apenas um novo nome para ele”. Assim sendo, articulando os pressupostos teórico-metodológicos da Sociolinguística Variacionista e do Funcionalismo Linguístico, mais particularmente da vertente norte-americana, a interface sociofuncionalista dedica-se à investigação da língua em uso.

A Sociolinguística ocupa-se do estudo da variação e da mudança linguísticas, assumindo a língua como sistema dinâmico e heterogêneo. O Funcionalismo, por seu turno, tem como objetivo explicar a língua a partir de contextos linguísticos e de contextos de uso; essa perspectiva concebe a língua como um mecanismo de interação social e admite que há uma relação intrínseca entre língua e sociedade. Ademais, os funcionalistas buscam analisar

dados reais da fala, deixando de trabalhar com textos inventados para um propósito específico. Em uma síntese das propostas do Funcionalismo, Cunha (2011, p. 158) pontua o seguinte: “a língua desempenha funções que são externas ao sistema linguístico em si” e “as funções externas influenciam a organização interna do sistema linguístico”. Bybee (2010, 2016), nessa mesma direção, já declarava que é por meio dos usos da língua que as estruturas linguísticas emergem e “o contexto social e cultural em que a língua é usada teria um impacto sobre as estruturas que são criadas” (BYBEE, 2016, p. 317). Desse modo, o falante adapta seus usos linguísticos de acordo com as mais variadas situações comunicativas; por esse motivo, em consonância com Tavares (2013), o uso é responsável tanto pela conservação quanto pela modificação da gramática. Nas palavras da autora,

No uso diário da língua, temos, por um lado, construções gramaticais que são repetidas, reforçando-se assim sua regularidade e seu caráter fixo, sistemático – o uso conserva a gramática; por outro lado, tais construções podem ser rearranjadas e remodeladas de modos diferentes, dando origem a construções inovadoras – o uso modifica a gramática. (TAVARES, 2013, p. 31)

A partir dessas constatações, Tavares (2013) explica que a gramática é suscetível a mudanças, uma vez que está sujeita ao uso que os falantes fazem dela no cotidiano e, conseqüentemente, às adaptações realizadas ao longo das interações diárias. Tais adaptações, segundo ela, ocorrem devido às diferentes experiências, particulares e individuais, que falantes e ouvintes têm com a língua e são necessárias para que haja compreensão mútua entre eles; desse processo, então, decorre a mudança, dado que

[...] adaptações feitas durante a interação, como tentativa de obtenção de êxito no processo de troca verbal, podem ocasionar o surgimento de novas estratégias para a constituição do discurso, que, se frequentemente repetidas, rotinizam-se, tornando-se parte da gramática da língua. Tornam-se, pois, construções gramaticais. Esse movimento de rotinização gramatical é denominado **gramaticalização**, caracterizado como o processo de regularização gradual pelo qual estratégias retóricas envolvendo itens lexicais e/ou itens gramaticais, inicialmente criativas e expressivas, tornam-se habituais por terem sido utilizadas recorrentemente em determinado tipo de contexto comunicativo. (TAVARES, 2013, p. 32, grifo nosso)

A gramaticalização, portanto, está no cerne das discussões sociofuncionalistas e é caracterizada como o processo pelo qual itens lexicais tornam-se gramaticais, podendo, mesmo depois de gramaticalizados, desenvolver outras funções gramaticais (HOPPER; TRAUGOTT, 1993; GONÇALVES, 2021). Esse processo, de acordo com Martelotta, Votre e

Cezario (1996), revela o caráter não estático da gramática e a busca por novas expressões, caracterizando, assim, o constante movimento e as eventuais mudanças da língua. Adiante, discorreremos mais detidamente sobre esse processo, com um olhar para a mudança dos RAD em estudo.

Não obstante a correlação entre a Sociolinguística e o Funcionalismo tenha causado questionamentos, muitos estudiosos defendem que a articulação entre essas abordagens é possível, apresentando alguns aspectos convergentes entre ambas as vertentes, por exemplo: i) a análise da língua conforme utilizada pelos falantes; ii) a visão de língua como heterogênea, variável e mutável; iii) as relações entre língua e sociedade; iv) a variabilidade como fator inerente à língua; v) maior ênfase na mudança, que é vista como contínua e gradual etc. (TAVARES, 2013; GONÇALVES, 2021). Em suma, as diferenças entre essas vertentes não impossibilitam a articulação teórica entre elas.

A abordagem sociofuncionalista, então, é a melhor opção para tratar de elementos discursivos, posto que essa abordagem “toma como objeto diferentes camadas ou variantes que partilham e/ou disputam determinada função, realizando o controle de grupos de fatores linguísticos, estilísticos e sociais passíveis de influenciar a escolha de uma delas pelos falantes” (TAVARES; GÖRSKI, 2012 apud VALLE, 2014, p. 177). Sendo assim, partimos desses pressupostos a fim de garantir um melhor tratamento e um controle mais apurado dos RAD e dos fatores considerados na análise.

2.3 A Sociolinguística Histórica: breves considerações

Agregando os princípios teórico-metodológicos da Sociolinguística Variacionista e da Linguística Histórica, a Sociolinguística Histórica constitui um campo híbrido, que interage com variadas áreas do conhecimento, e tem como objetivo descrever processos de variação e mudança linguísticas em diferentes sincronias (ROSA, 2015).

Nascida, aproximadamente, no final do século XVIII, a Linguística Histórica dedica-se ao estudo das mudanças que ocorrem nas línguas ao longo do tempo, investigando suas transformações e seu desenvolvimento. A teoria variacionista, de modo similar,

[...] assume a heterogeneidade sincrônica das línguas como sistemática e primordial, pressupõe o enraizamento da questão histórica nessa heterogeneidade, defende a necessidade de correlacionar língua e contexto social e busca sustentar suas hipóteses em amplos levantamentos de dados empíricos da comunidade de fala. (FARACO, 2006, p. 112)

Romaine (1982), pesquisando fenômenos linguísticos em registros escritos de outras épocas³, percebeu indícios de processos de variação nesses textos. Logo, se voltando para as mudanças das línguas no eixo do tempo, a linguista uniu os métodos da Sociolinguística Variacionista aos da Linguística Histórica, dando início à vertente conhecida como Sociolinguística Histórica. Para a autora, o principal objetivo dessa disciplina é proporcionar uma descrição das formas ou dos usos que apresentam variação em uma determinada comunidade, descrevendo o desenvolvimento desse processo.

Em consonância com Conde Silvestre (2007), o *Princípio do Uniformitarismo* desempenha um papel essencial na Sociolinguística Histórica. Desenvolvido, inicialmente, no âmbito da Geologia, esse princípio foi introduzido na Linguística por Whitney e por Brugmann no século XIX, os quais defendiam que, a partir da observação do presente, era possível explicar e compreender os processos que afetaram a língua no passado. Dessa maneira, segundo esse princípio, os fenômenos linguísticos observáveis nos tempos atuais não se distanciam tanto daqueles que se manifestaram em outras épocas (ROMAINE, 1982; LABOV, 1994; CONDE SILVESTRE, 2007). No âmbito da Linguística Histórica, esse princípio passou a formular que as forças que condicionam as mudanças linguísticas no tempo presente são as mesmas que atuaram no passado. Partindo desse princípio, então, os estudiosos podem inferir como os fenômenos de variação e mudança se desenvolveram em épocas passadas.

Além de defender que a teoria da variação e a Linguística Histórica têm uma íntima relação, Romaine (1982) destaca a relevância dos textos escritos para o estudo de fenômenos linguísticos. De acordo com a linguista, o predomínio da língua falada sobre a língua escrita em estudos sociolinguísticos é patente; contudo, a manifestação da variação se dá em ambos os meios, dado que fala e escrita são apenas representações diferentes de uma mesma língua. Como bem ressaltam Coelho e Nunes de Souza (2014, p. 164), “a análise do texto escrito é fundamental para uma pesquisa sociolinguística que se preocupe em investigar sincronias mais antigas, uma vez que o gravador [...] é um invento do século XX”. Conde Silvestre (2007), na mesma direção, explica que não havia meios de gravação e de reprodução da voz em épocas mais distantes. Todavia, Labov (1994) elucida que os documentos históricos são sobreviventes imprevisíveis e que, muitas vezes, acabam sendo afetados pela hipercorreção, por erros cometidos pelos escribas ou pela mescla de dialetos. É nesse sentido que Labov

³ Romaine (1982) investiga a relativização no escocês medieval (1530-1550), analisando um *corpus* composto por diferentes tipos de textos – textos em prosa (documentos oficiais, prosa literária e prosa epistolar) e textos poéticos (cortês, religiosos/moralizantes, cômicos).

(1994, p. 11, tradução nossa) declara: “a Linguística Histórica pode ser entendida como a arte de fazer o melhor uso de maus dados”⁴. Nesse caso, os eventos socioculturais são resgatados e reconstruídos, e o controle de categorias sociais (idade, escolaridade, sexo etc.) torna-se desafiador, visto que os informantes não estão à disposição dos pesquisadores (LOPES; RUMEU, 2018). Apesar das dificuldades encontradas pelos historiadores da língua, a Sociolinguística Histórica tem desenvolvido métodos e técnicas para correlacionar os fenômenos linguísticos aos fatores sociais.

Conde Silvestre (2007) explica que alguns autores, tomando como base a viabilidade dos tipos de registros escritos para o estudo da variação e da mudança linguísticas ao longo do tempo, distribuem esses materiais em categorias. Dentre os autores que se propõem a realizar tal tarefa, Conde Silvestre (2007) destaca o trabalho de Schneider (2007), que segmenta cinco categorias de materiais escritos a serem utilizados na investigação da Sociolinguística Histórica: i) registros diretos do discurso oral, sejam estes escritos simultaneamente ao ato da comunicação ou transcritos posteriormente com o auxílio de algum mecanismo de gravação; ii) registros escritos de enunciados que foram recordados; nesse caso, essas anotações são realizadas um tempo depois da emissão do discurso oral, dependendo da memória do indivíduo; iii) registros escritos nos quais há manifestação das variedades da língua falada de determinada época, como diários e cartas íntimas; iv) registros e observações dos usos e comportamentos linguísticos de outros falantes, ainda que não constituam transcrições de um discurso real; e v) registros escritos que foram inventados e imaginados. Cabem, nesta última categoria, as peças teatrais, que, embora não constituam uma situação linguística real, são, muitas vezes, representações de falantes reais.

Conforme observamos anteriormente, os estudos de Romaine (1982) e de outros pesquisadores comprovam que fontes escritas constituem materiais fecundos para as investigações linguísticas. É nessa direção que propomos, neste trabalho, o estudo dos MD em textos dramáticos.

2.4 O teatro brasileiro nas décadas de 1950, 1960 e 1970

Não é possível relembrar a história do teatro brasileiro sem percorrer um pouco da história do Brasil, posto que os contextos sócio-histórico e sociocultural exercem grande influência sobre a forma e a estrutura das peças e dos espetáculos teatrais (CEBULSKI, 2013). Contudo, cumpre destacar que não faremos uma retomada de todo o percurso histórico

⁴ “Historical linguistics can then be thought of as the art of making the best use of bad data” (LABOV, 1994, p. 11).

do teatro brasileiro e, tampouco, entraremos em detalhes acerca da história do Brasil, mas o intuito desta subseção é discorrer, sucintamente, a respeito do teatro nas décadas de 1950, 1960 e 1970 e dos principais eventos históricos que ocorreram nos referidos períodos.

Em decorrência da Primeira Guerra Mundial, que perdurou de 1914 a 1918, as décadas iniciais do século XX são marcadas por um distanciamento das produções europeias no teatro, impulsionando um movimento nacionalista, que resulta na valorização de temáticas regionais. Após a Segunda Guerra Mundial (1939-1945), algumas companhias de teatro são fundadas no Brasil, a saber: Teatro Brasileiro de Comédia (TBC); Teatro de Arena; Teatro Paulista dos Estudantes (TPE); Teatro Oficina; entre outros. Nesse período, surgem, também, grandes nomes da dramaturgia brasileira, tais como Nelson Rodrigues, Ariano Suassuna, Gianfrancesco Guarnieri, Dias Gomes, Oduvaldo Vianna Filho, Plínio Marcos etc., tratando de temas diversos, como religião, regionalismos, urbanização, política, e, muitas vezes, fazendo críticas sociais condizentes com o período (CEBULSKI, 2013).

Na segunda metade do século XX, há uma ascensão da dramaturgia feminina, da qual fazem parte autoras como Hilda Hilst, Consuelo de Castro e Leilah Assumpção, versando sobre assuntos diversos com contundência e intensidade. Fazendo uma breve menção aos acontecimentos históricos que caracterizam esse período, Cebulski (2013, p. 86) cita: “a revolução comunista e o ideário marxista, as duas grandes guerras do século XX (o sentimento do absurdo), o desenvolvimentismo de JK, a teoria psicanalítica de Freud, o movimento da contracultura, a ditadura militar e a abertura política dos anos 80 e 90”.

O período de 1950 a 1970 foi bastante significativo para o teatro brasileiro, principalmente porque ocorreram inúmeras mudanças no contexto sócio-histórico do país. A partir da década de 1950, as peças teatrais adquirem um cunho social e começam a tratar de temas nacionais, tornando-se cada vez mais realistas e focalizando personagens de origem humilde e de classes sociais mais baixas, a fim de trazer à tona as opressões sofridas pelas camadas mais largas da sociedade. Klafke (2016, p. 13-14) registra que

[...] a segunda metade dos anos 1950 no Brasil é um período de otimismo, e foi nesse cenário de efervescência que o teatro brasileiro passou por um momento instigante de diálogo entre arte e política. O final da década de 1950 é marcado pelo fortalecimento de um teatro politicamente comprometido no Brasil, de inspiração nacionalista e visando retratar e atingir o povo, buscando em figuras populares seus protagonistas ideais. [...] O povo é compreendido como fonte única da identidade nacional, e é identificado como aquele que é marginalizado e explorado. A tônica da produção teatral deste período, e em diante ainda nos anos 1960, é uma abordagem crítica da realidade do país. Há uma noção geral de que a abordagem da vida do trabalhador e do pobre retirante, bem como de sua

bagagem cultural, seria um caminho para, por um lado, criticar a negligência do Estado quanto à situação do povo brasileiro, e, por outro, divulgar e valorizar os valores culturais das classes populares.

Por conseguinte, nas peças produzidas nesse contexto, há retratos de sujeitos pertencentes aos mais variados segmentos sociais e uma maior representatividade da sociedade daquela época.

A partir da década de 1960, o Brasil vive um de seus períodos mais sombrios: a ditadura militar. Nas palavras de Schwarcz e de Starling (2018, p. 166), "uma ditadura é formada por mandantes arbitrários, oposicionistas tenazes e uma população que precisa sobreviver - parte dela atravessa em silêncio, com medo ou apenas conformada o tempo do arbítrio", e outros, complementando o pensamento das autoras, levantam suas vozes apesar das circunstâncias. Nesse momento, as peças assumem um importante papel de denúncia e de combate à repressão; aliás, muitas obras foram censuradas por se contraporem ao regime militar. O teatro, então, se torna subversivo e uma ferramenta de resistência contra os ideais promulgados naquele período. Surge, nesse contexto, um posicionamento nacionalista que vem acompanhado por uma militância teatral e por questionamentos políticos (PRADO, 1988). Prado (1988, p. 97) afirma que "os anos imediatamente anteriores e posteriores a 1964 enfatizaram a dramaturgia política, ainda mais que a social". No final da década de 1960, em contraponto, começam a surgir obras que não mais focalizam temas políticos, essas peças se voltam para a subjetividade do indivíduo e para os conflitos deste em relação ao outro e a si mesmo, contendo um teor altamente psicológico. Nesse sentido, o teatro abre espaço para que grupos de pessoas marginalizadas e oprimidas sejam ouvidos (PRADO, 1988).

Essas considerações são apenas fragmentos de um cenário bem maior e complexo. No entanto, a partir desse panorama geral do teatro do século XX, é possível ter uma certa consciência dos contextos representados nas peças teatrais selecionadas para a análise, bem como compreender o comportamento dos personagens e seus perfis sociais.

2.4.1 O texto teatral como fonte de análise para investigações linguísticas

Rosa (2019) ressalta que há variados estudos que comprovam a eficácia de fontes escritas para a investigação da variação e mudança linguísticas. Partindo dos estudos de Romaine (1982), a autora defende que "a língua escrita é uma fonte primária tão valiosa quanto a língua falada, visto que ambas são apenas formas diferentes de representar a mesma língua" (ROSA, 2019, p. 38). Nesse mesmo sentido, Tarallo (1997), ao tratar sobre o estudo

linguístico em fontes históricas, já propunha o uso de materiais escritos. Nas palavras do autor,

Além do atlas, textos escritos em prosa que potencialmente reflitam o vernáculo de um certo período de tempo poderão ser utilizados. Reflitam-se aqui cartas de cunho pessoal, diários e textos teatrais que tenham visado, em seu momento de criação, a um retrato de fala de diferentes camadas sociais da comunidade. (TARALLO, 1997, p. 71)

Entretanto, embora fontes escritas sejam recursos riquíssimos para as investigações linguísticas, Rosa (2019) chama a atenção para o fato de que alguns registros não representam a língua dos falantes reais, restringindo-se à variedade adotada pelos escritores. Preti (2000), de maneira similar, declara que peças de teatro e outras fontes literárias devem ser utilizadas com cautela como *corpora* para os estudos linguísticos, pois os autores desses textos tendem a estilizar a fala em vez de reproduzi-la fielmente. Em um outro estudo, o autor explica que

Como constitui uma manifestação escrita, a linguagem literária tem afinidades maiores com essa modalidade de língua. Por mais que se pretenda aproximá-la do fenômeno da oralidade, o escrito literário pressupõe uma elaboração por parte do escritor, ainda mesmo quando sua intenção seja a de aproximar o que escreve da naturalidade da fala. (PRETI, 1999, p. 218)

Posteriormente, porém, Preti (2004, p. 138) reconhece que os autores "procuram fazer do diálogo de suas personagens um elemento a mais para configurar uma época, aproximando-se, tanto quanto possível, da realidade falada de seu tempo". Para Schneider (2007), textos escritos utilizados em estudos sobre variação e mudança linguísticas devem seguir os seguintes requisitos: i) o texto deve apresentar usos linguísticos que se assemelham à oralidade e ao vernáculo; ii) considerando a correlação com fatores extralinguísticos, esses materiais devem ser provenientes de origens diversas – a exemplo disso, o autor afirma que o perfil social dos autores desses textos deve ser variado, bem como seu estilo linguístico; iii) o fenômeno a ser investigado deve apresentar variabilidade nesses textos; e iv) o último requisito diz respeito à frequência do fenômeno em estudo nesses textos, em razão do caráter quantitativo da investigação. Desses requisitos, os textos teatrais atendem ao primeiro e ao segundo, pois a linguagem usada nas peças se aproxima do vernáculo, e essas obras são produzidas por diferentes autores, com origens e estilos diversos.

Um dos argumentos mais fortes contra o uso de textos literários na análise de fenômenos linguísticos está ligado ao fato de que, muitas vezes, os personagens seriam retratados de modo estereotipado. Todavia, em concordância com Brandão (2022), se e quando isso se dá, a construção desses estereótipos nos ajuda a compreender a língua e a

sociedade de outros períodos, assim como os valores sociais que as formas linguísticas carregam. Além disso, segundo a autora, o teatro não possui apenas bases artísticas, mas também históricas e culturais. As peças, portanto, podem constituir um retrato da sociedade de outros períodos, reconstruindo hábitos e costumes de toda uma época; nessa mesma direção, a língua integra esse retrato, atuando como uma parte dessa reconstrução histórica (BRANDÃO, 2022).

Outra crítica ao uso de fontes escritas em análises linguísticas é o fato de que o acesso a essa modalidade se restringia a classes mais abastadas e, conseqüentemente, não representaria as variedades dos demais usuários da língua. Em contraponto a essa ideia, argumentamos que, ainda que os escritores desses textos pertençam a classes específicas e possuam certo nível de letramento, há peças de teatro que representam a fala de personagens distintos, pertencentes a camadas sociais diversas (ROSA, 2019; BRANDÃO, 2022).

Nesse sentido, pensando no uso de peças teatrais para os estudos sócio-históricos, pode-se afirmar que, na procura por uma representação realista dos membros da sociedade, os dramaturgos tentam reproduzir os usos linguísticos de indivíduos reais nos diálogos de seus personagens. Está claro que a fala, no teatro, não constitui apenas uma forma de expressão, mas é também o meio pelo qual as características dos personagens se fazem conhecidas ao público; por conseguinte, o teatro se constrói a partir do diálogo e, por meio deste, os personagens são construídos. Destarte, o diálogo entre personagens fictícios possui características de uma conversação natural; isso se dá porque os enunciados contidos nas peças, apesar de serem produzidos por determinados personagens, são elaborados por falantes reais, isto é, os autores, que retratam outros usuários reais da língua. Logo, ao levarmos tais aspectos em consideração, percebemos que estamos lidando com dados produzidos por pessoas reais, os quais foram apreendidos pelos dramaturgos e utilizados no discurso de seus personagens.

Considerando-se essas discussões, percebe-se que a escolha desses materiais não se dá de maneira aleatória. Montgomery (2007), então, pensando na construção de *corpus/corpora*, propõe as seguintes dimensões a considerar: textual; temporal; social; espacial e de representatividade. Na **dimensão textual**, é preciso que o investigador considere a natureza do material escrito, tendo em vista o gênero, o estilo e a proximidade do texto com a oralidade. Embora o autor veja a representação da fala em peças teatrais como estereotipada e exagerada, ele reconhece que esses textos retratam variedades linguísticas não padronizadas de determinadas regiões e períodos. Além disso, esses textos trazem a percepção dos autores a respeito de uma determinada época, de seus costumes e de seus usos linguísticos.

Na **dimensão temporal**, os pesquisadores devem considerar os períodos de produção do material selecionado e qual momento da língua esse texto representa. Assim sendo, o autor considera que cartas são vantajosas por apresentarem datas e locais de origem, ao passo que textos literários têm como referência a data de publicação, que, não necessariamente, corresponde à data de produção desses textos. Brandão (2022), contrariando o posicionamento do autor e defendendo o uso de peças teatrais, explica que, por meio de estudos biográficos e da literatura vigente acerca do assunto, é possível saber exatamente a data de produção e de publicação das peças; somado a isso, segundo a autora, o espaço entre a data de escrita e de publicação dessas peças não constituiria, normalmente, uma lacuna grande o bastante para descaracterizar os usos linguísticos atribuídos aos personagens.

A **dimensão social** se refere à autoria dos textos e às informações que possuímos sobre esses autores. Considerando que há textos que podem ter sido ditados ou escritos por outras pessoas, sob o uso de pseudônimos, é essencial que os pesquisadores reúnam informações dos autores desses textos. Nesse caso, o uso de peças é bastante positivo, posto que há dados biográficos disponíveis dos autores e há informações concernentes ao perfil social dos personagens nas próprias peças.

A **dimensão espacial** corresponde às regiões de origem dos textos e dos autores. Mais uma vez, obras biográficas podem nos auxiliar e muitas peças trazem rubricas a respeito do local de ambientação da trama, facilitando o trabalho do pesquisador.

Finalmente, a **dimensão de representatividade** diz respeito à quantidade de indivíduos envolvidos na produção textual e a quais categorias esses sujeitos estão relacionados. No tocante às peças teatrais, Brandão (2022) enfatiza a riqueza de perfis sociais que podem ser encontrados nesses textos, bem como a diversidade linguística retratada na fala de diferentes personagens. De acordo com a autora, esses perfis sociais fazem referência às diferentes camadas sociais da época, distanciando-se da unilateralidade dos escritores que, em sua maioria, são homens, brancos e de classe social abastada. No entanto – não obstante a produção teatral esteja, em um primeiro momento, restrita a um grupo específico –, mais tarde, surgem novas vozes, que ecoam nos palcos da dramaturgia brasileira. Logo, em concordância com Brandão (2022), a homogeneidade do perfil dos autores diz respeito a um problema histórico, mas os personagens criados por esses autores apresentam perfis e usos linguísticos repletos de variabilidade.

Lopes e Rumeu (2018) demonstram a importância de realizar um levantamento do perfil sociocultural dos redatores – dos personagens, neste caso – de *corpora* históricos. Propondo alguns caminhos metodológicos para a execução dessa tarefa, as autoras explicam

que, a partir de certas evidências textuais contidas nos próprios materiais investigados, é possível obter algumas informações do perfil social desses sujeitos. No caso das peças, essas informações podem ser explicitadas em uma breve descrição nas primeiras páginas ou, simplesmente, surgir ao longo da trama, em rubricas ou em diálogos entre os personagens. Em casos assim, entram em cena os recursos de caracterização do personagem apresentados por Prado (2000, p. 88): “o que a personagem revela sobre si mesma, o que faz, e o que os outros dizem a seu respeito”.

Para fechar esta subseção, em consonância com Berlinck, Barbosa e Marine (2008), os estudos históricos da língua, de maneira geral, e mais especificamente aqueles que tomam o texto dramático como fonte de dados, devem ser respaldados por uma investigação referente ao contexto sócio-histórico e à estrutura da sociedade do período de produção dos textos. Contudo, segundo elas, não se pode desconsiderar a contribuição dos autores para a representação de determinadas épocas e linguagens.

2.5 Marcadores discursivos: diferentes rótulos, diferentes abordagens

Há múltiplas abordagens de MD, bem como há diversos rótulos utilizados para denominá-los, quais sejam: *marcadores conversacionais*, *operadores argumentativos*, *articuladores textuais*, *operadores discursivos*, *conectivos discursivos* etc. (FREITAG, 2007; GUERRA, 2007; FRASER, 2009). Somado a isso, são inúmeros os elementos que constituem o grupo dos MD, que inclui conjunções (*mas*, *e*), advérbios (*então*, *agora*), interjeições (*ah*, *oh*), sons não lexicalizados (*uhn uhn*, *ahn*) etc. Devido à falta da especificidade da natureza e das propriedades dos MD, há uma tendência para a agregação de novos itens a esse grupo, que se torna cada vez mais heterogêneo. Essa pluralidade de elementos torna válida a assertiva de Pottier (1962 apud VALLE, 2014, p. 39), segundo a qual “são classificados como MDs todos os elementos discursivos ‘com os quais não se sabe o que fazer’”.

A heterogeneidade desse conjunto é um fator reconhecido por vários estudiosos; até mesmo os autores que possuem pontos de vista divergentes concordam que esse conjunto é um tanto vasto e amorfo. Fraser (2006), por exemplo, ressalta a multiplicidade de rótulos existentes para se referir aos MD e pontua que não há convergência a respeito de quais itens compõem essa classe. Blakemore (2002) declara que, uma vez que não há um consenso quanto aos itens que compõem o conjunto dos MD, é difícil saber se os diferentes rótulos utilizados nomeiam elementos distintos ou o mesmo fenômeno. Schiffrin (2006, p. 321, tradução nossa) explica que os inúmeros rótulos não são apenas nomenclaturas distintas para o mesmo objeto de estudo, mas “refletem diferentes formas de pensar a organização do que

acaba constituindo diferentes conjuntos de palavras e de expressões”⁵; para ela, as consequências dessa diversidade de rótulos são, igualmente, teóricas e práticas. Ademais, Schiffrin (2001, 2006), comentando a abundância de trabalhos que se debruçam sobre o estudo dos MD e as diferentes abordagens que emergem desse contexto, conclui que esses variados pontos de vista divergem no que diz respeito aos princípios mais básicos sobre o fenômeno. Fischer (2006), por seu turno, propõe a sistematização das mais variadas abordagens de MD; para tal fim, a linguista reúne diferentes autores, que discorrem sobre suas perspectivas acerca do fenômeno. Essa iniciativa, segundo a autora, se justifica pelo fato de que

[...] os estudos disponíveis até agora dificilmente são comparáveis; as abordagens variam em muitos aspectos diferentes: a(s) língua(s) e itens levados em consideração, a terminologia utilizada, as funções consideradas, os problemas enfocados e as metodologias empregadas. É necessário algum tipo de visão geral que nos permita classificar as diferentes direções, métodos e perspectivas de pesquisa. Esta coleção constitui uma tentativa inicial de apresentar tal caminho através da selva de diferentes abordagens. (FISCHER, 2006, p. 1, tradução nossa)⁶

Fischer (2006) pontua que a obra pode auxiliar os leitores a enxergarem as similaridades entre as perspectivas, tornando-as comparáveis. Nessa mesma direção, Guerra (2007) expõe que há, muitas vezes, uma falta de consenso no que tange à definição e à atuação dos MD. Segundo a autora, cujo trabalho teve como um dos objetivos sistematizar semelhanças e discrepâncias entre algumas abordagens de estudo dos MD, as concepções adotadas por essas diferentes abordagens não são excludentes, apenas diferem por seus interesses específicos. Tanto Guerra (2007) quanto Penhavel (2010) comparam e sistematizam algumas abordagens de MD, concebendo um mapeamento geral dessas diferentes perspectivas, já que a expressiva quantidade de concepções sobre o fenômeno pode, segundo Penhavel (2010), dificultar o trabalho com esses elementos devido à falta de uma organização sistemática e de um conceito geral e elementar.

Tendo isso em vista, Penhavel (2010), partindo da proposta de Fischer (2006), propõe uma classificação das abordagens de MD, estabelecendo três tipos. Essa classificação, em

⁵ “[...] they reflect different ways of thinking about the organization of what ends up being different sets of words and expressions” (SCHIFFRIN, 2006, p. 321).

⁶ “[...] the studies available so far are hardly comparable; the approaches vary with respect to very many different aspects: the language(s) under consideration, the items taken into account, the terminology used, the functions considered, the problems focused on, and the methodologies employed. Some kind of overview is needed that allows us to sort out the different research directions, methods, and perspectives. This collection constitutes an initial attempt to present such a path through the jungle of different approaches” (FISCHER, 2006, p. 1).

consonância com o autor, tem como principal finalidade promover um maior esclarecimento a respeito desses trabalhos.

- (i) Tipo A: abordagens que analisam como MD expressões com função de conexão, integradas a um enunciado matriz e que se referem a um aspecto desse enunciado;
- (ii) Tipo B: abordagens que analisam como MD expressões com função de gerenciamento da conversação, que constituem enunciados independentes e que se referem a determinados domínios comunicativos;
- (iii) Tipo C: abordagens que analisam como MD ambos os tipos de expressões. (PENHAVAL, 2010, p. 35)

No primeiro grupo (Tipo A), o autor inclui as abordagens de Fraser e de Blakemore, que assumem como MD elementos que desempenham funções relacionais e de sequenciamento, focalizando o princípio de conectividade dessas unidades. Além disso, em ambas as abordagens, considera-se a contribuição dos MD para a interpretação dos enunciados. No segundo grupo (Tipo B), encontra-se a abordagem de Fischer, que, diferentemente de Fraser e de Blakemore, não considera os MD como expressões de natureza relacional, mas, sim, como unidades ligadas à organização e ao gerenciamento da interação. Finalmente, o terceiro grupo (Tipo C) engloba as abordagens de Schiffrin, que investiga unidades de natureza conectiva (*but, and*) e unidades de natureza mais interacional (*I mean, y'know*), e de Risso, Silva e Urbano, que delimitam dois tipos básicos de MD: os *basicamente sequenciadores*, mais voltados para a articulação e para o amarramento das porções textuais, e os *basicamente interacionais*, mais direcionados para o gerenciamento da interação, sinalizando envolvimento interpessoais.

Dos três grupos propostos por Penhavel (2010), optamos pelo Tipo C, pois entendemos que os MD são elementos multifuncionais que operam na organização textual-interativa dos textos falados. Por essa razão, neste trabalho, nos fundamentamos nas abordagens de Schiffrin (1987, 2001, 2006) e de Risso, Silva e Urbano (2019).

2.5.1 A abordagem de Schiffrin

Desenvolvida no âmbito da Análise do Discurso, mais especificamente na esfera do modelo de coerência do discurso, o principal foco da abordagem de Schiffrin é verificar como os MD se comportam para tornar o discurso coerente. Schiffrin (1987), inicialmente, declara que o campo de análise do discurso é bastante amplo e ambíguo, pois suas raízes não provêm apenas da Linguística, mas de outras áreas do conhecimento. Devido à vastidão e à

ambiguidade desse campo, antes de partir para a análise dos MD, Schiffrin (1987) discute alguns aspectos referentes a essa área com o objetivo de estabelecer uma base teórica para o estudo dos MD.

Para ela, no âmbito da análise do discurso, os pressupostos centrais acerca da língua envolvem contexto e comunicação. Primeiramente, ela pontua que “a língua sempre ocorre em um contexto”⁷. Mencionando contextos culturais, sociais e cognitivos, a linguista explica que, para compreender como a língua é utilizada e estruturada, é preciso considerar os contextos nos quais ela está imersa. De acordo com ela,

[...] a língua sempre ocorre em algum tipo de contexto, incluindo contextos cognitivos, nos quais a experiência e o conhecimento passados são armazenados e utilizados, contextos culturais, que consistem em significados e visões de mundo compartilhados, e contextos sociais, por meio dos quais tanto o eu quanto os outros recorrem a ordens institucionais e interacionais para construir definições de situação e de ação (SCHIFFRIN, 1987, p. 4, tradução nossa).⁸

Em segundo lugar, “a língua é sensível aos contextos”⁹ e os reflete, uma vez que ajuda a constituí-los. A exemplo disso, a autora cita estudos sociolinguísticos, que levam em consideração as pressões internas e externas envolvidas na mudança da língua. Em terceiro lugar, “a língua é sempre comunicativa”¹⁰. Schiffrin (1987) pressupõe que a comunicação ocorra quando um remetente dá informações; logo, segundo ela, a língua é comunicativa porque é direcionada a alguém. Por fim, “a língua é projetada para a comunicação”¹¹, pois alguns aspectos estruturais da língua só podem ser explicados ao considerarmos que eles foram desenvolvidos para a interação face a face.

Posto isso, Schiffrin (1987) propõe um modelo de coerência na conversação, que é um modelo discursivo, focalizando a coerência local (*local coherence*). Nesse modelo, há dois tipos de estruturas não linguísticas e pragmáticas: a **estrutura de troca** (*exchange structure*), que diz respeito à organização dos turnos conversacionais, e a **estrutura de ação** (*action structure*), que se refere aos segmentos de atos de fala no discurso. As unidades que integram a **estrutura ideacional** (*ideational structure*) são tomadas como semânticas – não

⁷ “Language always occurs in a context” (SCHIFFRIN, 1987, p. 3).

⁸ “[...] I assume that language always occurs in some kind of context, including cognitive contexts in which past experience and knowledge is stored and drawn upon, cultural contexts consisting of shared meanings and world views, and social contexts through which both self and others draw upon institutional and interactional orders to construct definitions of situation and action” (SCHIFFRIN, 1987, p. 4).

⁹ “Language is context sensitive” (SCHIFFRIN, 1987, p. 3).

¹⁰ “Language is always communicative” (SCHIFFRIN, 1987, p. 3).

¹¹ “Language is designed for communication” (SCHIFFRIN, 1987, p. 3).

pragmáticas, como as estruturas de troca e de ação –, envolvendo relações coesivas, tópicas e funcionais. Para Schiffrin (1987, p. 26, grifo da autora, tradução nossa)¹²,

[...] as estruturas de ideias diferem das estruturas de ação e de troca, porque consistem em unidades linguísticas (proposições com conteúdo semântico), enquanto as estruturas de troca e de ação emergem de unidades (turnos e atos) que são realizadas pelo uso da língua, mas não são linguísticas *per se*.

O **esquema de participação** (*participation framework*) diz respeito às relações entre os interlocutores e destes com o enunciado. Em consonância com Schiffrin (1987), falante e ouvinte podem se relacionar de diversas maneiras um com o outro, assim como se relacionam com o que está sendo produzido, isto é, a conversação/fala. Esse esquema, então, abrange tanto a relação entre os interlocutores quanto a relação entre falante e enunciado. Dessa maneira, assim como a estrutura de troca e a estrutura de ação, o esquema de participação é considerado pragmático, posto que as relações entre falante e ouvinte desempenham um papel central.

Finalmente, o **estado da informação** (*information state*) focaliza as capacidades cognitivas dos interlocutores e envolve a organização e o gerenciamento do conhecimento e do metaconhecimento. De acordo com Schiffrin (1987), os estados da informação estão, constantemente, evoluindo durante o fluxo conversacional e constituem processos dinâmicos da interação. Esse domínio, para a autora, é pragmaticamente relevante, mas não é pragmático em si.

Realizando uma síntese de seu modelo, Schiffrin (1987, p. 29, grifo da autora, tradução nossa)¹³ escreve:

[...] meu modelo de discurso tem estruturas não linguísticas (de troca e de ação) e estruturas linguísticas (ideacionais). Falante e ouvinte estão relacionados entre si e com seus enunciados em um esquema de participação. Seus conhecimentos e metaconhecimentos sobre ideias são organizados e gerenciados em um estado da informação. A coerência local no discurso é, então, definida como o resultado do conjunto de esforços dos interactantes para **integrar** conhecimento, significado, fala e ação.

¹² “[...] idea structures differ from action and exchange structures because they consist of linguistic units (propositions with semantic content), whereas exchange and action structures emerge through units (turns and acts) which are realized by the use of language, but are not linguistic *per se*” (SCHIFFRIN, 1987, p. 26, grifo da autora).

¹³ “In sum, my discourse model has both non-linguistic structures (exchange and action) and linguistic structures (ideational). Speaker and hearer are related to each other, and to their utterances, in a participation framework. Their knowledge and meta-knowledge about ideas is organized and managed in an information state. Local coherence in discourse is thus defined as the outcome of joint efforts from interactants to **integrate** knowing, meaning, saying and doing” (SCHIFFRIN, 1987, p. 29, grifo da autora).

Em um estudo mais recente, Schiffrin (2001) explica que a produção de um discurso coerente requer que os falantes recorram a diferentes tipos de conhecimento comunicativo. Em seguida, a autora destaca quatro domínios nos quais os MD operam, quais sejam: i) *expressivo*; ii) *social*; iii) *cognitivo*; e iv) *textual*. Os domínios *expressivo* e *social* estão, segundo a autora, intimamente relacionados e dizem respeito à “habilidade de usar a língua para exibir identidades pessoais e sociais, para transmitir atitudes e performar ações e para negociar relacionamentos entre o eu e os outros”¹⁴. O domínio *cognitivo* se refere à “habilidade de representar conceitos e ideias por meio da língua”¹⁵, e o *textual* diz respeito à “habilidade de organizar formas e de transmitir significados dentro de unidades da língua que sejam maiores do que uma única sentença”¹⁶ (SCHIFFRIN, 2001, p. 54, tradução nossa).

Em se tratando dos MD, Schiffrin (1987, p. 31, grifo da autora, tradução nossa)¹⁷ os define como “elementos **sequencialmente dependentes** que agrupam unidades do discurso”. Estendendo essa definição, a autora aponta que

Marcadores discursivos não se referem apenas às propriedades linguísticas (por exemplo, significados semânticos e pragmáticos, origem, funções) de um conjunto de expressões usado com frequência e à organização das interações sociais e das situações nas quais essas expressões são usadas, mas também dizem respeito às competências cognitiva, expressiva, social e textual daqueles que as usam. Uma vez que as funções dos marcadores são tão amplas, toda e qualquer análise de marcadores – mesmo aquelas que focalizam apenas um aspecto relativamente restrito de seu significado ou uma pequena porção de seus usos – pode nos ensinar algo sobre seu papel no discurso. (SCHIFFRIN, 2001, p. 67, tradução nossa)¹⁸

Em relação às funções dessas unidades, a autora questiona se um mesmo MD pode desempenhar mais de uma função. Schiffrin (1987) conclui que, além de ter mais de uma função, um MD pode ter uma função predominante. Os MD, então, são usados em diferentes domínios discursivos (*estrutura de troca, estrutura de ação, estrutura ideacional, esquema de*

¹⁴ “[...] the ability to use language to display personal and social identities, to convey attitudes and perform actions, and to negotiate relationships between self and other” (SCHIFFRIN, 2001, p. 54).

¹⁵ “[...] ability to represent concepts and ideas through language” (SCHIFFRIN, 2001, p. 54).

¹⁶ “[...] ability to organize forms, and convey meanings, within units of language longer than a single sentence” (SCHIFFRIN, 2001, p. 54).

¹⁷ “I operationally define markers as **sequentially dependent** elements which bracket units of talk” (SCHIFFRIN, 1987, p. 31, grifo da autora).

¹⁸ “Discourse markers tell us not only about the linguistic properties (e.g. semantic and pragmatic meanings, source, functions) of a set of frequently used expressions, and the organization of social interactions and situations in which they are used, but also about the cognitive, expressive, social, and textual competence of those who use them. Because the functions of markers are so broad, any and all analyses of markers – even those focusing on only a relatively narrow aspect of their meaning or a small portion of their uses – can teach us something about their role in discourse” (SCHIFFRIN, 2001, p. 67).

participação e estado da informação) e, por mais variadas que possam ser as origens desses elementos, eles podem compartilhar certas funções no discurso.

Propondo uma outra dimensão de análise para o fenômeno, Schiffrin (1987, 2006) discorre sobre a dêixis, visto que, para a autora, todos os MD têm funções indexicais, ou seja, eles são “apontadores” de um determinado domínio contextual. De acordo com ela, “a vantagem de ver marcadores discursivos como indexicais é que muitas das características que parecem tão preocupantes – incluindo, mas não se limitando, a multiplicidade – são na verdade características regulares de expressões dêiticas” (SCHIFFRIN, 2006, p. 335, tradução nossa)¹⁹.

Por fim, a linguista destaca algumas condições que permitem que uma determinada unidade seja usada como marcador: i) independência sintática; ii) comumente utilizada em posição inicial de enunciado; iii) variedade de contornos prosódicos; e iv) atuação nos níveis local e global do discurso e em diferentes domínios discursivos.

Em resumo, os estudos desenvolvidos por Schiffrin (1987, 2001, 2006) contribuem imensamente para a compreensão dos MD e demonstram que a heterogeneidade dos itens que integram esse conjunto é, na verdade, um indício do caráter multifuncional desses elementos.

2.5.2 A Gramática Textual-Interativa

A Gramática Textual-Interativa (GTI) ou Perspectiva Textual-Interativa²⁰ parte da concepção de linguagem “como uma forma de ação, uma atividade verbal exercida entre pelo menos dois interlocutores, dentro de uma localização contextual, em que um se situa reciprocamente em relação ao outro, levando em conta circunstâncias de enunciação” (JUBRAN, 2019, p. 32). Fundamentando-se em subsídios teóricos da Análise da Conversação, da Linguística Textual e da Pragmática, a GTI toma o texto falado como objeto de estudo, propondo-se a analisar a construção desses textos e a investigar os elementos envolvidos em sua estruturação. Dentre os mecanismos envolvidos na organização do texto falado, encontram-se os MD, que foram explorados de maneira inédita no âmbito da GTI.

Inicialmente, Risso, Silva e Urbano (2019) explicam que os MD são um conjunto bastante amplo e diversificado, envolvendo expressões diversas; por isso, faltava um estudo

¹⁹ “The advantage of viewing discourse markers as indexicals is that many of the features that seem so worrisome - including, but not limited to multiplicity - are actually regular features of deictic expressions” (SCHIFFRIN, 2006, p. 335).

²⁰ As pesquisas para a elaboração da Gramática do Português Culto Falado no Brasil iniciaram-se em 1987. No âmbito do Projeto de Gramática do Português Falado (PGPF), um dos grupos de trabalho ficou responsável pelos estudos da organização textual-interativa dos textos falados, de onde surge a Perspectiva Textual-Interativa. Portanto, os resultados desses estudos provêm de pesquisas intensas e de um esforço conjunto de diversos pesquisadores ao longo dos anos (cf. JUBRAN, 2019).

que determinasse a natureza e a propriedade dos MD. Diante da amplitude de abordagens e de conceitos a respeito do fenômeno, os autores desenvolvem uma pesquisa inovadora, com o foco voltado para os traços definidores dos MD, visto que estudos anteriores revelaram não haver um padrão de classificação para o que poderia ser considerado MD. Segundo os autores, agregam-se, continuamente, novos itens ao conjunto, sem, contudo, estabelecerem-se traços que auxiliem na identificação de um elemento como marcador.

Adotando o rótulo *marcadores discursivos*, por lhes parecer mais abrangente, os pesquisadores, a fim de identificarem traços definidores dos MD, analisam quinze minutos de vários inquéritos do projeto NURC²¹. Nessa coleta de dados, os estudiosos consideraram itens de fontes gramaticais diversas, incluindo aqueles que são definidos como MD de maneira consensual ou não. Ao final do processo, os autores obtiveram um total de 1.298 ocorrências. Diante desse alto número de ocorrências, Risso, Silva e Urbano (2019) propõem dez variáveis, com os seus respectivos traços, segundo as quais essas ocorrências foram analisadas:

Padrão de recorrência - variável 1

Traços	1 – de 1 a 3 vezes (baixa frequência)
	2 – de 4 a 9 vezes (média frequência)
	3 – de 10 vezes em diante (alta frequência)

Articulação de segmentos do discurso - variável 2

Traços	1 – sequenciador tópico
	2 – sequenciador frasal
	0 – não sequenciador

Orientação da interação - variável 3

Traços	1 – secundariamente orientador
	2 – basicamente orientador
	0 – fragilmente orientador

Relação com o conteúdo proposicional - variável 4

Traços	1 – exterior ao conteúdo
	0 – não exterior ao conteúdo

²¹ O Projeto da Norma Urbana Linguística Culta (NURC), iniciado em 1969, teve como objetivo documentar e investigar a oralidade, mais especificamente a norma urbana culta, de cinco capitais brasileiras, quais sejam Porto Alegre, Salvador, São Paulo, Recife e Rio de Janeiro. Esse acervo tem sido acessado e utilizado por diversos pesquisadores, constituindo, assim, o *corpus* de inúmeros trabalhos acadêmicos (cf. CASTILHO, 2019).

2 – não se aplica

Transparência semântica - variável 5

Traços 2 – totalmente transparente
 1 – parcialmente transparente
 0 – opaco
 3 – não se aplica

Apresentação formal - variável 6

Traços 1 – forma única
 2 – forma variante

Relação sintática com a estrutura oracional - variável 7

Traços 1 – sintaticamente independente
 0 – sintaticamente dependente

Demarcação prosódica - variável 8

Traços 1 – com pauta demarcativa
 0 – sem pauta demarcativa

Autonomia Comunicativa - variável 9

Traços 1 – comunicativamente autônomo
 0 – comunicativamente não autônomo

Massa fônica - variável 10

Traços 1 – até três sílabas tônicas
 2 – além de três sílabas tônicas

O **padrão de recorrência**, variável 1, diz respeito à frequência das formas classificadas como MD nos textos falados, uma vez que esses elementos costumam ser bastante recorrentes no discurso.

A variável 2, **articulação de segmentos do discurso**, se refere à função de articulação e de sequenciamento dos MD, que atuam, muitas vezes, como recursos coesivos do texto falado; nesse caso, salientam-se a organização tópica e a organização da estrutura frásica. Constatou-se que os MD, em sua maioria, são sequenciadores tópicos ou não sequenciadores.

A **orientação da interação**, que constitui a terceira variável, relaciona-se intimamente à segunda variável, visto que os itens cujos traços sequenciadores são mais atenuados tendem

a operar como orientadores da interação. De acordo com Risso, Silva e Urbano (2019, p. 375), "todo mecanismo com estatuto textual, como os MD, cumpre sempre uma função orientadora da interação, ainda que fragilmente". Desse modo, todos os MD desempenham uma função interacional; no entanto, em alguns casos, mais especificamente quando os marcadores têm um função sequenciadora bem definida, a função interacional se manifesta em segundo plano ou fragilmente. Os MD, então, seguem duas tendências funcionais: "maior projeção da interação, quando o foco funcional não está no sequenciamento de partes do texto" e "maior projeção da articulação textual, quando o foco deixa de incidir no eixo da interação" (RISSO; SILVA; URBANO, 2019, p. 376).

A **relação com o conteúdo proposicional**, variável 4, leva em consideração a ligação desses elementos com o conteúdo informacional das porções do texto. Os resultados revelaram que a maior parte dos MD são exteriores ao conteúdo, porém, mesmo que não contribuam para o conteúdo referencial, essas formas auxiliam na organização textual.

A **transparência semântica**, variável 5, diz respeito ao sentido lexical das palavras, ou seja, o significado previsto pelos dicionários e pelas gramáticas. No caso das unidades investigadas, o traço mais recorrente foi o de transparência parcial, demonstrando que os MD possuem uma certa relação com o significado de origem, mas são adaptáveis ao contexto da situação comunicativa.

A análise da **apresentação formal**, variável 6, teve como objetivo verificar se as unidades sob análise eram representadas por uma forma fixa ou se apresentavam formas variantes. Ainda que as ocorrências dos dois traços tenham sido bastante aproximadas, os autores atestam que "as unidades em análise são normalmente cristalizações usadas automaticamente no discurso e não propriamente unidades formuladas *ad hoc*" (RISSO; SILVA; URBANO, 2019, p. 378, grifo dos autores).

A **relação sintática com a estrutura oracional**, variável 7, tem relação com as funções desempenhadas pelos MD na estrutura da oração. Os resultados apresentados pelos autores condizem com os de estudos anteriores: os MD são, em sua maioria, sintaticamente independentes.

A **demarcação prosódica**, variável 8, corresponde à presença de algum tipo de variação melódica em torno dos MD, o que diferenciaria e destacaria esses elementos dos demais segmentos do discurso.

O exame da **autonomia comunicativa**, variável 9, teve o intuito de observar se os marcadores são portadores de conteúdo proposicional, ou seja, se os MD seriam suficientes para constituírem enunciados.

Por fim, a última variável, **massa fônica**, diz respeito à extensão da forma, visto que os MD, normalmente, possuem forma reduzida, que se torna cada vez mais fixa.

A partir dessas variáveis, os autores puderam identificar quais os traços mais frequentes dos MD e notaram algumas oscilações, dado que nem todos os MD apresentam os traços considerados recorrentes. Ainda assim, os pesquisadores perceberam que há uma certa flexibilidade e que esta respeita alguns padrões de prototipicidade, que podem ser representados por seis **matrizes básicas**, conforme ilustrado no quadro 2. No quadro, encontram-se as dez variáveis consideradas pelos autores e as seis matrizes básicas (A, B, C, D, E e F); os números nas colunas das matrizes representam os traços das variáveis, exibidos anteriormente. Cada matriz, portanto, apresenta uma combinação diferente de traços.

Quadro 2 – Matrizes básicas propostas pela GTI

Variáveis		A	B	C	D	E	F
1	Padrão de recorrência	3	3	3	3	3	3
2	Articulação de segmentos do discurso	0	0	1	1	1	1
3	Orientação da interação	2	2	1	1	0	0
4	Relação com o conteúdo proposicional	1	1	1	1	1	1
5	Transparência semântica	1	1	1	1	1	1
6	Apresentação formal	2	1	1	2	2	1
7	Relação sintática com a estrutura oracional	1	1	1	1	1	1
8	Demarcação prosódica	1	1	1	1	1	1
9	Autonomia comunicativa	0	0	0	0	0	0
10	Massa fônica	1	1	1	1	1	1

Fonte: Adaptado de Risso, Silva e Urbano (2019, p. 383-384)

Esse resultado levou os autores a constatarem que o comportamento dos MD não é homogêneo e que esses elementos podem desempenhar inúmeras funções, constituindo um conjunto multifuncional. Dentre todos os traços, cinco foram apontados como estáveis, os quais são decisivos para se classificar um elemento como MD. São estes: i) exterioridade dos MD em relação ao conteúdo proposicional; ii) independência sintática; iii) ausência de autossuficiência comunicativa; iv) articulação tópica; e v) orientação da interação. Para Risso, Silva e Urbano (2019), a combinação desses traços forma o que eles chamam de **núcleo-piloto**, que diz respeito às características estáveis dos elementos que compõem o conjunto dos MD. Não obstante os pesquisadores proponham essa articulação de traços comuns, de acordo com eles, “os MD não chegam a constituir uma classe discreta, que se esgota nesses padrões e se delimita perfeitamente por eles” (RISSO; SILVA; URBANO, 2019, p. 387). A partir disso, os estudiosos indicam os MD como uma classe gradiente, que se distribui em uma espécie de *continuum*, no qual há elementos mais típicos (**MD prototípicos**)

e menos típicos (**MD não prototípicos**). No primeiro caso, os elementos incorporam todos os traços das matrizes básicas, ao passo que, no segundo caso, os elementos manifestam esses traços de maneira parcial, podendo variar, sem, contudo, ultrapassar uma determinada margem.

Além dos MD prototípicos e não prototípicos, Risso, Silva e Urbano (2019) apresentam um outro grupo: as **unidades limítrofes**. Em consonância com os autores, essas expressões realizam uma intersecção com o conjunto de MD; todavia, mesmo que essas unidades compartilhem alguns traços com os marcadores, há certas particularidades que as distanciam desse conjunto. Diferentemente dos MD não prototípicos, as unidades limítrofes se desviam de maneira significativa dos traços comuns ao grupo dos marcadores.

Percorrido todo esse trajeto, os autores concluem que

Os MD sinalizam, portanto, articulações textuais e relações interpessoais, com foco funcional em um ou outro desses aspectos, que particularizam dois grandes subconjuntos de MD: os *basicamente sequenciadores* e os *basicamente interacionais* (RISSO; SILVA; URBANO, 2019, p. 390, grifos dos autores).

De maneira sucinta, os MD *basicamente sequenciadores* (*então, mas, aí, assim* etc.) são mais atuantes na articulação textual, voltando-se à organização tópica dos textos e ao amarramento das porções textuais, e os MD *basicamente interacionais* (*ah, né?, entendeu?, olha, uhn uhn* etc.) operam mais acentuadamente na orientação da interação, voltando-se para o jogo de relações interpessoais. Segundo os autores, essas duas funções se inter-relacionam, havendo maior ou menor grau de projeção de uma ou de outra, ou seja, ainda que um elemento desempenhe função de sequenciador, haverá um grau leve de função interacional e vice-versa.

2.5.3 Um diálogo entre abordagens: a multifuncionalidade dos MD em foco

A partir da proposta de Penhavel (2010), optamos pelas abordagens de Schiffrin (1987, 2001, 2006) e de Risso, Silva e Urbano (2019), porque ambas reconhecem que os MD atuam na articulação textual e no gerenciamento da interação. Todavia, à primeira vista, a interface entre essas perspectivas pode parecer inexequível, pois, ainda que as duas abordagens considerem a atuação dos MD nos planos textual e interacional, as propostas são distintas. A principal divergência repousa no fato de que Risso, Silva e Urbano subdividem o conjunto dos MD em *basicamente sequenciadores* e *basicamente interacionais* – dando a impressão equivocada de que os MD desempenham apenas uma dessas funções, mas não

ambas; esse, porém, não é o caso, conforme veremos adiante –, enquanto Schiffrin analisa elementos de natureza diversificada, sem recorrer à subdivisão da classe dos MD. Apesar das divergências, há um elo entre as duas perspectivas, qual seja a multifuncionalidade dos MD.

Schiffrin (1987) analisa os MD *oh, well, and, but, or, so, because, now, then, y'know* e *I mean*. A linguista classifica os MD em estudo da seguinte forma: i) *oh*: marcador de gerenciamento de informações; ii) *well*: marcador de resposta; iii) *and, but, or*: conectivos; iv) *so, because*: marcadores de causa e efeito; v) *now, then*: advérbios de tempo; e vi) *y'know, I mean*: informação e participação. Dessa maneira, mesmo que a autora não elabore subconjuntos, ela atribui determinados papéis a cada MD investigado, salientando a multifuncionalidade e a versatilidade dessa classe cujos componentes podem atuar em diferentes planos do discurso e contribuir para o estabelecimento da coerência discursiva.

Jakobson (1960 apud SCHIFFRIN, 1987), após identificar as seis funções da linguagem, conclui que seria bastante difícil encontrar mensagens verbais que desempenhassem unicamente uma função. Com base nessa afirmação, Schiffrin (1987) chega à conclusão de que

O ponto de vista de Jakobson sugere não apenas que um único marcador pode ter múltiplas funções, mas que um marcador pode ter **uma função predominante**: talvez alguns marcadores sejam especializados para funções interacionais e outros para funções ideacionais. (SCHIFFRIN, 1987, p. 60, grifo nosso, tradução nossa)²²

Nessa direção, Risso, Silva e Urbano (2019, p. 390) reconhecem que “o maior peso do fator interacional corresponde normalmente a uma diluição do papel articulador e, inversamente, o crescimento da atuação sequenciadora convive com um grau mais atenuado de manifestação do jogo de relações interpessoais”. Portanto, as particularidades funcionais dos subconjuntos *basicamente sequenciador* e *basicamente interacional* não são excludentes, mas estão inter-relacionadas. Zanin e Gonçalves (2022, p. 65, grifo nosso), corroborando essas afirmações, atestam que

[...] as funções textual-interativas de MD devem ser entendidas em termos de **predominância**, assim estabelecida: quando a função de um MD é predominantemente textual, implica reconhecer que ela também comporta uma contraparte interacional menos saliente, e, contrariamente, quando a

²² “Jakobson's point suggests not only that a single marker might have multiple functions, but that a marker might have a predominant function: perhaps some markers are specialized for interactional functions, others for ideational functions” (SCHIFFRIN, 1987, p. 60).

função predominante é interacional, a contraparte textual é a que fica minimizada (GUERRA, 2007).

O que se percebe, a partir dessas considerações, é que a predominância é um fator essencial para a compreensão dessas funções, posto que a proeminência de uma função não acarreta o desaparecimento da outra. Todavia, tomando *but* como exemplo – MD que atua na estrutura ideacional, opera como recurso coesivo e sinaliza um movimento interacional –, Schiffrin (1987) reconhece que não é uma tarefa simples apontar qual a função predominante desse elemento. Apesar disso, segundo a autora, não há como negar que *but* opera, simultaneamente, em vários planos do discurso. Schiffrin (1987), então, não determina qual função seria mais predominante, mas admite que os MD são elementos multifuncionais; por outro lado, Risso, Silva e Urbano (2019), mesmo constatando que os planos textual e interacional estão inter-relacionados, apontam a função que aparenta ser mais predominante nos MD analisados em seu trabalho.

Em conformidade com Penhavel (2005), as funções sequenciadora e interacional são duas **macrofunções** dos MD; logo, no domínio dessas funções gerais, é possível distinguir uma variedade de **subfunções**. A exemplo disso, Urbano (2019) destaca as seguintes subfunções dos MD interacionais: i) fáticos de natureza imperativa e entonação exclamativa; ii) fáticos de natureza ou entonação interrogativa, produzidos após enunciado declarativo; iii) fáticos de natureza e entonação interrogativa, produzidos após enunciado interrogativo; iv) *feedbacks*; e v) início de respostas formais ou de comentários²³. Os MD, então, não podem ser encaixotados em uma única função, porque constituem uma classe gradiente e atuam em variados domínios funcionais.

Neste trabalho, consideramos que os RAD são elementos multifuncionais que operam tanto no plano interacional quanto no plano textual, os quais estão inter-relacionados. Ademais, a depender do contexto em que essas formas são usadas, há maior projeção de uma função sobre a outra, ocasionando a atenuação e não o apagamento da função menos predominante.

2.6 Delimitação do objeto

Conquanto existam traços que facilitem a identificação dos MD, reduzindo o abismo entre os elementos que integram essa classe, o conjunto ainda é composto por expressões muito diversificadas; essa heterogeneidade é notável, até mesmo, nos subconjuntos de MD,

²³ Guerra (2007) revisita essas subfunções, utilizando os seguintes rótulos: injuntiva; *checking*; *feedback*; iniciadora e interpelativa.

tais como o dos RAD. Em decorrência dessa pluralidade de constituintes, os trabalhos que se dedicam à investigação dos MD propõem a delimitação do objeto de estudo, tendo como base, principalmente, o critério funcional desses elementos, isto é, as funções que os MD desempenham no discurso. Por isso, decidimos delimitar o nosso objeto de estudo, selecionando apenas algumas unidades pertencentes ao subconjunto dos RAD.

Valle (2014), procurando selecionar um grupo consistente para a sua análise, elaborou quatro critérios unificadores para delimitar variáveis discursivas. O **critério de unidade funcional e de compartilhamento de contextos de uso** leva em consideração as funções desempenhadas pelas unidades discursivas e os contextos nos quais elas são utilizadas. O **critério de unidade conceptual e classe gramatical de origem** se volta para a natureza conceptual desses elementos e para a semelhança de significados entre eles, os quais permitem que esses itens compartilhem certas funções e contextos de uso. Sendo assim, elementos provenientes da mesma classe gramatical e com sentidos parecidos estariam mais sujeitos a desenvolver significados e funções similares. O **critério de relevância do item para a comunidade investigada**, conforme indicado pelo próprio nome, faz relação entre as formas linguísticas e as questões identitárias e estilísticas de determinados grupos sociais. Finalmente, o **critério de frequência de uso dos itens** corresponde à recorrência dos itens na amostra selecionada, isto é, diz respeito à quantidade de vezes que uma determinada expressão aparece, tornando-a quantitativamente significativa ou não. Esses critérios possuem diferentes pesos e, em consonância com a autora, o primeiro critério é o de maior peso, já que “a unidade funcional e o compartilhamento de contextos de uso é o principal fio condutor de análises variacionistas de unidades discursivas” (VALLE, 2014, p. 53).

Cabe ressaltar que, no caso de Valle (2014), todas as entrevistas da amostra²⁴ selecionada foram ouvidas para que, em seguida, o objeto de estudo fosse delimitado; nesse sentido, Valle selecionou sete formas (*sabe?*, *sabes?*, *entende?*, *entendeu?*, *entendesse?*, *tá entendendo?* e *tás entendendo?*) para compor sua análise. Neste trabalho, em contrapartida, as formas foram selecionadas antes da leitura das peças; logo, partimos dos dois primeiros critérios para a delimitação do nosso objeto. Assim sendo, os critérios de relevância para a comunidade e de frequência não foram considerados, pois não selecionamos uma comunidade ou grupo social específicos e não estipulamos uma frequência de ocorrências. Na verdade, consideramos que a frequência de ocorrências dos RAD constitui um fator relevante para os

²⁴ Para a realização de sua tese, Valle (2014) utilizou a amostra Brescancini-Valle (2001-2010), que integra o Banco de Dados VARSUL (Variação Linguística Urbana do Sul do País). A autora examinou 30 entrevistas com informantes da comunidade da Barra da Lagoa – Florianópolis/SC.

resultados da pesquisa, já que a recorrência de um elemento pode indicar certas preferências ou a fixação de uma determinada forma no vocabulário dos falantes, o que corresponde à variável *padrão de recorrência* dos estudos de Risso, Silva e Urbano (2019).

Baseando-nos nos critérios estabelecidos por Valle (2014), selecionamos os RAD derivados dos verbos de cognição *compreender*, *entender*, *perceber*, *saber* e *ver*, aos quais pode ser dispensado um tratamento variável, porque, além de terem em comum a mesma classe gramatical de origem, esses RAD possuem significados semelhantes e atuam em um mesmo domínio funcional. De acordo com Valle (2014, p. 56, grifos da autora):

Os verbos *ver*, *perceber*, *saber*, *entender* e *compreender* são de natureza conceptual semelhante, todos indicando processos mentais que, a princípio, diferem entre si em grau de complexidade: *ver* e *perceber*, voltados a atividades mais concretas, seriam menos complexos que verbos como *entender* e *compreender*, que envolvem processos mentais.

Halliday e Matthiessen (2004), por seu turno, propõem quatro subtipos de processos mentais: i) de percepção: relacionados à sensação ou à observação (degustar, ver, ouvir, cheirar, sentir etc.); ii) de cognição: relacionados à compreensão, à consideração e à elaboração ou aquisição do conhecimento (imaginar, acreditar, compreender, entender, perceber, saber etc.); iii) de desejo: relacionados à expressão de vontade (querer, desejar, esperar, concordar etc.); e iv) de afeição: relacionados a sentimentos e a emoções (odiar, adorar, amar, gostar, temer etc.) (HALLIDAY; MATTHIESSEN, 2004; COSTA; FURTADO DA CUNHA, 2016).

É válido frisar que Valle (2014) distingue os verbos *ver* e *perceber* dos demais, pontuando o caráter menos complexo e mais concreto dessas formas, ao passo que Halliday e Matthiessen (2004) incluem *perceber* (*realize*) no conjunto de processos mentais de cognição, juntamente com *compreender*, *entender* e *saber*, e colocam *perceber* (*perceive*) no conjunto de processos mentais de percepção. No Português Brasileiro, porém, não existe essa distinção de vocábulos, o que justifica o uso de *perceber* para indicar processos mentais de percepção e de cognição. As propostas dos autores, então, se complementam, sugerindo que *perceber* – também *ver*, nesse caso – desliza de um processo mais concreto e menos complexo para outro mais abstrato e com graus mais acentuados de complexidade. Conforme explicamos na introdução, por meio do estudo de Rost Snichelotto (2009), que investiga MD derivados dos verbos de percepção *ver* e *olhar*, esses verbos, que antes indicavam atividades concretas, adquirem sentidos abstratos, ligados à cognição. Martelotta, Votre e Cezario (1996, p. 39, grifo dos autores) referem-se a *ver* e a *perceber* como verbos de percepção, os quais “tendem

a ser usados metaforicamente no sentido de **saber**”. Heine *et al.* (1991 apud MARTELOTTA; VOTRE; CEZARIO, 1996), semelhantemente, colocam *ver* e *perceber* no grupo dos verbos de percepção sensorial, e *saber* e *entender* no conjunto dos verbos de processo mental. Tais características, na verdade, estão intimamente relacionadas ao processo de gramaticalização desses verbos, que será discutido adiante.

Com base nessas considerações, é possível afirmar que todas as fontes verbais escolhidas (*compreender*, *entender*, *perceber*, *saber* e *ver*) são verbos de cognição; todavia, alguns desses verbos, mais especificamente *ver* e *perceber*, são, em sua base, de percepção.

2.6.1 Os requisitos de apoio discursivo

Semelhantemente ao conjunto geral de MD, os RAD recebem diversos rótulos: *busca de aprovação discursiva* (BAD); *checking*; *marcadores de controle de contato*; *apêndices comprovativos*; entre outros (GUERRA, 2007; VALLE, 2014; URBANO, 2019). Nesta pesquisa, contudo, optamos por *requisitos de apoio discursivo* (RAD), que, segundo Valle (2014), é o rótulo mais difundido em estudos brasileiros. Os RAD englobam expressões diversas e de naturezas distintas, que, apesar disso, possuem certas similaridades e atuam de forma parecida no discurso. Dentre essas formas encontram-se expressões como *né?*, *sabe?*, *entendeu?*, *tá?*, *percebeu?*, *viu?*, *tá ligado?*, *sacou?* etc.

Para entender como funcionam e quais são as particularidades dos RAD, recorreremos a estudos prévios sobre o fenômeno (MARCUSCHI, 1986; SILVA; MACEDO, 1996; URBANO, 1999b; VALLE, 2001, 2014; GUERRA, 2007; FREITAG, 2009; FREITAG; SILVA; EVANGELISTA, 2017).

Marcuschi (1986) inclui os RAD na categoria de *recursos verbais*, os quais, segundo ele, constituem uma classe estereotipada e bastante recorrente, que não chega a contribuir com novas informações para o tópico discursivo, não obstante sejam cruciais para o funcionamento do discurso. Ao tratar sobre os tipos, funções e posições dos MD no discurso, o autor divide os *marcadores verbais* em dois subtipos (*sinais do falante* e *sinais do ouvinte*) e aponta duas funções (*conversacional* e *sintática*). A *função conversacional*, que corresponde aos MD *basicamente interacionais* na GTI, engloba os *sinais do falante* – que, de acordo com Marcuschi (1986), são responsáveis pela sustentação dos turnos, pelo preenchimento de pausas, pela organização e orientação do discurso etc. – e os *sinais do ouvinte* – que, geralmente, são realizados em sobreposição ao turno do falante com o objetivo de orientar, dando sinais de concordância, de discordância, de dúvida, de encorajamento ou de interrupção do turno do falante. A *função sintática*, por outro lado, corresponde aos MD *basicamente*

sequeenciadores na GTI e abrange os elementos responsáveis pelo encadeamento e pela segmentação das partes do texto. Quanto à posição, Marcuschi (1986) explica que os MD podem ser encontrados em diferentes posições dentro do turno conversacional ou entre turnos. O linguista admite que esses elementos “podem aparecer em várias posições: na troca de falantes, na mudança de tópico, nas falhas de construção, em posições sintaticamente regulares. Fundamentalmente, eles podem operar como *iniciadores* (de turno ou unidade comunicativa²⁵) ou *finalizadores*” (MARCUSCHI, 1986, p. 61, grifos do autor). Para o autor, os sinais do falante ocorrem em posições inicial, medial ou final de turno, ao passo que os sinais do ouvinte, por vezes, se sobrepõem aos sinais do falante, surgindo em momentos de concordância ou de discordância com o tópico.

Silva e Macedo (1996), referindo-se a esses marcadores como RAD, se debruçam no estudo desse subconjunto. As autoras notaram a significativa recorrência dessas unidades no *corpus*²⁶ investigado e, segundo elas, a principal função dessas formas seria a de manter o fluxo conversacional e de estabelecer harmonia entre os participantes. As autoras apontam, ainda, as seguintes características dos RAD: i) atuam na manutenção da conversação; ii) são utilizados com o intuito de pedir a aquiescência ou de se certificar da atenção do interlocutor; iii) possuem forma fixa ou apresentam pouca variabilidade; iv) mantêm a entonação/função interrogativa; e v) surgem, geralmente, em posição final de enunciado ou de turno. Por meio dos resultados da pesquisa, as autoras constataram que as variáveis sociais não exercem grande influência sobre os RAD; a esse respeito, as autoras levantaram duas hipóteses: primeiramente, fenômenos discursivos são menos influenciáveis por fatores extralinguísticos e, em segundo lugar, a escola, propagadora da gramática normativa, não foi capaz de sistematizar o discurso.

Assim como na investigação realizada por Silva e Macedo (1996), Urbano (1999b) nota a significativa recorrência de RAD no diálogo analisado²⁷. Referindo-se a esses elementos como *marcadores de teste de participação* ou *busca de aprovação discursiva* (BAD), o linguista classifica os RAD como sinais do falante, similarmente a Marcuschi (1986), incluindo-os no grupo de marcadores linguísticos verbais lexicalizados. Diferentemente de Silva e Macedo (1996), que afirmam que os RAD ainda são utilizados de

²⁵ A unidade comunicativa é, segundo Marcuschi (1986, p. 61-62), “um substituto conversacional para ‘frase’, ou seja, é a expressão de um conteúdo que pode dar-se, mas não necessariamente, numa unidade sintática tipo frase”.

²⁶ Silva e Macedo (1996) utilizam como *corpus* as 64 entrevistas da Amostra Censo, que constitui uma amostra estratificada da comunidade carioca. Para a execução das entrevistas, consideraram-se os seguintes aspectos: sexo, faixa etária e nível de escolaridade. Os entrevistados pertenciam a diferentes bairros e estratos sociais (cf. PAIVA; PAREDES SILVA, 2012).

²⁷ Urbano (1999b) analisa o trecho de um inquérito do Projeto NURC/SP.

modo a receber um estímulo ou resposta do interlocutor, indicando sua aquiescência, Urbano (1999b, p. 96) constata que “se trata de uma espécie de pergunta retórica, que abre expectativa de ‘resposta’, mas apenas no nível pragmático e dificilmente com implicações semânticas”. Desse modo, o interlocutor já não sente a necessidade de interagir com o falante diante do uso dos RAD. No que tange à posição desses elementos no discurso, o autor explica que os RAD, comumente, se posicionam em final de unidade entonacional – isto é, “a expressão linguística de uma informação ou ideia, atualizada e reconhecida num dado momento por meio de uma entonação, específica” (URBANO, 1999b, p. 96) – ou podem caracterizar a passagem, consentida ou não, de turno. Mais adiante, Urbano (1999b) pontua que esses marcadores são, frequentemente, acompanhados por pausas ou por marcadores de apoio ou de atenção, como *uhn uhn* e *ahn*.

Valle (2014), partindo da abordagem de Schiffrin (1987, 2001), afirma que os RAD, além de atuarem no plano interacional, operam no plano textual, dando ênfase a certas partes do texto, de modo a chamar a atenção do interlocutor para uma informação específica. Fundamentando-se em estudos prévios (ÖSTMAN, 1981; SCHIFFRIN, 1987; MACEDO; SILVA, 1989; GALUÉ, 2002; MARTELOTTA; LEITÃO, 1998; entre outros), a autora compõe um quadro²⁸, reunindo as principais funções dos RAD nos planos textual e interacional, os quais, segundo ela, se interseccionam. Com relação à influência de fatores extralinguísticos sobre a produção de RAD, Valle (2001, 2014), semelhantemente a Silva e Macedo (1996), declara que os RAD não apresentam diferenças sociais expressivas. Em consonância com Valle (2001, p. 147), “a escolha das formas dos RADs parece ser pouco condicionada por padrões sociais, estando mais relacionada a atitudes individuais de cada informante [...]”. Desse modo, a influência das variáveis de natureza social sobre fenômenos discursivos costuma ser pequena.

Guerra (2007), partindo dos estudos de Urbano (2006), propõe o rótulo *checking* para nomear o subconjunto dos RAD e explica que se trata de uma subfunção interacional cujos elementos estão voltados para a orientação do falante em direção ao ouvinte. Por meio dos *checkings*, o falante busca o “apoio” ou a “aprovação” do interlocutor para que possa dar continuidade ao seu discurso. De acordo com Guerra (2007), os *checkings* são utilizados como perguntas retóricas; ocorrem, geralmente, no interior dos turnos conversacionais; e, ao usá-los, os falantes não costumam ceder o turno para o interlocutor.

Freitag (2009) afirma que os RAD são elementos voltados para a organização da fala, operando nos seguintes planos: i) *interpessoal*: voltado para as relações entre os

²⁸ O referido quadro pode ser consultado em Valle (2014, p. 90).

interlocutores; ii) *interpessoal e textual*: atraindo a atenção do ouvinte para certas porções textuais; e iii) *rítmico*: atuando como formas automatizadas cuja função é demarcar, pontuar, o ritmo – nesse caso, o contorno interrogativo torna-se menos evidente. Um único RAD pode desempenhar as três funções, revelando seu caráter multifuncional.

Freitag, Silva e Evangelista (2017), partindo da abordagem de Schiffrin (2001), atestam que os MD operam nos domínios cognitivo, expressivo, social e textual. Os MD, então, estariam correlacionados a fatores sociais, estilísticos e identitários, os quais motivariam sua escolha e uso. As autoras focalizam os RAD *entendeu?*, *sabe?*, *né?* e *certo?*, referidos como marcadores discursivos interacionais, que, segundo elas, são multifuncionais e atuam no planejamento, na manutenção e na organização da interação. Nas palavras das linguistas,

Marcadores discursivos interacionais são elementos linguísticos com especificidades de estratégias discursivas organizadas de planejamento, replanejamento, correção ou manutenção da fala, e, por isso, contribuem para externalizar perspectivas e atitudes dos falantes, assegurando a ancoragem pragmática do conteúdo quando definem a força ilocutória, as atitudes assumidas em relação ao conteúdo, a checagem da atenção do ouvinte para a mensagem transmitida e a orientação do falante com relação à natureza do elo sequencial entre as unidades textuais (FREITAG; SILVA; EVANGELISTA, 2017, p. 62).

A partir desses estudos, é possível traçar um panorama geral da atuação e do funcionamento dos RAD no discurso. Dado que estamos partindo das abordagens de Schiffrin (1987, 2001, 2006) e de Risso, Silva e Urbano (2019), consideramos que os RAD desempenham diversas funções nos planos textual e interacional e podem surgir em variadas posições no discurso.

2.6.2 Gramaticalização dos RAD

Indo na contramão da TVML, que afirma que mudança pressupõe variação (WEINREICH; LABOV; HERZOG, 2006[1968])²⁹, os MD resultam do processo de mudança e, posteriormente, variam. Zanin e Gonçalves (2022, p. 67) atestam que,

Enquanto para variacionistas a variação é considerada sempre um estágio anterior da mudança (WEINREICH; LABOV; HERZOG, 1968), para estudiosos da gramaticalização, essa proposição se reverte, uma vez que, ao se gramaticalizar, uma forma passa a coexistir, em um mesmo domínio

²⁹ “Nem toda variabilidade e heterogeneidade na estrutura linguística implica mudança; mas toda mudança implica variabilidade e heterogeneidade” (WEINREICH; LABOV; HERZOG, 2006[1968], p. 126).

funcional, com outras que cumprem a mesma função, instaurando um típico caso de variação por estratificação funcional (HOPPER, 1991).

Contudo, os autores declaram que esse ponto de divergência entre a TVML e a Teoria da Gramaticalização (TG) não compromete a correlação entre as duas teorias, como demonstramos na seção referente ao Sociofuncionalismo (cf. 2.2). Sendo assim, embora não seja o objetivo deste estudo se ater ao histórico de discussões acerca da gramaticalização dos RAD ou desenvolver um exame minucioso sobre o percurso de mudança das formas em análise, discorreremos, de maneira sucinta, a respeito desse processo com o intuito de complementar a investigação desenvolvida neste trabalho. Nesse sentido, é extremamente relevante pensar na noção de gramaticalização e no fato de que os MD resultam desse processo.

No início do século XX, Antoine Meillet cunhou o termo gramaticalização (TRAUGOTT; HEINE, 1991; HOPPER; TRAUGOTT, 1993); segundo o estudioso, esse processo se dá quando um caráter gramatical é atribuído a um elemento antes autônomo. De acordo com Hopper e Traugott (1993), antes das análises desenvolvidas por Meillet, já havia estudos que se dedicavam à investigação desse processo, sem, contudo, usar a denominação *gramaticalização*, que seria empregada posteriormente. Ainda assim, para os autores, foi Meillet quem primeiramente reconheceu a relevância do processo de gramaticalização para os estudos linguísticos, cunhou o termo e dedicou estudos ao fenômeno. O linguista alterou o rumo que os estudos sobre a gramaticalização estavam tomando, tirando de foco a origem das formas gramaticais e trazendo à tona as transformações que essas formas sofriam. Meillet descreve dois processos pelos quais novas formas gramaticais podem emergir: *analogia* e *gramaticalização*. No primeiro caso, novos paradigmas tornam-se formalmente semelhantes a paradigmas já existentes, ao passo que, no segundo caso, uma palavra autônoma assume o papel de um elemento gramatical.

Na primeira metade do século XX, muitos linguistas assumiram um ponto de vista sincrônico em seus estudos, deixando fatores históricos, dentre os quais se encontram a gramaticalização, em segundo plano. Entretanto, Benveniste, aluno de Meillet, retoma o tópico em seus estudos, fazendo com que o tema volte a ser alvo de diversos trabalhos. A partir disso, o interesse pelo processo de gramaticalização se avolumou, acompanhado pelo aumento de estudos que tratavam do tópico, sob perspectivas sincrônica ou diacrônica, e rodeado por um caloroso debate a respeito de princípios teóricos envolvidos no processo. E, como era de se esperar, surgiram diversas e variadas teorias sobre a gramaticalização (HOPPER; TRAUGOTT, 1993).

Para Hopper e Traugott (1993), o termo gramaticalização possui dois significados: um deles está relacionado ao quadro de investigação adotado para investigar fenômenos linguísticos e o outro está ligado aos próprios fenômenos linguísticos. No primeiro caso,

[...] ‘gramaticalização’ se refere àquela parte do estudo da mudança linguística que se preocupa com questões do tipo: como itens e construções lexicais são introduzidos em determinados contextos linguísticos para servir a funções gramaticais ou como itens gramaticais desenvolvem novas funções gramaticais³⁰ (HOPPER; TRAUGOTT, 1993, p. 1, tradução nossa).

No segundo caso, “‘gramaticalização’ se refere mais especificamente aos passos pelos quais alguns itens tornam-se mais gramaticais ao longo do tempo”³¹ (HOPPER; TRAUGOTT, 1993, p. 2, tradução nossa). Traugott e Heine (1991), tratando da relação entre gramaticalização e mudança linguística, atestam que a gramaticalização é um tipo de mudança, mas nem todo processo de mudança é considerado gramaticalização. Em outra passagem do texto, os autores consideram as condições para que uma forma seja gramaticalizada; de acordo com eles, apenas algumas classes lexicais são suscetíveis a sofrer gramaticalização. Nas palavras dos linguistas, “para qualquer domínio gramatical, existe apenas um conjunto restrito de campos lexicais e, dentro deles, apenas um conjunto restrito de itens lexicais, que provavelmente serão fontes” (TRAUGOTT; HEINE, 1991, p. 8, tradução nossa)³². Há, portanto, três condições para que a gramaticalização aconteça: o contexto semântico, a saliência e a frequência. Em consonância com Traugott e Heine (1991), a frequência é o aspecto que, de fato, leva à gramaticalização, promovendo a fixação da forma, e, quanto mais gramaticalizada uma forma está, mais recorrente ela pode se tornar. Nessa mesma direção, Tavares (2013) menciona que a frequência e a rotinização de certas formas caracterizam o processo de gramaticalização, e Zanin e Gonçalves (2022, p. 67, grifos nossos) explicam que

[...] a mudança é disseminada gradualmente ao longo dos espectros linguístico e social, com incrementações contínuas em termos de **frequência de uso**, que, na perspectiva funcionalista, é fundamental para o estabelecimento e a manutenção da gramática, além de permitir também captar a difusão linguística e social da mudança. Na perspectiva variacionista, o **aumento de frequência** também é compreendido como

³⁰ “[...] ‘grammaticalization’ refers to that part of the study of language change that is concerned with such questions as how lexical items and constructions come in certain linguistic contexts to serve grammatical functions or how grammatical items develop new grammatical functions” (HOPPER; TRAUGOTT, 1993, p. 1).

³¹ “[...] ‘grammaticalization’ refers most especially to the steps whereby particular items become more grammatical through time” (HOPPER; TRAUGOTT, 1993, p. 2).

³² “[...] for any given grammatical domain, there is only a restrictive set of lexical fields, and within them only a restricted set of lexical items, that are likely to be sources” (TRAUGOTT; HEINE, 1991, p. 8).

índice de difusão linguística e social, e as variantes devem ter certa recorrência para que possam ser comparadas por meio de instrumental estatístico.

Posto isso, um tema frequente em estudos referentes à gramaticalização é a unidirecionalidade desse processo. Ao discutirem o assunto, Traugott e Heine (1991) destacam o aumento da abstratização. Segundo os autores, nos estágios iniciais da gramaticalização, há um aumento do significado pragmático, e, à medida que esse aumento ocorre, o conteúdo semântico pode sofrer uma certa redução. No entanto, “nem sempre é um caso simples de mais > menos semântico, mas, sim, de uma hierarquia implicacional do tipo semântico > pragmático > menos semântico-pragmático” (TRAUGOTT; HEINE, 1991, p. 5, tradução nossa)³³, o que corresponde a uma caracterização unidirecional, uma vez que, de acordo com eles, não se espera o processo inverso (pragmático > semântico > menos semântico-pragmático). Traugott (1982 apud HOPPER; TRAUGOTT, 1993) aponta que há fatores semânticos e pragmáticos envolvidos no processo de gramaticalização que levam à unidirecionalidade da mudança, especialmente quando há um percurso do significado mais concreto para o mais abstrato. Martelotta, Votre e Cezario (1996, p. 26), partindo das ideias do realismo experiencialista, sugerem que, a princípio, o pensamento “trabalha com conceitos adquiridos pelo contato com o mundo concreto” e, a partir dessa experiência, passa-se a compreender a realidade abstrata, isto é, o mundo das ideias. Os autores declaram que a metáfora é responsável pelo entendimento do mundo das ideias com base no mundo concreto; logo, “o processo metafórico tende a obedecer de modo unilateral a uma trajetória + concreto > - concreto [...]” (MARTELOTTA; VOTRE; CEZARIO, 1996, p. 26). Segundo os autores, o processo de gramaticalização resulta em noções mais abstratas. Traugott e Heine (1991) afirmam que essas mudanças de significado caracterizam não apenas a gramaticalização, mas o processo geral de mudança de significado. Ademais, os autores concluem que, apesar de o conceito de unidirecionalidade sugerir que haja apenas um caminho possível para a evolução, estudos que fazem referência a múltiplos domínios funcionais parecem mais próximos da realidade.

Além da gramaticalização, alguns autores apontam outro processo de mudança, conhecido como discursivização (MARTELOTTA; VOTRE; CEZARIO, 1996; GÖRSKI *et al.*, 2002). Em consonância com Görski *et al.* (2002), no processo de gramaticalização, construções e itens lexicais assumem um novo estatuto gramatical e, no processo de

³³ “It is not always a simple case of more > less semantic, but more often of an implicational hierarchy of the type semantic > pragmatic > less semantic-pragmatic” (TRAUGOTT; HEINE, 1991, p. 5).

discursivização, os elementos passam a assumir função de MD. De maneira mais detalhada, Martelotta, Votre e Cezario (1996, p. 24) definem ambos os processos da seguinte forma:

Gramaticalização e discursivização constituem processos especiais de mudança linguística. Gramaticalização leva itens lexicais e construções sintáticas a assumir funções referentes à organização interna do discurso ou a estratégias comunicativas. Discursivização leva o item a assumir função de marcador discursivo, modalizando ou reorganizando a produção da fala, quando a sua linearidade é momentaneamente perdida, ou servindo para preencher o vazio causado por essa perda da linearidade.

Na contramão dessa ideia, Traugott (1995) considera desnecessário haver dois processos de mudança, defendendo que os MD podem ser abrangidos pela gramaticalização, que, segundo ela, “é o processo por meio do qual materiais lexicais, altamente compelidos por contextos pragmáticos e morfossintáticos, tornam-se gramaticais”³⁴ (TRAUGOTT, 1995, p. 1, tradução nossa). Em concordância com Traugott (1995), entendemos que a mudança dos verbos que passam a atuar como RAD pode ser descrita e explicada pelo processo de gramaticalização, conforme demonstrado por estudos realizados anteriormente (ROST SNICHELOTTO, 2009; VALLE, 2014; ZANIN; GONÇALVES, 2022; entre outros).

Freitag (2009), consentindo com Traugott (1995), admite que os MD decorrem do processo de gramaticalização e não do de discursivização. Diante disso, focalizando os RAD, Freitag (2009, p. 7, grifos nossos) explica que a trajetória de mudança desses elementos envolve a abstratização, que tem início “em contextos de **perguntas plenas**, passando ao uso em **contextos totalmente interrogativos**, **pergunta semi-retórica**, em que o falante responde à sua própria pergunta, e por fim, a construção encontra-se em **contexto interrogativo totalmente retórico**”.

Para entender como esses MD se constituem, recorreremos aos cinco princípios de gramaticalização propostos por Hopper (1991), a saber: i) **estratificação**; ii) **divergência**; iii) **especialização**; iv) **persistência**; e v) **descategorização**. Esses princípios, segundo o autor, podem ser aplicados também aos estágios incipientes da gramaticalização e não apenas aos mais avançados.

A **estratificação** diz respeito à variação linguística, uma vez que as novas camadas que surgem dentro de um domínio funcional coabitam e coexistem com as camadas mais antigas. Neste trabalho, estamos lidando com o domínio funcional dos MD, do qual fizemos o recorte de um subdomínio, o dos RAD. Conforme ilustramos anteriormente, assim como o

³⁴ “[...] grammaticalization is the process whereby lexical material in highly constrained pragmatic and morphosyntactic contexts becomes grammatical” (TRAUGOTT, 1995, p. 1).

conjunto dos MD, que é extremamente amplo e heterogêneo, o subconjunto dos RAD comporta as mais variadas unidades. Dessa forma, além dos RAD investigados neste trabalho, há outros MD nesse subdomínio, os quais também resultam do processo de gramaticalização. Vejamos alguns exemplos:

(11)

KATIA — E além de tudo engraçadinho, **né?** Tenho lá eu culpa se você é estudante ou não? (*O cordão umbilical* – Mário Prata, 1970).

(12)

ARTHUR – Olha, eu vou até aquela churrascaria da esquina. Eu te dou duas horas. Duas horas, **tá?** (*Santidade* – José Vicente de Paula, 1967).

(13)

PIMENTEL – Zulmira, **é?** Espera, que eu vou fechar a porta (*A falecida* – Nelson Rodrigues, 1953).

Por meio desses exemplos, percebemos que *né?*, *tá?* e *é?*, apesar de atuarem em contextos e com funções semelhantes aos dos RAD em estudo, possuem naturezas conceptuais distintas. Tendo em vista que o nosso objetivo é dispensar um tratamento variável aos RAD, demonstrando a sua intercambialidade, optamos por formas que, além de atuarem no mesmo domínio funcional, são semelhantes quanto à classe gramatical de origem, à natureza conceptual e ao percurso de mudança.

O princípio da **divergência** sugere que, quando um elemento sofre gramaticalização, sua forma lexical original pode sofrer as mesmas alterações, constituindo, assim, formas etimologicamente semelhantes, mas funcionalmente discrepantes. No caso dos RAD em estudo, os verbos dos quais eles se originaram continuam em uso na língua; entretanto, esses verbos não se confundem com os RAD, conforme ilustrado a seguir:

(14)

BISPO - Moça, não se faça de louca. No seu caso não terá sucesso. Como não **compreende?** Você foi o nosso termômetro. [...] (*A empresa ou A possessa* – Hilda Hilst, 1967).

(15)

PEDRO - Pata, seu Vigário. Ainda não **entendeu?** (*A nova Helena* – Francisco Pereira da Silva, 1957).

(16)

PATRÍCIO — [...] Você tem que se valorizar. Senão o cara te chuta. Será que você não **percebe**? (*Toda nudez será castigada* – Nelson Rodrigues, 1965).

(17)

SECRETÁRIO – [...] O “Boca de Ouro” matou gente pra burro e quem sabe se ela não conta a você, com exclusividade, uma dessas mortes, um crime bacana? Hem, quem **sabe**? (*Boca de ouro* – Nelson Rodrigues, 1959).

(18)

ALZIRA – Era a Esmeralda. (Para ele) A Esmeralda **viu**??? (*O cão siamês ou Alzira Power* – Antônio Bivar, 1969).

Diferentemente dos RAD, nesses trechos os verbos não atuam na articulação textual ou na orientação da interação, não são sintaticamente independentes, não são exteriores ao conteúdo proposicional e, no tocante à transparência semântica, são usados no sentido denotativo, sendo, portanto, totalmente transparentes. Além disso, essas formas podem ser flexionadas para se adequar ao discurso, o que não ocorre com os RAD, que apresentam pouca ou nenhuma variabilidade (RISSO; SILVA; URBANO, 2019).

Feitas essas observações, no caso dos RAD em estudo, estamos considerando tanto formas simples, formadas por uma só palavra (*entendeu?*, *sabe?*, *compreende?*), quanto formas compostas, acompanhadas por outras unidades linguísticas, como advérbios (*compreendeu agora?*, *entendeu bem?*), verbos (*tá percebendo?*, *tá sabendo?*), pronomes (*você sabe?*, *me entende?*) ou outros MD (*mas compreendeu?*, *sabe*, *né?*). Consideremos os exemplos abaixo:

(19)

ALZIRA - [...] Eu nunca fui atrás e nem precisei de ninguém pra me satisfazer sexualmente, **está me entendendo?** Eu sempre resolvi o meu problema sozinha. [...] (*O cão siamês ou Alzira Power* - Antônio Bivar, 1969).

(20)

KÁTIA - Mora mais gente aqui, é? Tá pensando que eu sou do que, hein? Não sou destas, não, **tá sabendo?** (*O cordão umbilical* - Mário Prata, 1970).

(21)

HOMEM - [...] Eu não sou muuuuito macho não... (risada de Mariazinha) Assim de dar cinco... **entende, né?** Assim eu não sou não. (*Fala baixo, senão eu grito* – Leilah Assumpção, 1969).

(22)

PATRÍCIO - Ouve, Serginho, ouve a minha ideia. Passei a noite em claro, só pensando o seguinte: — teu pai se casar com a Geni.

SERGINHO - Com uma prostituta?

PATRÍCIO - Pois teu pai vai ser o marido e a prostituta vai ser a esposa!

SERGINHO - Esposa, como minha mãe?

PATRÍCIO - Esse casamento é preciso, sabe por quê? Porque você vai cornear seu pai!

Compreendeu agora?

SERGINHO - Tenho nojo dessa mulher! (*Toda nudez será castigada* - Nelson Rodrigues, 1965).

Com base nos exemplos dados, nota-se que as formas compostas podem ser caracterizadas pela carga semântica que ainda carregam, pela massa fônica proeminente e pela sujeição a flexões gramaticais (RISSO; SILVA; URBANO, 2019; ZANIN; GONÇALVES, 2022). Essas formas, então, não atingiram o estatuto de MD, mas não atuam mais como verbos plenos. Gonçalves (2021) explica que uma forma linguística pode ter diferentes graus de gramaticalização, já que, por se tratar de um processo histórico, a gramaticalização ocorre gradualmente, o que justifica o fato de algumas formas estarem totalmente gramaticalizadas e de outras se localizarem em estágios iniciais do processo. As formas compostas, portanto, estariam em estágios iniciais ou intermediários desse processo.

É por esse motivo que, neste trabalho, partimos do *continuum* proposto por Rizzo, Silva e Urbano (2019), que consideram a existência de MD mais ou menos prototípicos. Desse modo, defendemos que as formas compostas são marcadores menos prototípicos, pois, mesmo não sendo totalmente gramaticalizadas, apresentam características comuns aos RAD. Levando em consideração os traços identificadores dos MD (cf. Quadro 3), observamos que nem todas essas propriedades se aplicam às formas compostas, como a alta recorrência, a invariabilidade formal ou variabilidade restrita e a massa fônica reduzida. Por outro lado, podemos afirmar que as formas compostas são *exteriores ao conteúdo proposicional*, pois não contribuem para o significado referencial da proposição; são *sintaticamente independentes*, visto que não estão presas à estrutura oracional; *não possuem autonomia comunicativa*, porque dependem de um contexto para que façam sentido; e *atuam nos planos*

textual e interacional, focalizando determinadas porções textuais e chamando a atenção do interlocutor.

Quadro 3 – Traços identificadores dos MD

Traços identificadores
Traço 1: alta recorrência
Traço 2: exterioridade ao conteúdo proposicional
Traço 3: transparência semântica parcial
Traço 4: invariabilidade formal ou variabilidade restrita
Traço 5: independência sintática
Traço 6: demarcação prosódica
Traço 7: não autonomia comunicativa
Traço 8: massa fônica reduzida
Traço 9: articulação tópica
Traço 10: orientação interacional

Fonte: Adaptado de Risso, Silva e Urbano (2019, p. 381)

É interessante pontuar que os traços destacados anteriormente compõem o núcleo-piloto dos MD, que são os traços mais recorrentes desses elementos, conforme ilustrado no quadro 4. Dessa maneira, sugerimos que as formas compostas não operam mais como verbos plenos e já assumiram certas particularidades e funções de MD.

Quadro 4 – Núcleo-piloto de MD proposto pela GTI

Núcleo-piloto
Traço 1: Exterioridade ao conteúdo proposicional
Traço 2: Independência sintática
Traço 3: Ausência de autossuficiência comunicativa
Traço 4: Maior projeção da articulação tópica
Traço 5: Maior projeção da orientação da interação

Fonte: Adaptado de Risso, Silva e Urbano (2019, p. 386)

A **especialização** diz respeito ao estreitamento das escolhas de formas que pertencem a um mesmo domínio, ou seja, uma forma vai se tornando mais corriqueira e cristalizada. O tratamento variável dispensado aos RAD em estudo nos permitirá aferir contextos de especialização, posto que esses contextos são controlados em função de variáveis linguísticas e extralinguísticas.

A **persistência** é o princípio relacionado à manutenção de certos traços semânticos, pois, ao sofrer gramaticalização, uma forma pode manter alguns traços semânticos da forma de origem, dando indícios de seu percurso. É possível, então, que os RAD mantenham determinados traços das fontes verbais das quais se originaram. Como vimos há pouco,

algumas formas compostas carregam, ainda, traços morfossintáticos dos verbos fonte, variando quanto ao tempo, ao modo, à pessoa etc.

O princípio da **descategorização** sugere que as formas em processo de gramaticalização perdem os traços de categorias plenas, como substantivos e verbos, e assumem características de categorias secundárias, como adjetivos. Os RAD, ao contrário de seus verbos de origem, adquirem particularidades específicas, voltando-se à manutenção, ao planejamento e à organização do discurso (FREITAG; SILVA; EVANGELISTA, 2017); logo, à medida que os RAD tornam-se mais gramaticalizados, as marcas que caracterizavam os verbos começam a ser neutralizadas.

Diante disso, aplicando os princípios de Hopper (1991) aos marcadores sob análise, nota-se que os RAD originados de *compreender*, *entender*, *perceber*, *saber* e *ver* variam no domínio da requisição de apoio discursivo, coexistindo dentro de um mesmo domínio funcional (**princípio da estratificação**). Em alguns contextos, a variação pode deixar de acontecer, com a fixação de apenas uma das formas em coexistência (**princípio da especialização**). Além disso, as fontes das quais os RAD em estudo se originam permanecem em uso no sistema linguístico, mesmo com o surgimento dos RAD; todavia, esses verbos, ainda que etimologicamente semelhantes aos RAD deles provenientes, possuem diferentes funções (**princípio da divergência**). Nesse sentido, traços semânticos da fonte verbal podem ser mantidos nos RAD, isto é, na forma gramaticalizada (**princípio da persistência**). Por fim, os RAD perdem as características de verbos plenos e passam a assumir propriedades de marcadores que atuam no domínio da requisição de apoio discursivo (**princípio da descategorização**).

Lançando um olhar minucioso sobre a gramaticalização de *entender* e de *saber*, Valle (2014) aponta que os RAD oriundos de *saber* estariam mais gramaticalizados, com atuações voltadas à organização textual; os RAD provenientes de *entender*, em contraponto, estariam mais apegados aos significados fonte e às estratégias discursivas originais, sendo, assim, menos gramaticalizados. Tendo como parâmetro a redução/extensão e a fixidez das formas em estudo, bem como sua rotinização e frequência de uso, poderíamos afirmar que *ver*, por ser uma forma fixa, de uso generalizado e recorrente em nosso *corpus*, estaria avançado no processo de gramaticalização. *Saber* também estaria consideravelmente gramaticalizado, pois, apesar de apresentar alguma variabilidade formal, é mais usado na terceira pessoa do singular do presente do indicativo, sua forma mais comum (URBANO, 2019; ZANIN; GONÇALVES, 2022). *Compreender* e *entender*, por outro lado, estariam menos gramaticalizados, porque apresentam, ainda, algumas formas compostas, mais extensas e menos rotinizadas. Os RAD

advindos de *perceber* não são frequentes no *corpus*, mas, por apresentarem pouca variabilidade formal, é provável que estejam mais gramaticalizados do que *compreender* e *entender*.

Essa descrição corresponde a uma síntese de como os RAD se constituem, posto que não vamos nos ater à análise pormenorizada do processo de gramaticalização desses elementos. No entanto, a partir disso, é possível apreender de que modo os verbos de cognição tornaram-se RAD.

2.7 Sintetizando...

Neste capítulo, fizemos uma revisão dos subsídios teóricos do presente trabalho. Primeiramente, discutimos os pressupostos teórico-metodológicos da Sociolinguística Variacionista, apresentando os cinco problemas empíricos propostos por Weinreich, Labov e Herzog (2006[1968]) e retomando alguns conceitos essenciais dessa abordagem de estudos linguísticos. Em seguida, dissertamos a respeito do Sociofuncionalismo, que concilia a TVML e a Teoria da Gramaticalização (TG), e de sua relevância para o estudo dos RAD. Discorremos, também, sobre a Sociolinguística História, demonstrando a pertinência de fontes escritas para o estudo histórico da língua. Na sequência, fizemos algumas considerações referentes ao teatro brasileiro nas décadas de 1950 a 1970, traçando um panorama geral dos aspectos que marcaram esse período, e defendemos o uso das peças teatrais como fonte de análise para investigações linguísticas.

Concluídas essas discussões, voltamos o nosso foco para os MD, mencionando a heterogeneidade e a amplitude desse conjunto, que possui diferentes rótulos e é objeto de diferentes abordagens. Logo depois, nos ativemos aos estudos de Schiffrin (1987, 2001, 2006) e de Risso, Silva e Urbano (2019), abordagens adotadas neste trabalho, e estabelecemos um diálogo entre as duas perspectivas, revelando que é possível correlacioná-las. Ademais, explicamos como e por que delimitamos o objeto de estudo e sintetizamos alguns trabalhos para identificar as funções e as particularidades dos RAD. Por fim, buscando complementar a investigação desenvolvida neste trabalho, tratamos do processo de gramaticalização dos RAD com o intuito de esclarecer de que modo verbos passam a atuar como MD.

3 METODOLOGIA

Este capítulo destina-se ao detalhamento das etapas da pesquisa e dos procedimentos metodológicos adotados. Primeiramente, discorreremos acerca do percurso trilhado até a obtenção dos resultados, descrevendo cada uma das etapas da investigação. Nas subseções subsequentes, apresentamos o *corpus*, especificando quais peças e autores foram selecionados para a análise, descrevemos os procedimentos adotados para a realização da coleta e do tratamento dos dados e, finalmente, esmiuçamos as variáveis controladas.

3.1 As etapas da pesquisa

O primeiro passo para a execução desta pesquisa foi a realização de um levantamento bibliográfico. Fizemos pesquisas e discussões sobre estudos e obras que já tenham contemplado os MD, reunimos informações referentes ao teatro e ao conjunto de fatos históricos relacionados ao recorte temporal proposto (1950-1970) e selecionamos obras de cunho sociolinguístico, sócio-histórico e sociofuncionalista para fundamentar esta pesquisa.

Para compor o *corpus* do trabalho, selecionamos peças de teatro produzidas por autores brasileiros em meados do século XX, mais especificamente entre as décadas de 1950 e 1970. Os detalhes acerca das peças que constituem esse *corpus* estão contidos nas próximas subseções. Na sequência, partindo da leitura das peças selecionadas, realizamos uma categorização do perfil social dos personagens. Para tal fim, partimos da categorização desenvolvida pelo *Núcleo de Pesquisas em Sociolinguística de Araraquara* (SoLAr), no âmbito do projeto intitulado *Língua EmCena*, que tem por objetivo categorizar os perfis sociais de personagens de peças teatrais brasileiras produzidas nos séculos XIX, XX e XXI com a finalidade de correlacioná-los a fenômenos linguísticos. O grupo, partindo de fontes históricas, propôs, para identificação do perfil social dos personagens, cinco variáveis, a saber: gênero/sexo, fase da vida, instrução, classe social e etnia, as quais são utilizadas na categorização deste trabalho. Ademais, a presente pesquisa, somando-se a estudos prévios (cf. BRANDÃO, 2022) e em desenvolvimento, traz contribuições ao referido projeto, demonstrando que as peças são materiais relevantes para os estudos (socio)linguísticos.

Em seguida, partimos para a coleta dos RAD nas peças teatrais. Não obstante tenhamos contado com o auxílio de mecanismos de pesquisa e de busca textual pertencentes a visualizadores de textos, a coleta foi executada, quase integralmente, de maneira manual, mediante a leitura de cada uma das obras escolhidas.

Definidas as variáveis linguísticas e extralinguísticas que possam motivar a escolha e o uso dos RAD, partimos para a análise e para a quantificação dos dados. Desse modo, os resultados obtidos mediante o levantamento foram processados estatisticamente por meio do programa R (CORE TEAM, 2022) e interpretados a partir das hipóteses e dos pressupostos teóricos.

3.2 O *corpus*: peças de teatro produzidas entre as décadas de 1950 e 1970

O *corpus* do trabalho é composto por peças teatrais produzidas por escritores brasileiros entre as décadas de 1950 e 1970. Para selecionar as peças, consideramos alguns critérios: o período de produção, a autoria e a temática dessas peças, as quais deveriam ser condizentes com a realidade da época. Cumpre destacar que a extensão das peças não foi um critério a ser considerado na seleção dessas obras, pois realizamos uma delimitação bastante precisa do objeto de estudo, e o foco da pesquisa é, justamente, verificar quais forças linguísticas e extralinguísticas estão envolvidas no uso dos RAD em estudo.

Assim sendo, foram selecionadas 39 peças, produzidas por 14 autores diferentes, a saber: Abílio Pereira de Almeida; Antônio Bivar; Ariano Suassuna; Augusto Boal; Consuelo de Castro; Dias Gomes; Francisco Pereira da Silva; Hilda Hilst; José Vicente de Paula; Leilah Assumpção; Mário Prata; Nelson Rodrigues; Oduvaldo Vianna Filho e Plínio Marcos. No quadro abaixo, apresentamos as peças, seus autores e o ano de publicação dessas obras:

Quadro 5 - O *corpus*

Autores	Peças	Datas
Abílio Pereira de Almeida	Em moeda corrente do país	1960
Antônio Bivar	Abre a janela e deixa entrar o ar puro e o sol da manhã	1968
	Cordélia Brasil	1967
	O cão siamês ou Alzira Power	1969
Ariano Suassuna	A pena e a lei	1959
	Auto da compadecida	1955
	Farsa da boa preguiça	1960
	O arco desolado	1952
	O casamento suspeitoso	1957
	O rico avarento	1954
	O santo e a porca	1957
	O seguro	1957
Augusto Boal	O cavalo e o santo	1954
Consuelo de Castro	À prova de fogo	1968

Dias Gomes	A invasão	1960
	O pagador de promessas	1960
Francisco Pereira da Silva	A nova Helena	1957
	Cristo proclamado	1958
Hilda Hilst	A empresa ou A possessa	1967
	As aves da noite	1968
	O novo sistema	1968
	O verdugo	1969
	O visitante	1968
José Vicente de Paula	O assalto	1969
	Os convalescentes	1970
	Santidade	1967
Leilah Assumpção	Fala baixo, senão eu grito	1969
Mário Prata	O cordão umbilical	1970
Nelson Rodrigues	A falecida	1953
	Boca de ouro	1959
	O beijo no asfalto	1960
	Os sete gatinhos	1958
	Otto Lara Resende ou Bonitinha, mas ordinária	1962
	Perdoa-me por me traíres	1957
	Toda nudez será castigada	1965
	Viúva, porém honesta	1957
Oduvaldo Vianna Filho	Chapetuba Futebol Clube	1959
Plínio Marcos	Dois perdidos numa noite suja	1966
	Homens de papel	1968

Fonte: própria

Das peças escolhidas, algumas foram extraídas de acervos disponíveis no meio virtual, voltados para fins educacionais ou para amantes de teatro. Dentre esses acervos, destacamos a *Coleção Aplauso*³⁵, organizada pela Imprensa Oficial, que se dedica à preservação da memória da cultura nacional, homenageando artistas, dramaturgos e roteiristas diversos, e disponibilizando, além da biografia desses expoentes da cultura, a íntegra de peças teatrais e de roteiros de cinema. As demais peças foram retiradas de um acervo organizado pelo SoLAR, no interior do, já referido, projeto *Língua EmCena*. As peças brasileiras, que se distribuem ao longo dos séculos XIX, XX e XXI, foram formatadas e categorizadas pelos integrantes do grupo cujo intuito é que essas obras sejam utilizadas em pesquisas (socio)linguísticas e

³⁵ Disponível em: <https://aplauso.imprensaoficial.com.br/>

sócio-históricas. Nesse sentido, parte-se da periodização proposta por historiadores que integraram a coleção *História do Brasil Nação*: a) **Período 1:** 1808-1830; b) **Período 2:** 1831-1889; c) **Período 3:** 1890-1930; d) **Período 4:** 1931-1964; e e) **Período 5:** 1965-2010. É válido destacar que o grupo principia da década de 1831 – que corresponde ao segundo período, uma vez que há poucos registros de produção de peças teatrais por autores brasileiros antes desse momento (LEITE, 2022) – e ultrapassa o ano de 2010, incluindo peças mais contemporâneas no *corpus*, produzidas até 2020.

As categorizações dos perfis sociais dos personagens desenvolvidas pela equipe também se fundamentam em estudos históricos acerca do Brasil e de sua sociedade. Cada uma das variáveis propostas para a categorização (gênero/sexo, fase da vida, instrução, classe social e etnia) foi estudada de maneira aprofundada pelos integrantes do grupo, de modo a compreender como a sociedade era estratificada nesses períodos, e, a partir disso, elaboraram-se critérios para categorizar o perfil social dos personagens, conforme exposto no quadro 6.

Quadro 6 - Variáveis e critérios propostos pelo projeto *Língua EmCena*³⁶

Variáveis	Gênero/sexo	Feminino Masculino
	Fase da vida	Fase 1 / Fase 2 / Fase 3 Fase 1 / Fase 2a / Fase 2b / Fase 3
	Instrução	Sem instrução Instrução básica Instrução média Instrução avançada
	Classe social	Classe 1 Classe 2 Classe 3 Classe 4
	Etnia	Branco Não branco

Fonte: própria

Essas variáveis independentes, então, permitem caracterizar o perfil social dos personagens, que será detalhado mais adiante. O projeto reúne peças representativas das sociedades de épocas anteriores e grandes nomes do teatro brasileiro – Martins Pena, França

³⁶ Com a finalidade de se adequar à realidade social dos períodos investigados, a variável fase da vida sofre uma ramificação; por essa razão, há duas propostas de categorização: até 1964, os personagens são categorizados em fases 1, 2 e 3 e, após esse período, os personagens passam a ser categorizados em fases 1, 2a, 2b e 3.

Júnior, Ariano Suassuna, Coelho Neto, Júlia Lopes de Almeida, Dias Gomes, Joracy Camargo, Mário Bortolotto, Nelson Rodrigues etc.

3.3 Coleta e tratamento dos dados

Atualmente, há inúmeros programas que realizam buscas textuais, de modo a facilitar o trabalho do pesquisador, principalmente quando o fenômeno investigado é extremamente recorrente na amostra delimitada. Entretanto, neste estudo, nós partimos de uma coleta manual e recorremos aos próprios recursos de pesquisa e de busca textual de visualizadores de textos, como *Adobe Acrobat* e *Word*. Visto que os verbos dos quais os RAD em estudo derivam continuam a ser usados na língua, foi necessário se atentar aos contextos de uso, de modo a não confundir o verbo pleno com o MD. A exemplo disso, os trechos abaixo mostram *saber* atuando como RAD (23) e como verbo pleno (24 e 25).

(23)

JÚLIA – Zé... eu estou numa confusão... se você imaginasse... Eu tento te esquecer... eu tento, mas eu não consigo. Por mais que eu faça força, não consigo te tirar da cabeça, **sabe?** [...]. (*À prova de fogo* – Consuelo de Castro, 1968).

(24)

ZÉ – Estou sim. Mas agora eu já estou legal. Você me acalma. **Sabe** que tem os olhos mais bonitos que já vi? (*À prova de fogo* – Consuelo de Castro, 1968).

(25)

JÚLIA – [...] Desde que você e a Rosa entraram aqui, o Zé não para de paquerar a menina. E olha que ele é fogo, hein. Chegou, doutrinou, comeu. **Sabe como?** Ele diz que o preconceito sexual atrasa a revolução. (*À prova de fogo* – Consuelo de Castro, 1968).

Durante a leitura das peças, os RAD em estudo foram destacados e, na sequência, mediante uma busca textual, verificamos se havia outras ocorrências do fenômeno. A leitura, o destaque dos RAD e a verificação por meio do localizador de palavras foram realizados em todas as peças.

3.4 Variáveis dependente e independentes

O nosso objeto de estudo, conforme esclarecido previamente, são os RAD originados de *compreender*, *entender*, *perceber*, *saber* e *ver*. Essas formas, por compartilharem certas

características e serem intercambiáveis, são tratadas como variantes que atuam em um mesmo domínio funcional, isto é, na requisição de apoio discursivo.

Nossa variável dependente é controlada a partir das fontes verbais das quais os RAD em estudo se originaram, de modo que todas as variantes desses elementos sejam consideradas na análise (cf. Quadro 7).

Quadro 7 - Os RAD considerados na análise

RAD	Fontes verbais	Variantes
	Compreender	Exemplos: <i>compreendeu? compreende? compreendestes?</i>
	Entender	Exemplos: <i>tá entendendo? entendeu? entende?</i>
	Perceber	Exemplos: <i>percebeu? percebete? tá percebendo?</i>
	Saber	Exemplos: <i>sabe? sabia? tá sabendo?</i>
	Ver	Exemplos: <i>viu?</i> ³⁷

Fonte: própria

Assim como ilustramos na introdução, os RAD derivados desses verbos podem ser substituídos não apenas por suas variantes, mas também uns pelos outros, sem que o sentido da proposição seja prejudicado. Vejamos alguns exemplos:

(26)

SIGISMUNDO – [...] Ouvi tudo, **entendes?** Tua conversa com essa prostituta e depois com teu pai. [...] (*O arco desolado* – Ariano Suassuna, 1952).

(27)

VITOR - [...] Estou te pagando a tua hora contada, marcada no despertador, estou fuçando numa peça de máquina pra obrigar a parar, **tá me entendendo?!** Junta tudo e se arranca! Eu aceito morrer por você e você vive por mim. [...] (*O assalto* – José Vicente de Paula, 1969).

(28)

VERDUGO - É muito difícil para mim. É assim como se eu tivesse que cortar uma árvore, **você entende?** Eu nunca derrubei uma árvore, eu não saberia, é difícil, não é o meu ofício (*O verdugo* – Hilda Hilst, 1969).

³⁷ Não encontramos variação dos RAD provenientes da fonte verbal *ver* em nosso corpus.

Temos, nesses três exemplos, variantes do RAD derivado de *entender*. Cada um dos trechos provêm de peças e de autores diferentes, mas, ao analisarmos esses excertos, percebemos que é possível usar as três variantes (*entendes?*, *tá me entendendo?* e *você entende?*) em cada contexto, sem perda de sentido proposicional, conforme ilustrado abaixo.

i) [...] Ouvi tudo, **tá me entendendo?** / **você entende?** Tua conversa com essa prostituta e depois com teu pai [...].

ii) [...] Estou te pagando a tua hora contada, marcada no despertador, estou fuçando numa peça de máquina pra obrigar a parar, **entendes?! / você entende?! Junta tudo e se arranca! Eu aceito morrer por você e você vive por mim [...].**

iii) É muito difícil para mim. É assim como se eu tivesse que cortar uma árvore, **entendes? / tá me entendendo?** Eu nunca derrubei uma árvore, eu não saberia, é difícil, não é o meu ofício.

É possível, também, substituir as variantes do RAD proveniente de *entender* pelas variantes dos outros RAD em estudo, evidenciando a intercambialidade dessas formas.

i) Ouvi tudo, **compreendes? / percebes? / sabes? / vês?** Tua conversa com essa prostituta e depois com teu pai [...].

ii) [...] Estou te pagando a tua hora contada, marcada no despertador, estou fuçando numa peça de máquina pra obrigar a parar, **tá compreendendo?! / tá percebendo?! / tá sabendo?! / tá vendo?! Junta tudo e se arranca! Eu aceito morrer por você e você vive por mim [...].**

iii) É muito difícil para mim. É assim como se eu tivesse que cortar uma árvore, **compreende? / percebe? / sabe? / viu?** Eu nunca derrubei uma árvore, eu não saberia, é difícil, não é o meu ofício.

Os exemplos confirmam a intercambialidade das formas, mas também demonstram que, embora o sentido da proposição seja mantido, pode haver uma perda de efeitos de sentido, porque a troca de uma forma pela outra poderia comprometer o sucesso da construção das *personas*, já que os RAD são utilizados para compor as falas de personagens com diferentes perfis sociais. Podemos concluir, então, que, apesar de serem equivalentes, esses elementos não necessariamente causariam o mesmo efeito se fossem substituídos uns pelos

outros, indicando que há uma distinção entre a macrofunção dos RAD e seus potenciais significados sociais.

Posto isso, baseando-nos nos resultados de estudos prévios e nos pressupostos teórico-metodológicos adotados neste trabalho, especificamos um conjunto de fatores linguísticos e extralinguísticos correlacionados ao fenômeno, os quais constituem as nossas variáveis independentes. As variáveis independentes de natureza linguística são: i) a apresentação formal dos RAD – simples ou composta; ii) a posição dos RAD no turno conversacional – medial ou final; e iii) a presença ou ausência de respostas após o uso dos RAD. Quanto às variáveis independentes de natureza social, estamos considerando: i) gênero/sexo, fase da vida, instrução, classe social e etnia, que caracterizam o perfil social dos personagens; ii) a relação entre os personagens/interlocutores – íntima/não íntima e simétrica/assimétrica; iii) a ambientação das peças; e iv) a autoria dessas obras. O quadro 8 sintetiza as variáveis independentes relacionadas ao fenômeno investigado.

Quadro 8 - Variáveis independentes

Variáveis linguísticas	Variáveis extralinguísticas
Apresentação formal	Perfil social dos personagens: ● Gênero/sexo ● Fase da vida ● Instrução ● Classe social ● Etnia
Posição	Relação entre os personagens/interlocutores
Presença/ausência de respostas	Ambientação
	Autoria

Fonte: própria

Iremos discorrer mais detidamente acerca das variáveis independentes nas próximas subseções, apresentando as nossas hipóteses a respeito de cada uma delas.

3.4.1 Apresentação formal

Observando estudos anteriores, percebemos que alguns pesquisadores apontam os RAD como formas simples, fixas e reduzidas (SILVA; MACEDO, 1996; VALLE, 2014; RISSO; SILVA; URBANO, 2019). Embora uma forma possa apresentar alguma variabilidade, como é o caso da maior parte dos RAD considerados em nossa análise, outras tendem a se tornar cada vez mais cristalizadas, sendo reproduzidas da mesma maneira por todos os

falantes, o que equivale ao caso de *viu?*, que não apresenta variação nas peças que compõem o *corpus*. Outro exemplo é o RAD *entendeu?*, que, por vezes, é utilizado na terceira pessoa do singular e no pretérito perfeito do indicativo; a tendência é que essa forma torne-se fixa ou sofra redução – o que já pode ser observado em usos como *tendeu?*³⁸, dando indícios de uma redução fonética desse RAD.

Assim, para que todos os dados fossem incluídos na análise, nós consideramos tanto *formas simples* quanto *formas compostas*. São exemplos de formas simples RAD formados por uma única expressão, constituindo formas curtas (*compreendeu?*, *sabe?*, *viu?*, *entende?*, *percebeu?* etc.).

(29)

ERNESTO - Eu sou um cidadão responsável, responsável, **entendeu?** Eu quero deixar isso bem claro (*O cão siamês ou Alzira Power* – Antônio Bivar, 1969).

(30)

DUDU – [...] Eu já vou embora, **sabe?** Boa noite (*O cavalo e o santo* – Augusto Boal, 1954).

As formas compostas, em contraponto, correspondem aos RAD acompanhados por elementos linguísticos como pronomes, verbos ou, até mesmo, outros MD (*você me compreende?*, *tá percebendo?*, *entende, né?* etc.). Ainda que essas formas não tenham atingido o estatuto pleno de marcador, elas já compartilham certos contextos de uso com os RAD, constituindo, portanto, casos menos prototípicos de MD.

(31)

ESTUDANTE – [...] No laboratório teu trabalho deve ser como uma oração, **você compreende?** Não espere respostas imediatas, entendeu? Olha, você já não é mais criança... Eu preciso te falar bem claro, não espera resposta... entrega-te... como uma oração (*As aves da noite* – Hilda Hilst, 1968).

(32)

DIDI – [...] se começar a me encher muito, te ponho pra fora. **Tá percebendo?** E para de andar de calcinha e soutien aqui dentro porque isto aqui é um apartamento de respeito. O que que a velha aí da frente não vai pensar? (*O cordão umbilical* – Mário Prata, 1970).

³⁸ Em Valle (2014), há ocorrências dessa forma.

Nossa hipótese em relação a essa categoria diz respeito à maior ou à menor integração da forma ao papel de RAD. Consideramos, então, que as formas simples estejam mais gramaticalizadas e, por conseguinte, apresentem mais propriedades particulares à categoria dos RAD (SILVA; MACEDO, 1996; VALLE, 2014; RISSO; SILVA; URBANO, 2019). Nesse caso, as formas simples seriam mais frequentes do que as formas compostas, que estariam, ainda, em processo de gramaticalização, rumo ao estatuto pleno de RAD.

3.4.2 Posição

Em se tratando da posição dos RAD, Silva e Macedo (1996), assim como Marcuschi (1986), apontam a preferência desses elementos pela posição final de enunciado ou de turno. Valle (2014), em contrapartida, abrangendo ainda mais a colocação dessas unidades no discurso, observa as seguintes posições: inicial; entre orações; entre constituintes; em final de enunciado; em final de turno e em meio a truncamento. Em trabalho anterior (VALLE, 2001), a autora verificou que os RAD tinham preferência pela posição intra-oracional, resultado que se mantém em sua tese (VALLE, 2014).

Neste trabalho, contudo, visando maior clareza de resultados, optamos apenas pelo turno conversacional como critério de análise da posição dos RAD. Por estarmos lidando com peças teatrais, a estrutura do turno é mais bem delimitada e definida, uma vez que, diferentemente da fala, não há pausas antes ou depois desses elementos e, tampouco, pode-se considerar a entonação da voz ou a presença de estímulos por parte do ouvinte. Berlinck e Brandão (2021, p. 236, grifo das autoras) ressaltam que,

[...] as *peças* têm como componente principal o diálogo entre personagens. Elas trazem uma representação, por certo, mas moldada na estrutura do texto falado. Assim, vemos a alternância de turnos dos interlocutores: cada fala pede uma reação, uma resposta, e essa é dada na sequência dos turnos conversacionais ali reconstruídos.

O turno pode ser definido como cada um dos segmentos produzidos pelos interlocutores em uma conversação e diz respeito à alternância desses indivíduos nos papéis de falante e de ouvinte (GALEMBECK, 1999; CASTILHO, 2019). Logo, nós controlamos os RAD em posições *medial* e *final* de turno, conforme exemplificado em (33) e (34).

(33)

LEÔNIDAS – [...] Eu tenho uma tia, **sabe?** Ela é uma das médiuns que mais recebem lá na minha cidade. [...] (*Cordélia Brasil* – Antônio Bivar, 1967).

(34)

MARLI – [...] É porque tenho prazer em ver você vestido com a roupa que eu dei, com os sapatos que eu comprei e com a carteira recheada de notas que eu ganhei pra você. Tenho orgulho, **sabe?** (*O pagador de promessas* – Dias Gomes, 1960).

É válido frisar que, inicialmente, havíamos considerado, além das posições medial e final, a posição inicial; no entanto, devido à quase inexistente ocorrência de RAD nessa posição, optamos por retirá-la da análise.

Em consonância com estudos anteriores (VALLE 2001, 2014; TRAPP, 2014), ao controlar essa variável, esperamos que os RAD ocorram mais em posição medial, favorecendo a sustentação do turno e destacando a atuação desses elementos no plano textual, embora a função interacional permaneça tangível, visto que atuam, também, no gerenciamento da interação. Esperamos, ainda, encontrar dados em posição final de turno, que, segundo as nossas expectativas, evidenciariam o caráter interacional dos RAD, bem como o envolvimento do falante com o ouvinte ou, nesse caso, dos personagens com seus interlocutores. Nossa expectativa é que formas compostas, por terem mais material linguístico, ocupem principalmente a posição final e que formas simples ocorram, predominantemente, intraturno.

3.4.3 Presença/ausência de respostas

Estudos prévios a respeito dos RAD demonstram que a maioria desses elementos não são acompanhados por estímulos³⁹ por parte do ouvinte (URBANO, 1999a; VALLE, 2001, 2014; TRAPP, 2014), o que reforçaria o caráter retórico dos RAD, em que não se espera uma manifestação verbal do interlocutor em resposta ao marcador.

Valle (2014), investigando uma amostra de fala, considera a presença ou ausência de estímulos antes ou depois dos RAD, bem como a presença de respostas a esses elementos. Posto que estamos lidando com textos escritos, não é possível investigar a presença de estímulos, os quais, muitas vezes, se sobrepõem ao enunciado do falante. Nesta categoria, então, estamos considerando os casos em que há respostas ao uso dos RAD; sendo assim, consideramos dois tipos de respostas, conforme a proposta de Valle (2014): *plenas*, por exemplo, “Sim”, “Não”, “Claro”, e *com repetição do verbo fonte* – “Entendi!”, “Compreendo!”, “Sei” etc. À guisa de exemplo, trazemos os excertos (35) e (36).

³⁹ Consideramos estímulos formas como *ahn, uhn, uhn uhn*, as quais são chamadas de *feedbacks* por Urbano (2019).

(35)

MARIAZINHA - Ah... Eu pensava que eles passavam o dia inteiro pensando em mim, e ao mesmo tempo não sabiam que eu existia, **entende?**

HOMEM — Entendo (*Fala baixo, senão eu grito* – Leilah Assumpção, 1969).

(36)

VITOR – Eu não digo trepar só. Eu falo viver, no sentido genérico: vagabundear, com o que a gente gosta, **me entende?**

VARREDOR – Claro (*O assalto* – José Vicente de Paula, 1969).

Nossa hipótese é que, na maioria das ocorrências, não haja presença de respostas após o uso dos RAD, dando indícios do esvaziamento da ilocução interrogativa desses elementos e reforçando o seu papel como MD.

3.4.4 Perfil social dos personagens

Conforme mencionado previamente, a categorização dos perfis sociais dos personagens baseia-se na proposta do projeto *Língua EmCena*. Cabe ressaltar que as peças trazem representações de costumes, de hábitos e de usos linguísticos de determinada época; logo, não se trata de uma situação real, mas de um evento criado pelos autores cujo modelo é a vida real (PRETI, 2004).

Uma das características mais marcantes do texto teatral é a presença de rubricas que orientam os leitores/atores. Em algumas peças, as características dos personagens são bem descritas, contendo, por exemplo, descrições quanto à ocupação, à idade, à etnia e outros aspectos relevantes. O excerto abaixo, extraído da peça *O assalto*, exemplifica essa situação:

VITOR: Nº 5.925.800 de uma **organização bancária**, neurótico, estranho, fuma muito, pinga colírio no olho nervosamente, como se duma hora para outra fosse ficar cego. Tem **25 anos**, é **branco**, sem vitalidade, frágil, está à beira da loucura, da loucura que leva ao hospício. (PAULA, 2010[1969], p. 127, grifos nossos)

Ademais, há peças em que tais descrições estão contidas na própria trama, sendo reveladas por meio dos diálogos entre os personagens. No entanto, há, também, peças que não trazem informações referentes a essas categorias e, tampouco, abrem espaço para realizarmos inferências; nesses casos, não nos resta escolha a não ser utilizar o rótulo *não se aplica* (NA).

De maneira geral, todas as variáveis e critérios considerados pelo projeto foram estabelecidos mediante o estudo da época e de como a sociedade era estratificada, a fim de

não provocar equívocos e anacronismos (BERGS, 2012). Assim sendo, as categorias foram pensadas de acordo com as sociedades de cada século e, quando necessário, os critérios adotados foram revisados e adaptados.

Diante disso, mesmo que estudos anteriores (SILVA; MACEDO, 1996; VALLE, 2001, 2014) tenham demonstrado que as macrocategorias sociais exercem pouca influência na produção dos RAD, nossa hipótese é que haja relação entre os RAD e as variáveis gênero/sexo, fase da vida, instrução, classe social e etnia, primeiramente, por estarmos lidando com peças teatrais, e, em segundo lugar, porque essas categorias são usadas na construção dos personagens, constituindo *personas*. É possível, então, que personagens com certos perfis sociais usem RAD específicos.

3.4.4.1 Gênero/sexo

As discussões a respeito do tratamento do gênero/sexo em estudos (socio)linguísticos não são recentes e não se esgotam. Desse modo, pensar nessa categoria é uma tarefa difícil e um tanto complexa, pois vai além de aspectos biológicos, envolvendo questões sociais (FREITAG; SEVERO, 2015).

Neste trabalho, essa categoria segue a dicotomia *feminino* e *masculino*, em conformidade com a categorização dos personagens proposta no projeto *Língua EmCena*. Todavia, em concordância com Brandão (2022), essa situação chama a atenção para questões de gênero e de como elas eram tratadas nos períodos em que as peças foram escritas. Segundo a autora, nesses períodos, a discussão a respeito do assunto ainda não havia emergido e, por serem, em sua maioria, escritas por autores masculinos, as peças reproduzem, muitas vezes, um ponto de vista patriarcal acerca dos papéis de cada gênero na sociedade.

Em vista dessas considerações, o modo como os personagens são retratados nas peças pode promover reflexões profícuas no que tange ao papel dos gêneros na sociedade do século passado.

3.4.4.2 Fase da vida

Os integrantes do *Língua EmCena* optaram pelo rótulo *fase da vida* por se tratar de um conceito mais geral, que abrange o tempo cronológico de vida dos personagens e os papéis desempenhados por eles em cada uma das etapas da vida. Além da realidade social dos personagens, o grupo levou em consideração a melhoria da qualidade de vida e o aumento da expectativa de vida das pessoas ao longo do tempo. Outro aspecto relevante para a seleção

dos critérios é definido pelo papel que as sociedades desses períodos atribuem a cada uma dessas fases.

Anteriormente, mencionamos que os membros do projeto *Língua EmCena* adotaram a periodização proposta na coleção *História do Brasil Nação*, considerando os seguintes períodos: i) 1831-1889; ii) 1890-1930; iii) 1931-1964; e iv) 1965-2010. Em decorrência disso, para se adequar à realidade social desses períodos, os critérios da variável fase da vida sofrem uma pequena modificação de um momento para outro: a partir de 1964, a fase 2 é subdividida em fases 2a e 2b. O recorte temporal adotado neste estudo, de 1950 a 1970, abrange o período no qual ocorre a ramificação da fase 2 (de 1964 em diante); por isso, em concordância com as decisões do grupo, decidimos manter a ramificação. Portanto, para que não haja equívocos, não estamos propondo que a variável fase da vida tenha sofrido mudanças significativas entre as décadas de 1950 e 1970, estamos apenas informando que, a partir de 1964, os personagens foram categorizados de outro modo (cf. Quadro 9).

Quadro 9 - Critérios da variável fase da vida

Períodos	Fases	Idades
1950 - 1963	Fase 1	de 0 a 15 anos
	Fase 2	de 16 a 50 anos
	Fase 3	de 51 em diante
	NA	não se aplica
1964 - 1970	Fase 1	de 0 a 17 anos
	Fase 2a	de 18 a 30 anos
	Fase 2b	de 31 a 59 anos
	Fase 3	de 60 em diante
	NA	não se aplica

Fonte: própria

De maneira bastante resumida, a **fase 1** agrega crianças, adolescentes e jovens. A **fase 2** é composta por jovens e adultos; entretanto, os papéis desempenhados pela juventude começam a se distinguir dos papéis assumidos por personagens adultos, daí a ramificação em **fase 2a**, que reúne jovens solteiros ou recém-casados, estudando ou com filhos recém-nascidos, e em **fase 2b**, que engloba personagens mais experientes, socioeconomicamente independentes, com ou sem filhos. Por fim, na **fase 3**, incluem-se os personagens idosos/velhos; a maior diferença de um período para outro é o aumento da expectativa de vida, de modo que personagens que seriam considerados idosos/velhos aos 50 anos passam a ser classificados como adultos.

3.4.4.3 Instrução

Os rótulos que denominam os níveis da educação contemporânea (educação infantil, ensino fundamental I e II, ensino médio e ensino superior) não poderiam ser utilizados para categorizar os personagens de séculos passados, por um lado, porque a educação é influenciada pelos meios político, social e econômico de um determinado período e, por outro, porque foram diversas as mudanças implementadas no sistema educacional até alcançarmos os moldes estabelecidos atualmente.

Por esse motivo, o grupo opta pelo rótulo *instrução*, definindo os seguintes critérios: i) ***sem instrução***: refere-se a personagens analfabetos e sem qualquer indício de instrução formal; ii) ***instrução básica***: diz respeito a personagens cuja única informação está ligada ao fato de saberem ler ou escrever – tais casos são identificados, por exemplo, quando os personagens escrevem cartas, leem livros, revistas, jornais etc.; iii) ***instrução média***: engloba personagens cujos ofícios demandam determinada especialização ou conhecimentos específicos; iv) ***instrução avançada***: quando os personagens cursaram ou estão cursando o ensino superior; e v) ***não se aplica (NA)***: quando não há informações suficientes para categorizar o personagem.

3.4.4.4 Classe social

Para além do poder aquisitivo dos personagens, essa categoria foi pensada a partir de diversos fatores relacionados a esses sujeitos, considerando aspectos como o estilo de vida e a ocupação desses personagens, isto é, as atividades desempenhadas por eles, seja um ofício, seja cuidar da casa ou, até mesmo, o fato de estar desempregado.

Baseando-se em estudos prévios (RIBEIRO, 1995; entre outros), os integrantes do projeto *Língua EmCena* propuseram os seguintes critérios para categorizar os personagens: i) ***classe 1***: alta sociedade / alta burguesia / classe dominante; ii) ***classe 2***: classes médias / pequena burguesia / classes emergentes; iii) ***classe 3***: classe subalterna / proletariado / camponeses; iv) ***classe 4***: oprimidos / marginais; e v) ***não se aplica (NA)***: quando não há informações suficientes para categorizar o personagem. Dentro dessas categorias, então, distribuem-se personagens diversos, de sujeitos da alta sociedade a maltrapilhos.

3.4.4.5 Etnia

O conceito de etnia é um tanto complexo e, em se tratando do contexto histórico brasileiro, não podemos desconsiderar a intensa variedade de raças e de culturas que se

mesclaram, formando o mosaico étnico-cultural brasileiro. De maneira geral, a etnia corresponde ao conjunto de características que definem um povo; contudo, devido à densa mistura racial, esse conceito não traduziria bem a realidade brasileira. Conforme pontuado por Osório (2003, p. 19),

Não existe uma classificação internacional para raças ou para etnias. Nos diferentes países, conceitos como etnia, tribo, nação, povo e raça recebem conteúdos locais, pois as bases importantes para a delimitação das fronteiras entre grupos sociais são produzidas pela história de cada sociedade.

Segundo Schwarcz (2012), o início do século XX foi marcado por profundas mudanças na sociedade brasileira e pela intensificação do amálgama de culturas e de costumes. De acordo com a autora, a população encontrava-se dividida entre brasileiros (nacionais), africanos e imigrantes – em sua maioria europeus. Essa mistura, somada ao crescimento e à urbanização acelerados, provocou intensos choques culturais, gerando descontrole e desestabilidade.

Não nos cabe, nesta pesquisa, refazer todo o percurso histórico acerca das raças que se estabeleceram no Brasil ao longo dos séculos, mas essas breves considerações se fazem necessárias para que possamos compreender as dificuldades encaradas pelos integrantes do projeto *Língua EmCena* na tentativa de definir os critérios caracterizadores da etnia dos personagens.

Além de brancos e pretos, como resultado da miscigenação, começam a surgir os caboclos, os mulatos, os cafuzos – nomenclaturas datadas do período colonial, algumas das quais consideradas racistas –, que podem ser incluídos dentro de um rótulo maior: os pardos. Os recenseamentos realizados ao longo do século XX (GOMES, 2013) classificavam a população por meio das seguintes cores: brancos, pretos, amarelos e pardos; além disso, contabilizavam-se os indivíduos sem declaração de cor. Em decorrência dessas vastas categorias, o grupo optou pelos rótulos **branco** e **não branco**, os quais abrangem as diversas raças encontradas no país e, em especial, por essa divisão corresponder a uma distinção que marca profundamente nossa sociedade, como uma herança ainda presente do período escravocrata.

No caso das peças, a representatividade de personagens não brancos é muito pequena e, quando esses personagens aparecem, tendem a ocupar as camadas mais baixas da sociedade, atuando como personagens secundários. Ainda assim, essa variável será considerada na análise, para que possamos observar se os RAD são, em maior ou menor medida, influenciados por essa categoria.

3.4.5 Relação entre os personagens/interlocutores

Em consonância com Bell (1984), o falante modela seu discurso de acordo com a sua audiência, isto é, o(s) seu(s) ouvinte(s). Dessa maneira, considerar as relações entre os personagens e seus interlocutores pode nos revelar se esse fator motiva a escolha dos RAD. Brown e Gilman (1960), em um estudo acerca dos pronomes, sugerem que há correlação entre o fenômeno linguístico e as dimensões de poder e de solidariedade. Explicitando a proposta dos autores, Coelho e Nunes de Souza (2014, p. 171) declaram que,

[...] o poder se manifesta de diferentes formas: força física, riqueza, idade, sexo, papel institucionalizado na igreja, no Estado, no exército ou na família etc. Uma relação regida pela dimensão de solidariedade, por sua vez, é uma relação simétrica. Frequentar a mesma escola, ter os mesmos pais e ter a mesma profissão são relações de solidariedade.

Diante disso, os usos linguísticos dos interlocutores estariam intimamente relacionados às suas escolhas estilísticas, guiadas pelo tipo de relação que mantêm entre si. Consideramos, então, as diversas relações entre os personagens e seus interlocutores, tendo em vista os papéis desempenhados por esses sujeitos nas mais variadas instâncias sociais (família, trabalho, amizades etc.).

Para tal fim, adotamos os seguintes critérios: relação *simétrica* ou *assimétrica* e relação *íntima* ou *não íntima*. Consideramos simétricas as relações entre sujeitos de mesmo *status* e nível social, os quais partilham certas características quanto à idade, à escolaridade, à profissão, à classe social etc. As assimétricas, em contrapartida, correspondem às relações entre personagens pertencentes a grupos sociais distintos, diferindo quanto ao *status* e ao nível social – nesse caso, os personagens pertencem a classes sociais distintas, um exerce autoridade política ou profissional sobre o outro etc. São consideradas íntimas as relações entre personagens com algum grau de parentesco (irmãos, primos, casais, pais, filhos etc.) ou entre personagens que partilham confidências e possuem uma amizade de longa data. Por fim, as relações não íntimas se referem a personagens desconhecidos, sem vínculo familiar ou de amizade (BROWN; GILMAN, 1960; COELHO; NUNES DE SOUZA, 2014).

Nossa hipótese é que os RAD sejam mais recorrentes em situações de intimidade e de simetria, sendo, portanto, menos empregados em contextos de menor grau de intimidade e de assimetria.

3.4.6 Ambientação

Está claro que as línguas variam de uma região para outra. Essa variação, embora seja mais perceptível no léxico, atinge todos os níveis da língua (COELHO *et al.*, 2019). No caso dos RAD, não seria diferente: assim como os demais fenômenos linguísticos, esses elementos podem apresentar variabilidade de uma região para outra. O RAD *não tem?*, por exemplo, é bastante típico da fala florianopolitana, sendo, praticamente, inexistente na fala de indivíduos pertencentes a outras regiões do Brasil (VALLE, 2001, 2014).

O controle dessa variável toma como ponto de partida o fato de que as peças são ambientadas em determinadas regiões (Rio de Janeiro, Taperoá - PB, São Paulo etc.), e seus autores, buscando retratar os personagens de forma realista, baseiam-se nos usos linguísticos de falantes reais, que, originalmente, pertencem a essas localidades. Por serem diversas as localidades em que as peças teatrais são ambientadas, optamos pelos amálgamas ***nordeste*** e ***sudeste***, com o propósito de alcançar resultados mais consistentes.

Nossa hipótese é que algumas formas sejam mais frequentes em uma dessas regiões, ao passo que outras estejam mais bem distribuídas, sendo usadas, em proporções semelhantes, nas diferentes localidades. Esperamos que os RAD oriundos de *entender* e de *saber* sejam os menos segmentados em relação à região devido ao padrão de recorrência dessas formas em estudos anteriores, o que pode ser um indício de que sejam expressões comuns na fala. Quanto às demais formas (*compreender*, *perceber* e *ver*), esperamos que elas sejam utilizadas em locais específicos.

3.4.7 Autoria

Ainda que os textos dramáticos sejam materiais relevantes para a pesquisa sociolinguística e sócio-histórica, não podemos nos esquecer que as peças de teatro não são equivalentes a uma comunidade de fala, pois dizem respeito a uma realidade criada, isto é, uma representação gerada por um usuário real da língua – o autor. Dessa forma, mesmo que o modelo para a elaboração dessas representações seja a vida real, trata-se, sobretudo, da percepção do autor a respeito da língua, dos costumes e dos hábitos da sociedade de determinado período.

Por essa razão, iremos considerar a autoria das peças como uma variável independente em nosso trabalho, posto que a percepção desses dramaturgos poderá refletir nos usos linguísticos dos personagens das peças. Portanto, nossa hipótese é que certos RAD sejam mais utilizados nas peças de determinados autores, revelando características linguísticas não apenas

dos personagens representados, mas dos próprios escritores, que são falantes reais, inseridos em comunidades de fala e em grupos sociais específicos.

3.5 Sintetizando...

Neste capítulo, realizamos a descrição das etapas desenvolvidas ao longo desta pesquisa. O percurso foi iniciado pelo levantamento bibliográfico de obras que constituem repertórios pertinentes para este estudo. Em seguida, selecionamos peças teatrais brasileiras para compor o nosso *corpus*, as quais foram lidas e categorizadas, a partir da proposta do projeto *Língua EmCena*. Ademais, descrevemos como se deu a coleta e o tratamento dos dados e apresentamos as variáveis a serem controladas, bem como as hipóteses a respeito de cada uma delas.

O quadro 10 sintetiza todas as variáveis independentes e seus respectivos critérios.

Quadro 10 – Sistematização das variáveis independentes

Variáveis		Critérios
Linguísticas	Apresentação formal	Composta Simple
	Posição	Medial Final
	Presença/ausência de respostas	Ausência Presença: • resposta plena • resposta com repetição do verbo fonte
Extralinguísticas	Gênero/sexo	Feminino Masculino
	Fase da vida	Fase 1 / Fase 2 / Fase 3 Fase 1 / Fase 2a / Fase 2b / Fase 3
	Instrução	Sem instrução Instrução básica Instrução média Instrução avançada
	Classe social	Classe 1 Classe 2 Classe 3 Classe 4
	Etnia	Branco Não branco
	Relação entre os personagens/interlocutores	Íntima / Não íntima Assimétrica / Simétrica
	Ambientação	Nordeste Sudeste
	Autoria	Abílio Pereira de Almeida Antônio Bivar Ariano Suassuna Augusto Boal Consuelo de Castro Dias Gomes Francisco Pereira da Silva Hilda Hilst José Vicente de Paula Leilah Assumpção Mário Prata Nelson Rodrigues Oduvaldo Vianna Filho Plínio Marcos

Fonte: própria

4 ANÁLISE E DISCUSSÃO DOS DADOS: UM OLHAR PARA OS RAD NAS PEÇAS TEATRAIS

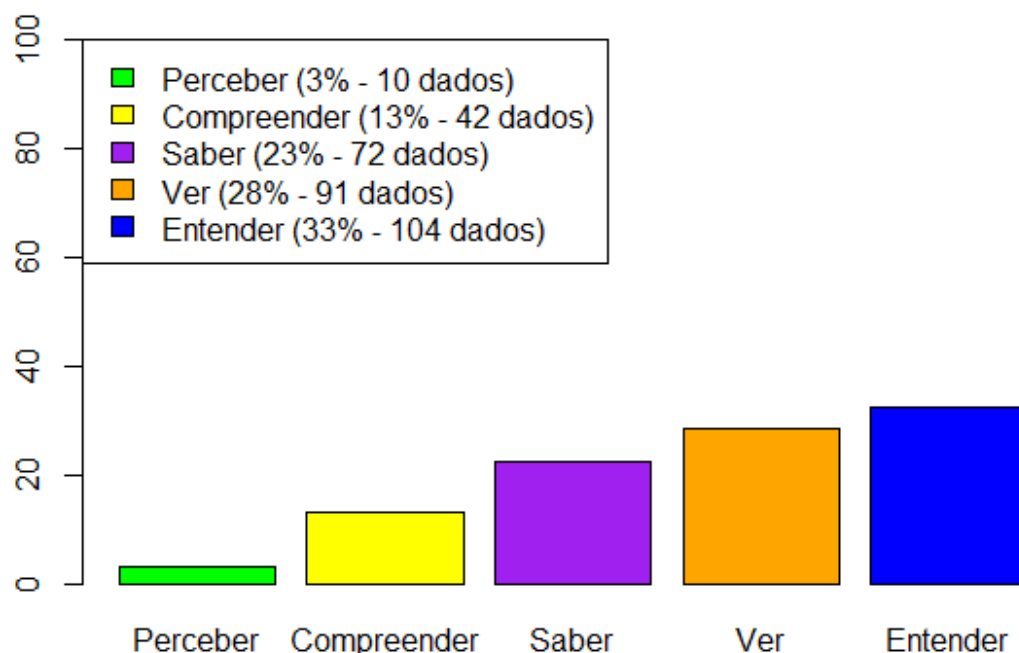
Esta seção destina-se à análise do fenômeno e à avaliação do comportamento das variáveis. Primeiramente, traçamos uma visão geral dos resultados, para, em seguida, esmiuçarmos cada uma das variáveis controladas e sua correlação com o fenômeno em estudo. Dessa forma, apresentamos os resultados das variáveis de natureza linguística – apresentação formal, posição e presença/ausência de respostas – e, logo depois, passamos a discutir os resultados das variáveis de natureza extralinguística – gênero/sexo, fase da vida, instrução, classe social e etnia, as quais correspondem ao perfil social dos personagens, relação entre os personagens/interlocutores, ambientação e autoria.

As discussões desenvolvidas são respaldadas nos pressupostos teóricos, tendo em vista, também, os resultados de estudos prévios, de maneira que possamos comparar o comportamento dos RAD com o conhecimento já construído sobre o tema. A partir disso, em concordância com a hipótese levantada previamente, esperamos demonstrar que é possível dispensar um tratamento variável aos RAD e que o uso dessas expressões é motivado por fatores linguísticos e extralinguísticos.

4.1 Visão geral dos resultados

Antes de exibir os resultados, cabe destacar que, dos 401 personagens das peças, apenas 123 utilizam RAD em suas falas, o que significa que somente estes personagens são considerados em nossa análise e, portanto, as ocorrências dos RAD, bem como a categorização dos perfis sociais, dizem respeito a esses 123 personagens.

Após a coleta dos dados, realizamos rodadas estatísticas por meio do programa R (CORE TEAM, 2022) para verificar a frequência de uso dos RAD nas peças teatrais. Foram constatadas **319 ocorrências**, o que, sob uma perspectiva quantitativa, constitui uma amostra pequena, mas, sob um ponto de vista sócio-histórico, corresponde a uma quantidade significativa de elementos típicos da oralidade em textos escritos. Diante disso, o gráfico 1 representa a frequência dos RAD nos textos selecionados.

Figura 1 – Distribuição dos RAD no *corpus*

Fonte: própria

O gráfico revela que os RAD derivados dos verbos *entender* (104 - 33%), *ver* (91 - 28%) e *saber* (72 - 23%) são os mais recorrentes nas peças analisadas. No que concerne à frequência, os RAD originados de *perceber* (10 - 3%) e de *compreender* (42 - 13%) são os menos utilizados nas peças. Esse resultado condiz, em parte, com a constatação de Valle (2014), mais especificamente quando a autora discute os critérios para a delimitação de variáveis discursivas; em sua amostra⁴⁰, os RAD provenientes de *compreender* e de *perceber* foram raros ou inexistentes. O que diferencia os nossos dados dos da autora é o fato de que o RAD *viu?* encontra-se entre os mais utilizados nas peças, ao passo que, na amostra investigada pela pesquisadora, essa forma quase não aparece. Silva e Macedo (1996), adotando uma perspectiva variacionista para o estudo dos MD, apontam os RAD *sabe?* e *entendeu?* como alguns dos mais frequentes em sua amostra⁴¹. Urbano (2019), inserindo as expressões que atuam no domínio da requisição de apoio discursivo no grupo dos *fáticos de natureza ou entonação interrogativa*, constata que os RAD *sabe?*, *entendeu?* e *viu?* estão entre os mais numerosos, ficando apenas abaixo de *né?*.

⁴⁰ Para a realização de sua tese, Valle (2014) utilizou a amostra Brescancini-Valle (2001-2010), que integra o Banco de Dados VARSUL (Variação Linguística Urbana do Sul do País). A autora examinou 30 entrevistas com informantes da comunidade da Barra da Lagoa – Florianópolis/SC.

⁴¹ Silva e Macedo (1996) utilizam como *corpus* as 64 entrevistas da Amostra Censo.

Em relação à baixa frequência dos RAD derivados de *compreender* e de *perceber*, é possível que a percepção dos autores, atrelada a fatores regionais, esteja motivando a escolha de certos RAD em detrimento de outros. Sendo assim, podemos supor que esses RAD sejam pouco utilizados nas regiões onde as peças são ambientadas e, na busca por uma representação que se aproxime da realidade linguística dos falantes, os autores acabaram não adotando essas formas.

Cabe, ainda, destacar um aspecto relevante no tocante ao uso dos MD derivados de *saber*. Valle (2001, p. 160) já havia verificado as seguintes ocorrências:

[...] **Sabe**, eu vejo jovens nas/ nas praças aí, eu também tive minha época de/ de estar num/ num/ nesses/ nesses/ nesses negócios de rolo. E é/ eu acho triste, eu acho que não/ não devia ser assim, preferia que fosse como era antes. (FLP12MJC:166)

[...] **Sabe**, a droga, ela/ como falei ontem também, pra minha namorada, ela tipo, vai te cegando, entende? (FLP12MJC:354)

Nesses trechos, em conformidade com Valle (2001), *sabe* sinaliza traços de operadores argumentativos, realiza um movimento catafórico e é usado em situações nas quais os falantes apresentam um posicionamento acerca de determinado assunto. Nas palavras da autora, “este uso nos parece bastante comum atualmente e pode estar indicando um passo adiante de *sabe*? no estágio de gramaticalização” (VALLE, 2001, p. 160, grifo da autora). Similarmente, nas peças analisadas, notamos esses usos e listamos alguns deles abaixo:

i) TONHO - **Sabe**, Paco, às vezes eu até penso que você é um bom chapa. (*Dois perdidos numa noite suja* – Plínio Marcos, 1966).

ii) GORILA - Fui eu que trouxe o doutor Deodato aqui e depois ajudei a arrumar as coisas... **Sabe**, doutor Deodato precisa de mim. Sou cabo eleitoral dele. Uma palavra minha é ordem. E a mocinha vai tão bem... era uma pena... (*A invasão* – Dias Gomes, 1960).

iii) HELONEIDA - **Sabe**, Geni, às vezes eu penso que estou num purgatório, mas com os últimos acontecimentos eu tenho quase a certeza de estar no inferno mesmo. (*Abre a janela e deixa entrar o ar puro e o sol da manhã* – Antônio Bivar, 1968).

iv) VITOR - **Sabe**, não tolero cheiro de cigarro fumado. Me intoxica. (*O assalto* – José Vicente de Paula, 1969).

v) ROSA - Gostava, sim. **Sabe**, na roça, o homem é feio, magro, sujo e mal vestido. Ele até que era dos melhores. Tinha um sítio. (*O pagador de promessas* – Dias Gomes, 1960).

Nos excertos apresentados, *sabe* parece chamar a atenção do interlocutor para o enunciado que será introduzido, atuando, conforme apontado por Valle (2001), como um termo catafórico. Nos exemplos (i) e (iii), os vocativos *Paco* e *Geni*, usados logo após *sabe*, reforçam a ação de chamar a atenção do ouvinte para o que será dito. Nesse sentido, concordamos com a conclusão de Valle (2001), segundo a qual tais aspectos podem ser indícios do avanço desse MD no processo de gramaticalização. Ademais, por surgirem em peças produzidas no século XX, é possível afirmar que esses usos não são recentes, fator que também corrobora o estágio mais avançado de *sabe* no processo de gramaticalização.

Diante dessas considerações, nas próximas seções, apresentamos os resultados da correlação dos RAD com as variáveis de natureza linguística e extralinguística, verificando se, de fato, esses fatores motivam a escolha e o uso dos marcadores sob análise.

4.2 Variáveis de natureza linguística

Antes de discutir os resultados obtidos a partir da correlação dos RAD com cada uma das variáveis de natureza linguística, apresentamos um quadro com os valores de significância dessas variáveis sobre o fenômeno. Os valores exibidos foram obtidos mediante testes de qui-quadrado, avaliando as variáveis individualmente.

Quadro 11 – Valores dos testes de qui-quadrado: variáveis linguísticas⁴²

Variáveis linguísticas	Qui-quadrado
Apresentação formal	X-squared = 51.008, df = 4, p-value = $2.224e^{-10}$
Presença/ausência de respostas	X-squared = 25.344, df = 4, p-value = $4.29e^{-05}$
Posição	X-squared = 10.184, df = 4, p-value = $3.744e^{-02}$

Fonte: própria

O teste de qui-quadrado tem por objetivo avaliar se há distinções entre as proporções dos dados, calculando a diferença entre os valores observados e esperados. Nesse sentido, “X-squared” equivale à identificação do tipo de teste realizado (qui-quadrado), “df” diz respeito ao grau de liberdade e “p-value” corresponde ao valor de significância, que é significativo quando está abaixo de 0,05 (5%) (BRANDÃO, 2018). Os valores demonstram que todas as variáveis linguísticas são estatisticamente significativas. A seguir, então, realizamos análises mais pormenorizadas a respeito de cada variável e, posteriormente, fazemos cruzamentos entre as variáveis, a fim de avaliar possíveis relações.

⁴² O quadro foi organizado de acordo com os valores de significância das variáveis (*p-value*) – da mais significativa para a menos significativa.

4.2.1 Apresentação formal

Ao controlarem a apresentação formal dos MD, Risso, Silva e Urbano (2019) constataram que esses elementos apresentam, em proporções semelhantes, tanto formas únicas e fixas (49,2%) quanto formas variantes (50,8%); todavia, segundo eles, esse resultado não altera o fato de que essas formas tendem à cristalização e à automatização no discurso, tornando-se formulaicas. Por essa razão, em nossa análise, consideramos a distinção entre *formas simples* e *formas compostas*. Vale enfatizar, contudo, que, conforme discutimos na seção dedicada à gramaticalização dos RAD (cf. 2.6.2), entendemos que nem todas as formas coletadas atingiram o estatuto pleno de marcador, ou seja, elas ainda não possuem todas as características de um MD prototípico; ainda assim, decidimos selecioná-las porque elas demonstram estar em um processo incipiente de gramaticalização, ilustrando a trajetória de mudança dos verbos rumo à função de RAD. À luz da GTI, essas formas constituiriam casos menos prototípicos de MD, partilhando algumas propriedades com as unidades que já alcançaram esse estatuto.

Assim sendo, visando esclarecer quais são as *formas compostas* coletadas em nosso *corpus*, agrupamos as variantes dos RAD de acordo com os seus modificadores, apresentando, na sequência, alguns exemplos para fins de contextualização.

Quadro 12 – Formas compostas dos RAD

Modificadores	Variantes
Advérbios	Compreendeu agora ? Entende agora ? Entendeu agora ? Entendeu bem ? Num sabe?
Forma estendida	Sabe como é ?
MD	Entende, né ? Mas compreendeu? Sabe, né ?
Pronomes	Entende eu ? Está me entendendo? Me entende? Me entendeu? O senhor compreende? O senhor entende? O senhor sabe? Os senhores entendem? Tá me entendendo? Tu compreendes? Você compreende? Você compreendeu? Você entende? Você me compreende? Você me entende? Você me entendeu? Vocês me entendem?
Verbos	Está me entendendo? Tá entendendo? Tá me entendendo? Tá percebendo? Tá sabendo?

Fonte: própria

- **Advérbios:** agora, bem, não

(37)

FREDERICO - Você vai voltar à sua vidinha de antes. A gente vai casar. Você vai tornar a andar no seu carro.

ROSA - Que é que você acha que mudou em mim?

FREDERICO - Faz um mês que você está plantada aqui; contra o teu pai, contra a polícia, contra um milhão de coisas. **Entendeu agora?**

ROSA - Você está distorcendo tudo. Não é nada disto. (*À prova de fogo* – Consuelo de Castro, 1968).

(38)

TONHO - Eu que mando, *entendeu?* Você só faz o que eu mandar! **Entendeu bem?** Eu que mando. (*Dois perdidos numa noite suja* – Plínio Marcos, 1966).

(39)

CAFUNÉ - Sem jogá é o melhó do campo, **num sabe?** O que foi aí? (*Chapetuba Futebol Clube* – Oduvaldo Vianna Filha, 1959).

No exemplo (37), ao olharmos para o contexto da interação, percebemos que a expressão *entendeu agora?* remete a algo que foi dito anteriormente, isto é, o personagem busca esclarecer o que foi falado há pouco, daí o uso do advérbio *agora*. Nesse caso, o significado de *entender* está bastante apegado ao verbo fonte; entretanto, a posição, o contexto de uso, a conjugação e, principalmente, a ausência de manifestação verbal em resposta à pergunta são particularidades dos RAD. No exemplo (38), o advérbio *bem* é utilizado para intensificar o MD *entendeu?*, que, por sinal, já havia sido utilizado na fala do personagem. Assim como no exemplo anterior, o RAD surge em sua forma mais comum – na terceira pessoa do singular do pretérito perfeito do indicativo (ZANIN; GONÇALVES, 2022). No último exemplo, temos uma estrutura que equivale a uma *tag question*. Urbano (2019), ao discutir os RAD *não é verdade?*, *não é?/num é?* e *né?*, nota que esses marcadores equivalem às interrogativas *tag*, como nos seguintes exemplos: “L2 – bom porque fome é mundial *não é?*” e “Inf. – o pessoal deixou de ir a teatro *né?*” (URBANO, 2019, p. 473, grifos do autor). A expressão *num sabe* poderia ser substituída por *não/num é*, *né* ou pelos RAD em estudo sem perda de sentido proposicional, revelando, mais uma vez, um contexto propício para a ocorrência dos RAD.

- **Forma estendida:** sabe como é?

(40)

LEÔNIDAS - Você sabe que o signo de Aquário é o mais evoluído? As pessoas do seu signo são cem anos adiantadas sobre as pessoas dos outros signos e, quem sabe, até mesmo sobre as do próprio signo, **sabe como é?** São geralmente altas... (*Cordélia Brasil* – Antônio Bivar, 1967).

Esse caso corresponde ao que Valle (2014) chama de extensão da forma, em que o RAD é complementado, formando uma pergunta. O autor poderia simplesmente utilizar *sabe?*, que seria suficiente no contexto da interação, mas opta por estender a forma. De todo modo, *sabe?* foi usado em sua forma mais típica e frequente – terceira pessoa do singular do

presente do indicativo (URBANO, 2019) –, e o fato de a expressão *sabe como é?* ser utilizada no mesmo contexto que os RAD remete ao princípio da estratificação, segundo o qual, uma vez gramaticalizada, uma forma passa a coabitar com outras em um mesmo domínio funcional (HOPPER, 1991).

- **MD:** mas, né?

(41)

SELMINHA – [...] Outro dia. A fechadura do banheiro estava quebrada. Arandir empurra a porta e vê Dália nua. Sem querer, naturalmente, e nem ele podia imaginar que. **Mas compreendeu?** Pelada. Completamente! Tinha acabado de tomar banho. [...] (*O beijo no asfalto* – Nelson Rodrigues, 1960).

(42)

DIDI - Olha, se for muito pode dizer. Eu andei conversando com o meu pai e a gente ficou pensando... É evidente que... **Sabe, né?** A gente, é evidente... (*O cordão umbilical* – Mário Prata, 1970).

Em ambos os exemplos, temos casos de agrupamento de MD ou de ocorrências conjuntas (RISSO, 2019). No exemplo (41), percebemos que há uma interrupção abrupta, destacada pela presença do ponto final após *que* (“Sem querer, naturalmente, e nem ele podia imaginar que.”), e uma retomada, produzida pelo MD *mas*, dando continuidade ao diálogo.

No exemplo (42), temos o agrupamento de dois RAD: *sabe?* e *né?*. Nesse trecho, o personagem discute o preço do aluguel de um apartamento com o amigo, com quem iria dividir o valor do imóvel; nota-se que, temendo a reação do colega, o personagem hesita em revelar o preço – atitude que é reforçada pela presença das reticências e dos dois RAD. Nesse caso, é provável que ambos os RAD estejam sendo usados como atenuadores, isto é, “um recurso estratégico que busca a aceitação por parte do ouvinte” (BENTES; FERREIRA-SILVA; MARIANO, 2013, p. 296). Em um estudo da fala de Mano Brown, Bentes, Ferreira-Silva e Mariano (2013) constataram que os MD, principalmente aqueles que têm o foco voltado para o eixo interacional, são usados como recursos atenuadores cujo papel é obter a aquiescência e a aceitação do interlocutor. Retornando ao exemplo apresentado, o personagem recorre a recursos atenuadores (*sabe?* e *né?*) com a finalidade de obter a aceitação e a compreensão do colega.

Em vista disso, nos dois exemplos, é possível identificar particularidades dos RAD: *compreendeu?* está conjugado na terceira pessoa do singular do pretérito perfeito do indicativo, característica similar à do RAD *entendeu?*, que, segundo Urbano (2019), é uma forma bastante recorrente, e *sabe?* também surge em sua forma mais comum, flexionado na terceira pessoa do singular do presente do indicativo. Além disso, o fato de esses RAD estarem atuando em conjunto com outros MD acentua, ainda mais, traços como a exterioridade ao conteúdo proposicional e a independência sintática desses elementos.

- **Pronomes:** eu, me, senhor(es), tu, você

(43)

SELMINHA - Uma coisa que me dá vontade de morrer. Como é que um homem pode desejar outro homem. (veemente e voltando-se para a irmã) Dália, **você entende? Entende eu?** Sei que, agora, quando um homem olhar para o meu marido. Vou desconfiar de qualquer um, Dália! [...] (*O beijo no asfalto* – Nelson Rodrigues, 1960).

(44)

VARREDOR - É louco. Tantã. Não bate, **me entende?** Quebra as coisas em casa, quando dá a crise, bate nas meninas, bate na mãe, fica batendo a cabeça na parede, de noite, e gritando, não deixa ninguém dormir, apronta e desapronta à vontade, os cambaus! (*O assalto* - José Vicente de Paula, 1969).

(45)

JUIZ VELHO - Ele falou demais. **O senhor compreende?** E boca deve ter uma medida. (*O verdugo* – Hilda Hilst, 1969).

(46)

VERDUGO – [...] O homem entrou no meu peito, **os senhores entendem?** Ele falava que era preciso...amor...ele falava... (*O verdugo* – Hilda Hilst, 1969).

(47)

CORCUNDA - [...] Com o tempo, um certo limo
Se faz na nossa carne. [...]
Ele a si mesmo se transforma
E te faz criatura alegre ou triste.
Te faz acreditar no que perdura

Ou em tudo que te parece real
 Mas que não existe. (Pausa) **Tu compreendes?**
 (*O visitante* - Hilda Hilst, 1968).

(48)

GERVÁSIO – [...] Não há a encruzilhada para se escolher. **Você me entende?** O caminho já está traçado. É mais simples. (*Em moeda corrente do país* - Abílio Pereira de Almeida, 1960).

Em todos esses exemplos, há casos menos prototípicos de MD: as formas são totalmente transparentes semanticamente, já que o significado do verbo de origem é conservado, e estão flexionadas (*compreendes, entendem, entende* etc.), o que não é comum aos RAD, cuja tendência é fixar-se em uma forma (RISSO; SILVA; URBANO, 2019). Ainda assim, com exceção do exemplo (47), nos demais excertos, os RAD surgem em posição medial de turno e não há manifestação verbal por parte do interlocutor após o uso desses elementos, o que pode estar caracterizando o esvaziamento da ilocução interrogativa dessas formas. Somado a isso, em todos os casos, a ausência de autonomia comunicativa é perceptível, visto que essas formas dependem do contexto comunicativo para que façam sentido; podemos afirmar, ainda, que elas são alheias à estrutura oracional e, por conseguinte, sintaticamente independentes. Apesar de não serem MD prototípicos, essas formas não atuam como verbos plenos, mas ilustram a trajetória percorrida pelos verbos em direção ao estatuto de RAD, pois, conforme discutimos anteriormente, o processo de gramaticalização é gradual e possui diferentes estágios (GONÇALVES, 2021). Essas formas, então, estariam ainda no início do processo.

- **Verbos:** (es)tá

(49)

KÁTIA - Depois? Vou embora já-já. Tenho meu coronel, **tá sabendo?** Não chego depois das seis, não. [...] (*O cordão umbilical* – Mário Prata, 1970).

O único verbo que acompanha os RAD nas peças é *estar*, que costuma ser seguido por gerúndio (*tá percebendo?*, *(es)tá entendendo?*, *tá sabendo?*) e tende a ser usado em sua forma reduzida (*tá*).

Concluída essa síntese, na tabela 1, apresentamos as variantes de maior frequência em nosso *corpus*.

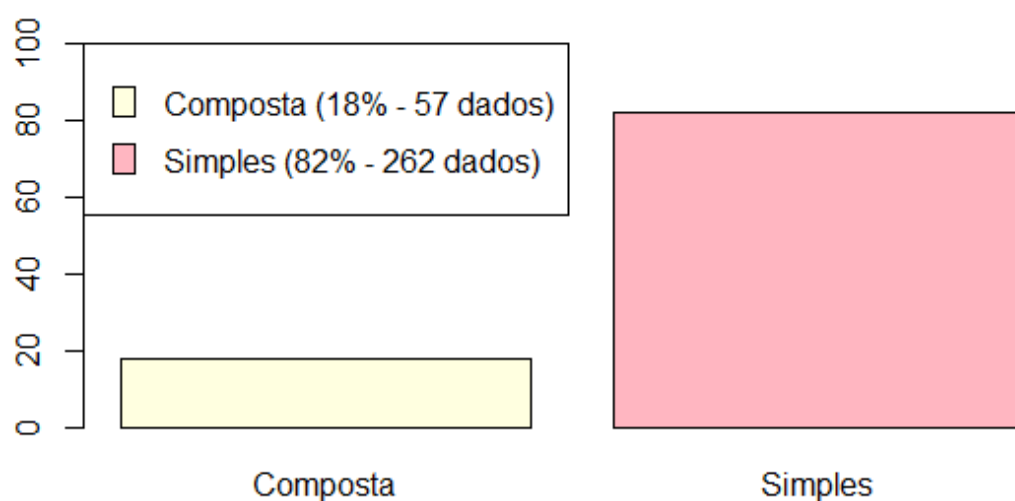
Tabela 1 – Os RAD mais frequentes no *corpus*

Variantes	Distribuição	
	Nº/T	%
-		
Compreendeu?	20/231	9%
Entende?	37/231	16%
Entendeu?	25/231	11%
Percebeu?	6/231	2,5%
Sabe?	52/231	22,5%
Viu?	91/231	39%
TOTAL	231/231	100%

Fonte: própria

Apesar da variabilidade encontrada no *corpus*, os RAD de maior frequência são formas que aparecem em outros estudos de MD e são, reconhecidamente, tomadas como RAD. Tais formas, então, estariam mais gramaticalizadas e teriam mais traços identificadores de MD, como alta frequência, massa fônica reduzida, invariabilidade formal ou variabilidade restrita (RISSO; SILVA; URBANO, 2019). Assim sendo, temos a seguinte distribuição: das 42 ocorrências de RAD derivados de *compreender*, 20 são de *compreendeu?*; das 104 ocorrências de *entender*, 37 são de *entende?* e 25 de *entendeu?*; das 10 ocorrências de *perceber*, 6 são de *percebeu?*; das 72 ocorrências de *saber*, 52 são de *sabe?*; e, das 91 ocorrências de *ver*, todas são de *viu?* – o único RAD que não apresenta variação no *corpus*.

Feitas essas considerações, na figura 2, apresentamos os resultados da variável apresentação formal, considerando as 319 ocorrências dos RAD em estudo.

Figura 2 – Apresentação formal dos RAD

Fonte: própria

Os resultados corroboram a nossa hipótese, segundo a qual as formas simples (262 – 82%) seriam mais frequentes do que as formas compostas (57 – 18%), sinalizando uma maior integração desses elementos ao papel de RAD. A tabela 2 apresenta essa distribuição de maneira mais especificada.

Tabela 2 – Distribuição dos RAD de acordo com a apresentação formal

RAD	Composta		Simples		TOTAL	
	Nº/T	%	Nº/T	%	Nº/T	%
-						
Compreender	19/42	45%	23/42	55%	42/319	13%
Entender	29/104	28%	75/104	72%	104/319	33%
Perceber	1/10	10%	9/10	90%	10/319	3%
Saber	8/72	11%	64/72	89%	72/319	23%
Ver	0/91	0%	91/91	100%	91/319	28%
TOTAL	57/319	18%	262/319	82%	319/319	100%

Fonte: própria

A tabela revela que os RAD derivados de *entender* (75 – 72%), *saber* (64 – 89%) e *ver* (91 – 100%) possuem, em sua maioria ou de maneira geral, como no caso de *viu?* (91 – 100%), formas simples, constituídas por uma única palavra e sem acompanhamento de outros termos. Em relação aos RAD provenientes de *perceber*, embora haja uma baixa quantidade de dados para que possamos realizar afirmações mais consistentes, nota-se a predominância de formas simples (9 – 90%) sobre formas compostas (1 – 10%), constituindo um resultado significativo. Por outro lado, os resultados dos RAD oriundos de *compreender* contrastam com os anteriores, pois há uma distribuição bastante equilibrada entre formas simples (23 – 55%) e formas compostas (19 – 45%), com leve predominância das simples. Esse resultado sugere que *compreender* ainda é uma fonte significativa de formas menos gramaticalizadas. A partir disso, podemos deduzir que as formas simples – não todas, mas a maioria, como as exibidas na tabela 1 – correspondem aos RAD mais prototípicos e, portanto, mais avançados no processo de gramaticalização, e as demais formas, deixando de desempenhar papel de verbo pleno, estariam em estágios iniciais ou intermediários desse processo.

4.2.2 Posição

Assumindo o turno conversacional como unidade de análise, consideramos as posições medial e final dos RAD, exemplificadas por meio dos excertos abaixo:

(50)

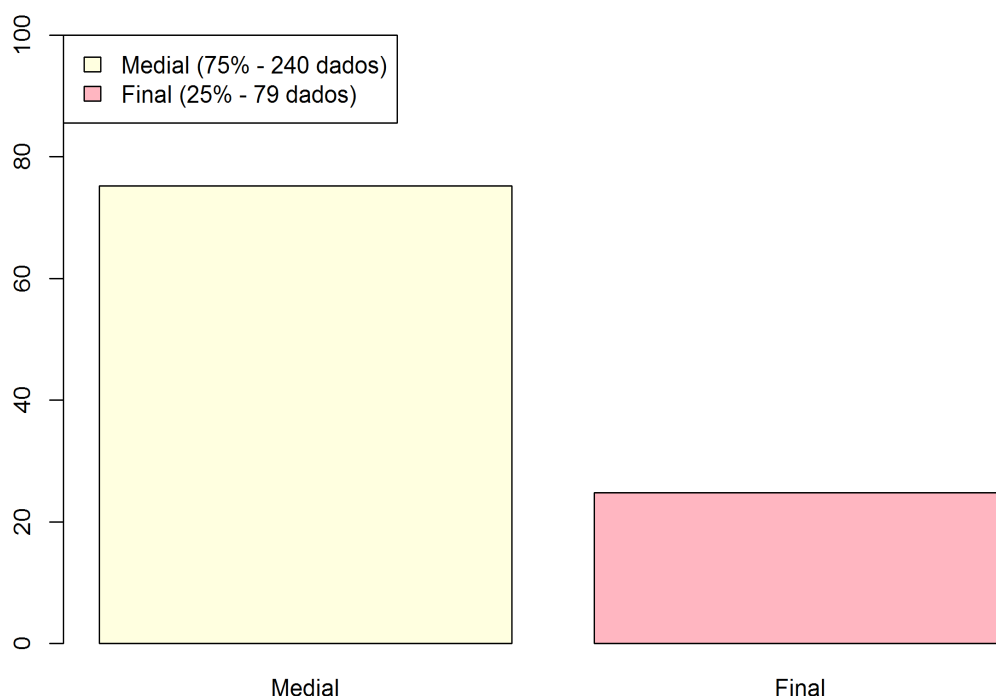
FREI ROQUE - Ah, Cancão miserável, falando da defuntência dos outros mais pobres do que ele! Pois agora eu vou, **sabe?** Mas vou da raiva em que estou, está ouvindo? Onde está o cavalo? Pelo menos essa desgraça presta? (*O casamento suspeito* - Ariano Suassuna, 1957).

(51)

BENEDITO - Não se meta não, **viu?** (*A pena e a lei* – Ariano Suassuna, 1959).

Diante disso, a figura 3 apresenta os resultados concernentes à posição dos RAD no *corpus* investigado.

Figura 3 – Posição dos RAD no turno conversacional



Fonte: própria

Por meio desse gráfico, percebemos que, em concordância com a nossa hipótese, a posição medial (240 – 75%) é priorizada. Há, também, conforme esperávamos, algumas ocorrências em final de turno (79 – 25%). Nessa direção, de acordo com a hipótese inicial, os RAD são mais frequentes em posição medial, favorecendo o desenvolvimento e a sustentação do turno – aspectos que destacam a função articuladora desses elementos, a qual, geralmente, não recebe destaque – e atuando, simultaneamente, no gerenciamento da interação, uma vez que os RAD não perdem seus traços interacionais. Também havíamos sugerido que as formas compostas ocorreriam mais em posição final e que as formas simples ocorreriam mais em posição medial. Vejamos a tabela 3:

Tabela 3 – Posição e apresentação formal dos RAD

Posição	Composta		Simples		TOTAL	
-	Nº/T	%	Nº/T	%	Nº/T	%
Medial	39/240	16%	201/240	84%	240/319	75%
Final	18/79	23%	61/79	77%	79/319	25%
TOTAL	57/319	18%	262/319	82%	319/319	100%

Fonte: própria

De um modo geral, vemos que tanto formas simples (201 - 84%) quanto formas compostas (39 - 16%) ocorrem, predominantemente, em posição medial de turno, mas há, de fato, uma leve tendência às formas compostas serem, proporcionalmente, mais frequentes na posição final (23%) do que na medial (16%). Posto isso, a tabela 4 informa os resultados de modo mais detalhado, tendo em vista a distribuição dos RAD nas duas posições consideradas.

Tabela 4 - Distribuição dos RAD de acordo com a posição no turno

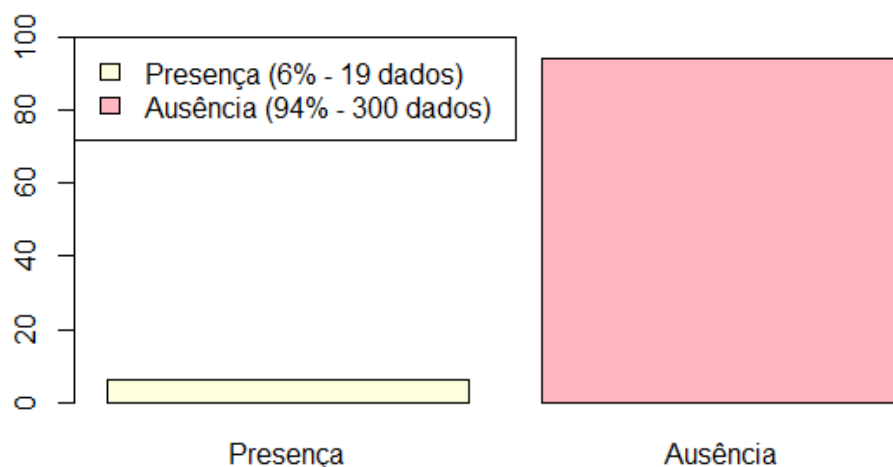
RAD	Medial		Final		TOTAL	
-	Nº/T	%	Nº/T	%	Nº/T	%
Compreender	28/42	67%	14/42	33%	42/319	13%
Entender	73/104	70%	31/104	30%	104/319	33%
Perceber	9/10	90%	1/10	10%	10/319	3%
Saber	63/72	87,5%	9/72	12,5%	72/319	23%
Ver	67/91	74%	24/91	26%	91/319	28%
TOTAL	240/319	75%	79/319	25%	319/319	100%

Fonte: própria

Observando a tabela, notamos que todos os RAD ocorrem, majoritariamente, em posição medial, até mesmo os RAD provenientes de *perceber*, que possuem apenas dez ocorrências. É possível concluir que a posição medial de turno é uma característica comum aos RAD analisados. Em posição final, em contrapartida, sobressaem-se os RAD originados de *entender* (31 – 30%), de *compreender* (14 – 33%) e de *ver* (24 – 26%).

4.2.3 Presença/ausência de respostas

Na conversação, é comum que o ouvinte dê indícios, verbais ou não verbais, de que está atento ao discurso do falante. Esses estímulos, na maioria das vezes, são itens que integram o subconjunto dos *feedbacks* – *ah*, *ahn*, *uhn uhn* etc. (URBANO, 2019); no entanto, na modalidade escrita, tais recursos não costumam ser encontrados, motivo pelo qual não foram controlados nesta pesquisa. Optamos, então, pelo controle de presença ou ausência de respostas, que podem ser plenas (*sim*, *não*, *claro* etc.) ou com repetição do verbo fonte (*sei*, *entendi*, *compreendo* etc.). O gráfico 4 expõe os resultados dessa variável.

Figura 4 – Presença/ausência de respostas após o uso dos RAD

Fonte: própria

A hipótese a respeito da presença/ausência de respostas – segundo a qual, na maioria das ocorrências, não haveria presença de respostas após o uso dos RAD, sinalizando o esvaziamento da ilocução interrogativa desses elementos e reforçando o seu estatuto de MD – se confirma. Das 319 ocorrências de RAD no *corpus*, apenas 19 – 6% são seguidas por respostas. Os excertos abaixo exemplificam, respectivamente, um caso de resposta plena e um caso de resposta com repetição de verbo – os RAD estão destacados e as repostas sublinhadas.

(52)

ESCUDEIRO 2 - Pois é. Então, de acordo com eles, o núcleo atômico... não é uma estrutura rígida, **compreendeu?**

PAI - Perfeitamente. Está claro. (*O novo sistema* – Hilda Hilst, 1968).

(53)

LULA - Claro que não, mas... a gente tinha menos que isto. Foi uma conquista. É preciso lutar pra conservar. **Entende?**

MALU - Entendo não. Lutar pra conservar uma coisa que a gente devia acabar com ela. Conservar o que devia ser destruído. (*A invasão* – Dias Gomes, 1960).

Cumprе destacar que os RAD provenientes de *perceber*, de *saber* e de *ver* não são seguidos por respostas em nenhuma de suas ocorrências, como é possível verificar na tabela 5.

Tabela 5 - Distribuição dos RAD de acordo com a presença/ausência de respostas

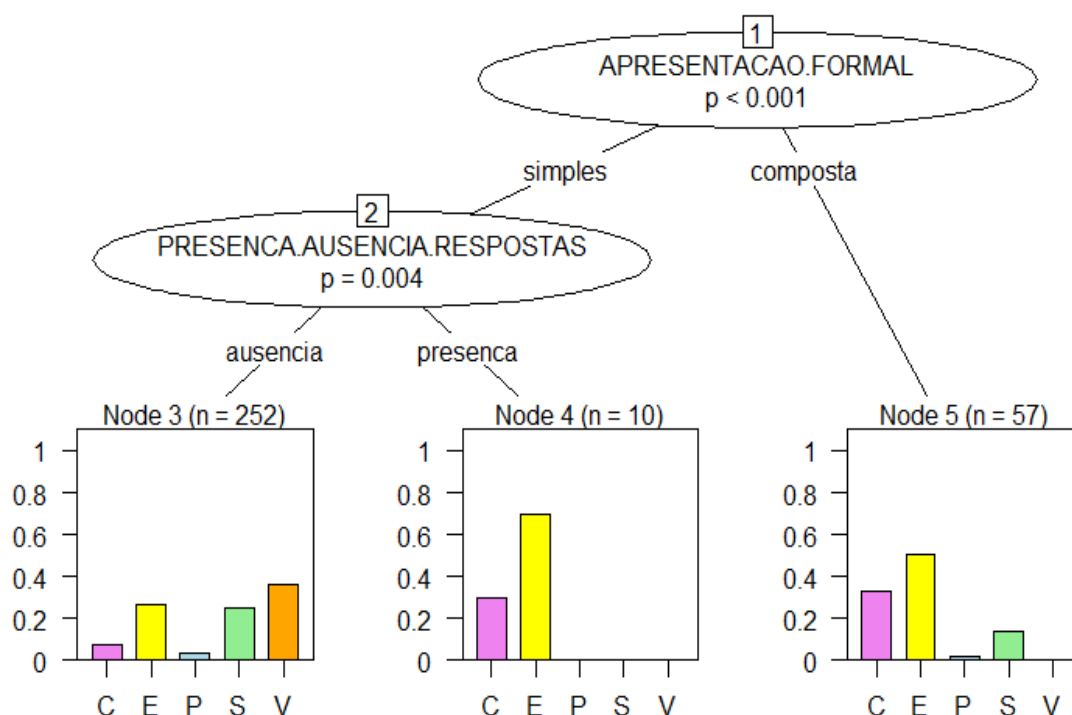
RAD	Ausência de respostas		Presença de respostas		TOTAL	
	Nº/T	%	Nº/T	%	Nº/T	%
-						
Compreender	35/42	83%	7/42	17%	42/319	13%
Entender	92/104	88%	12/104	12%	104/319	33%
Perceber	10/10	100%	0/10	0%	10/319	3%
Saber	72/72	100%	0/72	0%	72/319	23%
Ver	91/91	100%	0/91	0%	91/319	28%
TOTAL	300/319	94%	19/319	6%	319/319	100%

Fonte: própria

Dos 19 RAD sucedidos por respostas, todos são derivados de *compreender* e de *entender* e ocorrem, quase integralmente, em posição final, relação que fica clara ao realizarmos o cruzamento das variáveis posição e presença/ausência de respostas, tópico da próxima subseção.

4.2.4 Cruzamentos: variáveis de natureza linguística

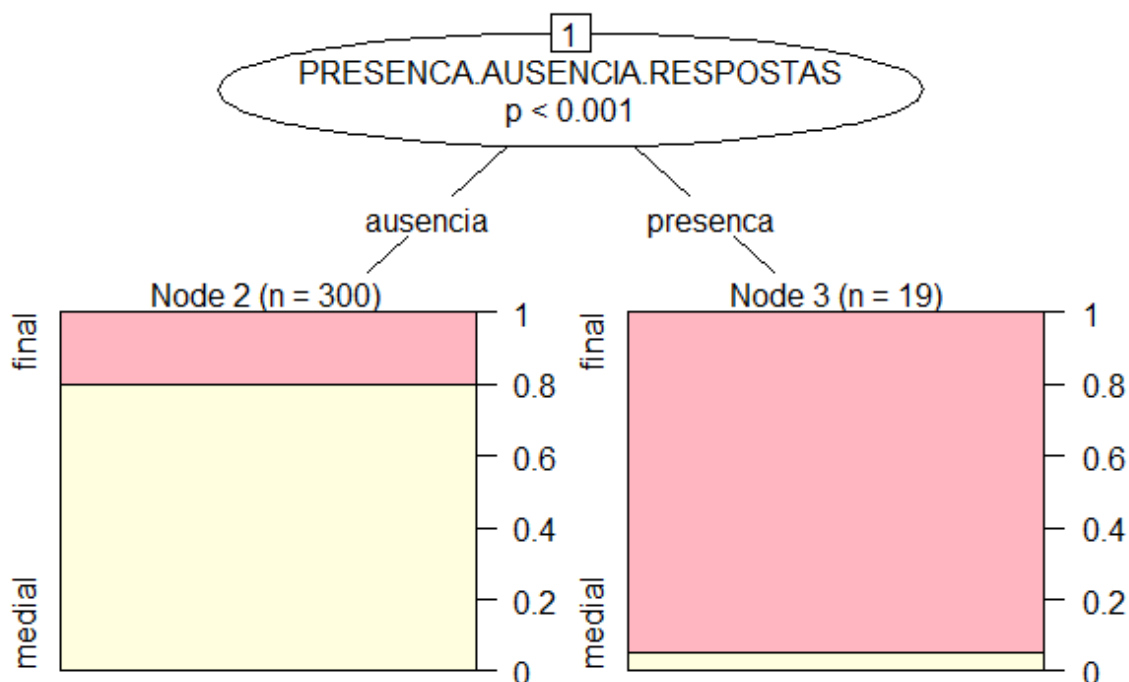
A figura 5 representa o cruzamento das variáveis apresentação formal, posição e presença/ausência de respostas. A partir desse cálculo, é possível visualizar a relevância dessas variáveis sobre o fenômeno em estudo, isto é, sobre a nossa variável dependente. Semelhantemente aos testes de qui-quadrado, considera-se relevante a variável cujo valor-*p* não excede 0,05 (5%).

Figura 5 - Diagrama arbóreo: variáveis linguísticas⁴³

Fonte: própria

Ao incluirmos as três variáveis de natureza linguística no modelo, o gráfico revela que a distribuição está sendo condicionada pela apresentação formal e pela presença/ausência de respostas, que se mostram mais significativas do que a variável posição. Temos, então, o seguinte padrão no que diz respeito às formas compostas dos RAD: *entender* > *compreender* > *saber* > *perceber* e não há formas compostas originadas de *ver*. Os RAD que possuem forma simples, por outro lado, mantêm interação significativa com a presença/ausência de respostas. Nos casos em que há ausência de respostas, sobressaem-se os RAD oriundos de *ver*, *entender* e *saber* e, quando há presença de respostas, destacam-se os RAD derivados de *compreender* e de *entender*. Na seção anterior, apontamos que há 19 RAD sucedidos por respostas e, nesta seção, ao correlacionar essa variável com a apresentação formal, descobrimos que 10 desses 19 dados possuem forma simples. Esses resultados indicam que a presença de respostas após o uso dos RAD não é motivada pela extensão da forma, mas, possivelmente, pela posição em que o marcador se encontra no turno. Nesse sentido, trataremos, agora, da relação entre a posição e a presença/ausência de respostas; a figura 6 traz os resultados desse cruzamento.

⁴³ C = compreender; E = entender; P = perceber; S = saber e V = ver.

Figura 6 - Diagrama arbóreo: posição e presença/ausência de respostas

Fonte: própria

Com base nesse gráfico, percebemos que a ausência ou presença de respostas está intimamente relacionada à posição do RAD no turno. Os RAD que são seguidos por respostas ocorrem, quase em sua totalidade, na posição final de turno; em contrapartida, os RAD que não apresentam respostas após seu uso estão, em grande parte, na posição medial. De acordo com Trapp (2014), a posição medial favorece a sustentação das falas, e a posição final sinaliza a passagem de turno e a mudança de papéis entre os interlocutores. Podemos inferir, portanto, que a posição final propicia a interrupção do fluxo discursivo, contribuindo para a troca de papéis entre os interlocutores e dando ênfase às relações interpessoais.

Em síntese, os RAD encontrados no *corpus* apresentam mais ocorrências de formas simples, ocupam, preferencialmente, a posição medial de turno e, na maior parte dos casos, não são sucedidos por respostas. Tendo em vista essas considerações, ao retomarmos o núcleo-piloto proposto pela GTI, vemos que os RAD em estudo possuem os traços mais estáveis dos MD.

Quadro 13 – Núcleo-piloto de MD proposto pela GTI

Núcleo-piloto
<i>Traço 1:</i> Exterioridade ao conteúdo proposicional
<i>Traço 2:</i> Independência sintática
<i>Traço 3:</i> Ausência de autossuficiência comunicativa
<i>Traço 4:</i> Maior projeção da articulação tópica
<i>Traço 5:</i> Maior projeção da orientação da interação

Fonte: Adaptado de Risso, Silva e Urbano (2019, p. 386)

Desse modo, acerca dos RAD em análise, consideramos que: eles não contribuem para o conteúdo referencial do discurso - até mesmo as formas compostas não atuam nesse plano -, mas operam no plano da enunciação, cooperando para a organização textual-interativa (**traço 1 - exterioridade em relação ao conteúdo proposicional**); esses RAD, incluindo as formas compostas, não desempenham funções importantes na estrutura oracional (**traço 2 - independência sintática**); eles não são comunicativamente autônomos, pois mesmo as formas mais extensas, como aquelas que são acompanhadas por pronomes, necessitam de um contexto comunicativo para que façam sentido (**traço 3 - ausência de autonomia comunicativa**); ainda que os RAD sejam, predominantemente, orientadores da interação, eles também operam, de maneira secundária, no plano textual, atuando como focalizadores, preenchedores, finalizadores, entre outras funções (cf. VALLE, 2014, p. 90) – neste caso, sobressaem-se os RAD mais prototípicos e mais gramaticalizados, caracterizados pela redução da massa fônica e pelo esvaziamento da ilocução interrogativa (**traço 5 - maior projeção da orientação da interação**). Os RAD, então, podem desempenhar diversos papéis no discurso.

4.3 Variáveis de natureza extralinguística

A presente seção se dedica à discussão dos resultados das variáveis extralinguísticas. Antes disso, apresentamos, no quadro 14, os valores de significância de cada uma dessas variáveis sobre os RAD, obtidos por meio de testes de qui-quadrado.

Quadro 14 – Valores dos testes de qui-quadrado: variáveis extralinguísticas⁴⁴

Variáveis extralinguísticas	Qui-quadrado
Autoria	X-squared = 302.84, df = 52, p-value < 2.2e ⁻¹⁶
Ambientação	X-squared = 67.688, df = 4, p-value = 6.979e ⁻¹⁴
Classe social	X-squared = 44.306, df = 12, p-value = 1.354e ⁻⁰⁵
Fase da vida (1964 - 1970)	X-squared = 33.187, df = 8, p-value = 5.698e ⁻⁰⁵
Etnia	X-squared = 16.378, df = 4, p-value = 2.551e ⁻⁰³
Instrução	X-squared = 26.049, df = 12, p-value = 1.057e ⁻⁰²
Fase da vida (1950 - 1963)	X-squared = 7.2049, df = 4, p-value = 1.254e ⁻⁰¹
Relação entre os personagens/interlocutores (simétrica vs. assimétrica)	X-squared = 7.0408, df = 4, p-value = 1.337e ⁻⁰¹
Relação entre os personagens/interlocutores (íntima vs. não íntima)	X-squared = 4.9497, df = 4, p-value = 2.925e ⁻⁰¹
Gênero/sexo	X-squared = 2.5344, df = 4, p-value = 6.385e ⁻⁰¹

Fonte: própria

Dentre as variáveis de natureza extralinguística, não foram consideradas significativas: gênero/sexo, relação entre os personagens/interlocutores – íntima vs. não íntima e simétrica vs. assimétrica – e fase da vida (1950 - 1963), cujos valores são superiores a 0,05 (5%). Ainda assim, discorreremos brevemente sobre essas variáveis nas próximas seções.

4.3.1 As macrocategorias sociais: o perfil dos personagens

Tratamos, a seguir, das variáveis utilizadas para categorizar o perfil social dos personagens das peças teatrais, avaliando cada uma individualmente. Antes de exibirmos os resultados, é importante destacar que os personagens neutralizados (NA) por falta de informação foram desconsiderados nos gráficos e no cruzamento das variáveis com os RAD; entretanto, apresentamos, em notas de rodapé, a quantidade de personagens excluídos da análise. Vale lembrar, também, que nossa hipótese é que – por se tratarem de textos teatrais, que buscam retratar indivíduos reais, bem como seus costumes e usos linguísticos – personagens com certos perfis sociais usarão alguns RAD em detrimento de outros.

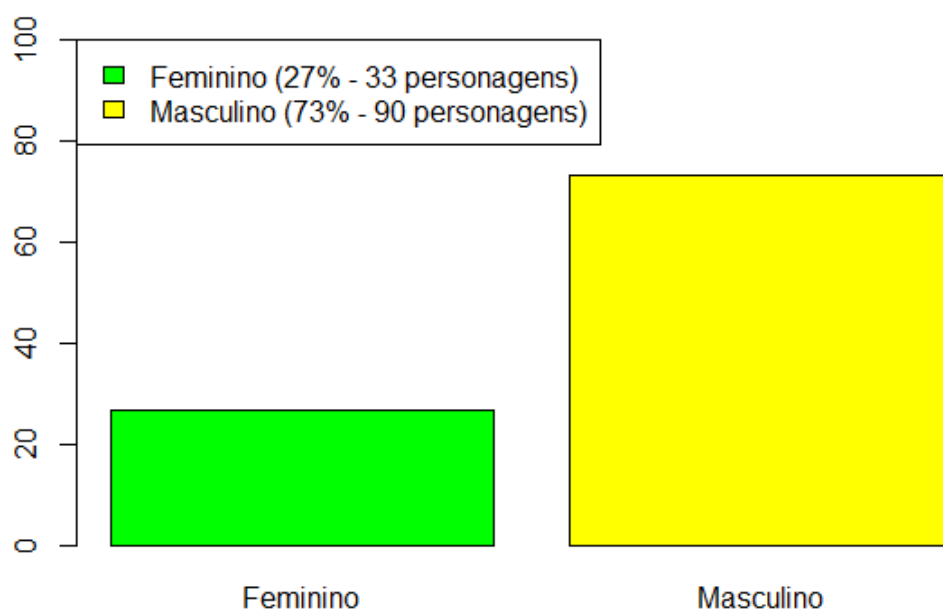
4.3.1.1 Gênero/sexo

Conquanto essa variável não seja estatisticamente significativa em relação ao uso dos RAD nas peças, conforme apontado pelo teste de qui-quadrado (cf. Quadro 14), faremos uma breve discussão, de cunho mais qualitativo, a respeito da representação dos gêneros/sexos nas

⁴⁴ O quadro foi organizado de acordo com os valores de significância das variáveis (*p-value*) – da mais significativa para a menos significativa.

peças teatrais. Dessa maneira, a figura 7 apresenta a distribuição dos personagens de acordo com seu gênero/sexo.

Figura 7 – Distribuição dos personagens de acordo com o gênero/sexo⁴⁵



Fonte: própria

Dos 123 personagens que utilizam RAD em suas falas, 33 – 27% são do sexo feminino e 90 – 73% do sexo masculino. Esse resultado desproporcional pode ser reflexo da baixa representatividade das vozes femininas dentro e fora do meio literário nos séculos passados, revelando a desigualdade entre os gêneros no contexto histórico brasileiro, situação que foi se modificando ao longo do tempo, com a emancipação das mulheres e a alteração dos papéis desempenhados por elas na sociedade (BRANDÃO, 2022).

Busatto (2003), em um estudo sobre a representação da mulher em peças de teatro produzidas no início do século XX, verificou que, mesmo em peças de autoria feminina, não havia uma representação de mulheres instigantes, que infringiam as normas da época. Segundo ela, o teatro desse período “[...] não era um teatro transgressor, comprometido com as transformações do seu tempo, era antes um teatro comportado, e assim bem aceito, teatro ao gosto burguês” (BUSATTO, 2003, p. 243). Por outro lado, nas peças que integram o nosso *corpus*, produzidas entre 1950 e 1970, notamos que o papel de personagens femininas começa a sofrer alterações, pois as mulheres passam a ocupar ambientes externos e a assumir novas responsabilidades, para além do meio doméstico. Em consonância com Busatto (2003, p. 231),

⁴⁵ Em relação à variável gênero/sexo, não há personagens neutralizados (NA).

[...] o trabalho feminino fora do lar ainda era hostilizado pela família e sociedade. Os pais desejavam que suas filhas encontrassem um bom partido para casar e que permanecessem no lar [...]. No entanto, o avanço industrial e urbano, o acesso à escolaridade, a veiculação pela imprensa dos espaços ocupados pelas mulheres após a Primeira Guerra Mundial e o crescente movimento feminista garantiram à mulher diferentes postos em funções consideradas adequadas ao público feminino: professora, enfermeira, datilógrafa, taquígrafa, secretária, telefonista, operária das indústrias têxtil, de confecções e alimentícia.

Há, portanto, um processo gradual de transformação e de emancipação da mulher. A título de ilustração, nas peças que compõem o *corpus*, encontramos mulheres que se unem a movimentos contrários à ditadura militar, mulheres que se opõem a determinados sistemas e manifestam sua opinião com ousadia, mulheres que se submetem a trabalhos indesejáveis para sustentar a família e mulheres que se dedicam ao cuidado do lar, dos filhos e do marido. Destarte, mesmo que tenhamos baixa representatividade de vozes femininas nas peças, percebemos que as personagens femininas assumem papéis mais marcantes e adquirem perfis sociais mais diversificados.

4.3.1.2 Fase da vida

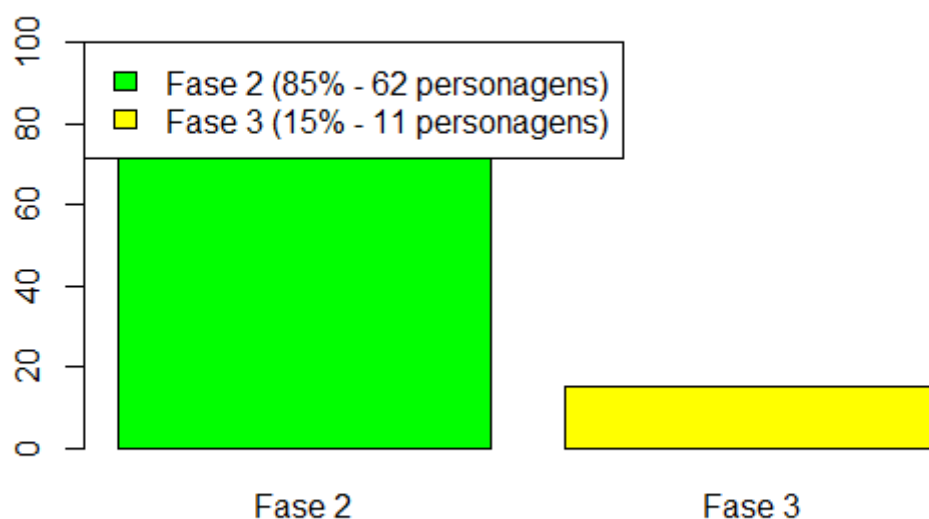
Para nortear as discussões, retomamos, a seguir, os critérios referentes à fase da vida, lembrando que, devido às mudanças de papéis e de funções desempenhados por indivíduos em diferentes fases da vida, essa variável sofre uma pequena alteração de um momento para outro, a fim de se adequar à realidade social de cada um dos períodos considerados pelos integrantes do projeto *Língua EmCena* (1831-1889; 1890-1930; 1931-1964 e 1965-2010). Dessa maneira – dado que nosso recorte temporal (1950-1970) abarca o período no qual ocorre a ramificação da fase 2 (a partir de 1964) –, de 1950 a 1963, categorizamos os personagens em fases 1, 2 e 3, e, de 1964 a 1970, os personagens foram categorizados em fases 1, 2a, 2b e 3. Destacamos, mais uma vez, que não estamos propondo que a variável fase da vida tenha passado por transformações significativas entre essas décadas, mas decidimos manter a ramificação para que o trabalho esteja alinhado às decisões tomadas no âmbito do projeto *Língua EmCena*.

Quadro 15 – Critérios da variável fase da vida

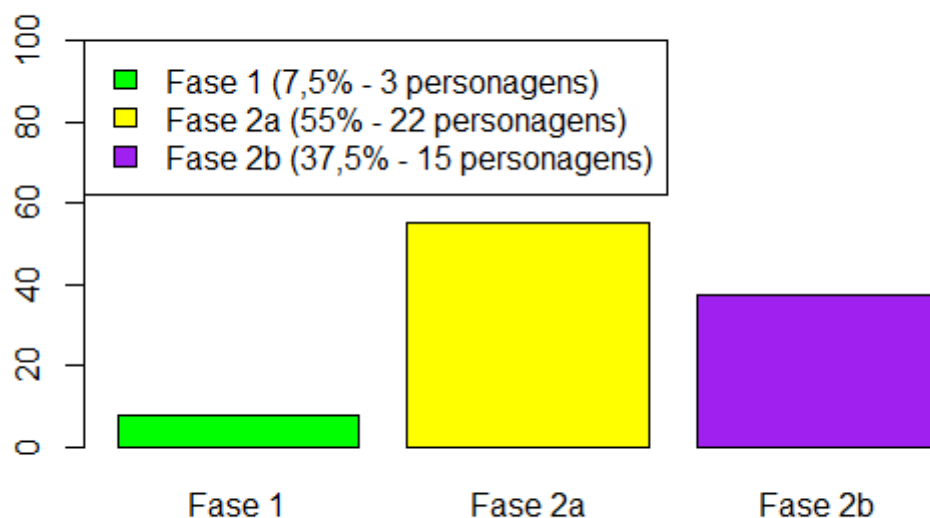
Períodos	Fases	Idades
1950 - 1963	Fase 1	de 0 a 15 anos
	Fase 2	de 16 a 50 anos
	Fase 3	de 51 em diante
	NA	não se aplica
1964 - 1970	Fase 1	de 0 a 17 anos
	Fase 2a	de 18 a 30 anos
	Fase 2b	de 31 a 59 anos
	Fase 3	de 60 em diante
	NA	não se aplica

Fonte: própria

Os testes de qui-quadrado (cf. Quadro 14) mostraram que a variável fase da vida (1950 – 1963) não é estatisticamente relevante, diferentemente da variável fase da vida (1964 – 1970), que foi considerada significativa. Ao verificar os resultados do processamento estatístico dessas variáveis, percebemos que, no primeiro período (1950 – 1963), não há personagens de fase 1 e, no segundo período (1964 – 1970), não há personagens de fase 3, como é possível observar nas figuras 8 e 9.

Figura 8 - Distribuição dos personagens de acordo com a fase da vida (1950 – 1963)⁴⁶**Fonte:** própria

⁴⁶ Seis personagens foram neutralizados (NA).

Figura 9 - Distribuição dos personagens de acordo com a fase da vida (1964 – 1970)⁴⁷

Fonte: própria

As figuras revelam que a maior parte dos personagens se distribui entre as fases 2 (62 – 85%), 2a (22 – 55%) e 2b (15 – 37,5%), que correspondem aos jovens e aos adultos. Crianças e adolescentes, que são incluídos na fase 1 (3 – 7,5%), e idosos, que correspondem à fase 3 (11 – 15%), não foram os principais produtores de RAD nas peças selecionadas.

Em se tratando da correlação dessa variável com o uso dos RAD, obtivemos os seguintes resultados:

Tabela 6 – Distribuição dos RAD de acordo com a fase da vida dos personagens

Períodos	Fases	Compreender		Entender		Perceber		Saber		Ver		TOTAL	
		Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%
-	-												
1950 a 1963	Fase 2	15/160	9%	47/160	29%	6/160	4%	32/160	20%	60/160	38%	160/295	54%
	Fase 3	4/15	27%	5/15	33%	1/15	7%	0/15	0%	5/15	33%	15/295	5%
1964 a 1970	Fase 1	3/4	75%	1/4	25%	0/4	0%	0/4	0%	0/4	0%	4/295	1%
	Fase 2a	1/61	2%	26/61	42%	1/61	2%	23/61	38%	10/61	16%	61/295	21%
	Fase 2b	16/55	29%	24/55	43%	2/55	4%	12/55	22%	1/55	2%	55/295	19%
TOTAL		39/295	13%	103/295	35%	10/295	3%	67/295	23%	76/295	26%	295/295	100%

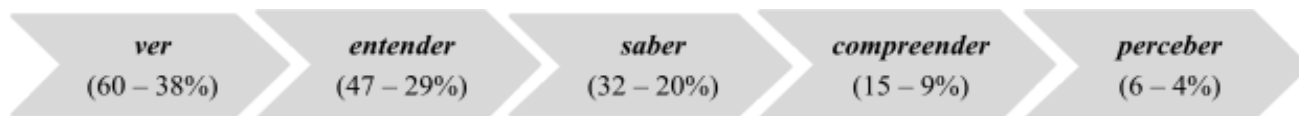
Fonte: própria

Para não nos atermos apenas aos personagens de fases 2, 2a e 2b, que foram responsáveis pela maior produção de RAD, discorreremos sobre os usos dos personagens pertencentes a todas as fases. Os resultados demonstram que o uso de RAD entre personagens de fase 1 é quase inexistente, constituindo apenas 1% das ocorrências. Esses personagens utilizam apenas dois tipos de RAD, os derivados de *compreender* (3 – 75%) e os derivados de *entender* (1 – 25%), porém, devido à baixa incidência de produção de RAD, não é possível

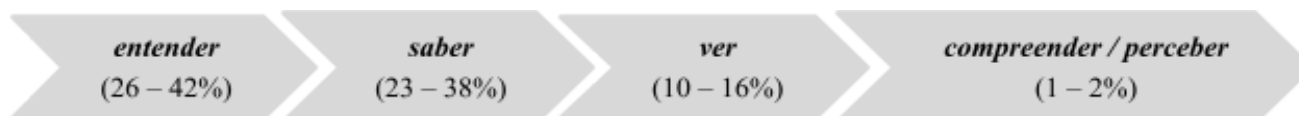
⁴⁷ Quatro personagens foram neutralizados (NA).

fazer afirmações mais substanciais. Quanto aos personagens das demais fases, percebemos a seguinte ordem de produção:

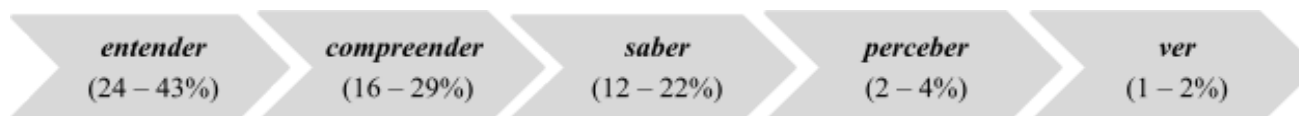
Fase 2:



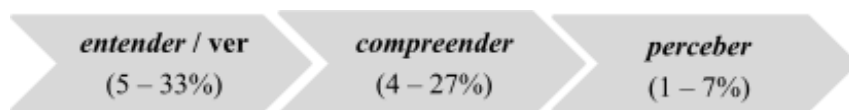
Fase 2a:



Fase 2b:



Fase 3:



Acerca desses resultados, é interessante sinalizar alguns aspectos. Em primeiro lugar, diferentemente dos personagens de fases 2a e 2b, que fazem maior uso dos RAD provenientes de *entender*, os personagens de fase 2 produzem mais o RAD *viu?*. Em segundo lugar, personagens de fase 3 produzem RAD oriundos de *entender* e de *ver* na mesma proporção (33%), assemelhando-se, ao mesmo tempo, às fases 2a e 2b e à fase 2. Em terceiro lugar, nos casos das fases 2 e 2a, as maiores ocorrências são dos RAD derivados de *entender*, *saber* e *ver*; todavia, os personagens de fase 2b tiveram como segunda maior produção os RAD originados de *compreender*, equiparando-se aos personagens de fases 1 e 3. Esse resultado é notável, pois revela uma relação entre a fase mais jovem (fase 1) e as fases mais velhas (fases 2b e 3). É válido pontuar, também, que, conforme exposto no quadro de critérios da variável fase da vida (cf. Quadro 15), a fase 2b abrange personagens de 31 a 59 anos, ao passo que a fase 3 agrega os de 51 anos em diante; é possível, então, que a semelhança de produção de RAD entre as fases 2b e 3 seja resultado da proximidade de idades.

4.3.1.3 Instrução

Os critérios propostos para essa variável estão sintetizados no quadro 16.

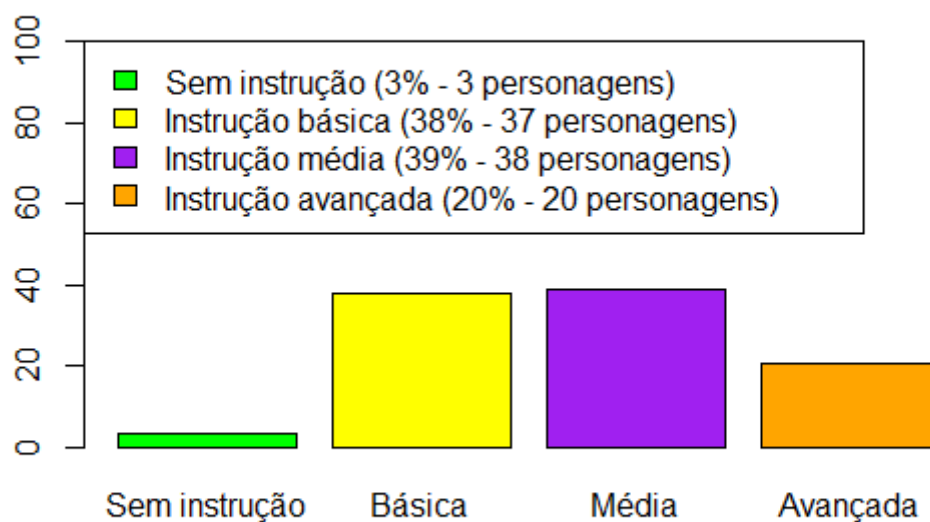
Quadro 16 – Critérios da variável instrução

Instrução	Sem instrução	Personagens analfabetos e sem qualquer indício de instrução formal
	Instrução básica	Personagens que sabem ler ou escrever
	Instrução média	Personagens que possuem especializações ou conhecimentos específicos
	Instrução avançada	Personagens que cursam o ensino superior ou têm profissões que exigem nível superior de escolaridade
	NA	Não se aplica

Fonte: própria

A categorização do nível de instrução dos personagens revelou que a maior parte deles possui algum tipo de formação identificável.

Figura 10 - Distribuição dos personagens de acordo com o nível de instrução⁴⁸



Fonte: própria

De acordo com o gráfico, os personagens se distribuem, principalmente, entre as instruções média (38 – 39%) e básica (37 – 38%); há uma quantidade significativa de personagens com instrução avançada (20 – 20%), e os sem instrução constituem uma minoria (3 – 3%). Ao correlacionar essa variável ao uso dos RAD, alcançamos os seguintes resultados:

⁴⁸ Vinte e cinco personagens foram neutralizados (NA).

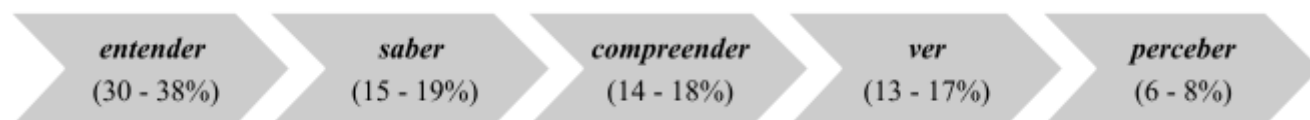
Tabela 7 - Distribuição dos RAD de acordo com o nível de instrução dos personagens

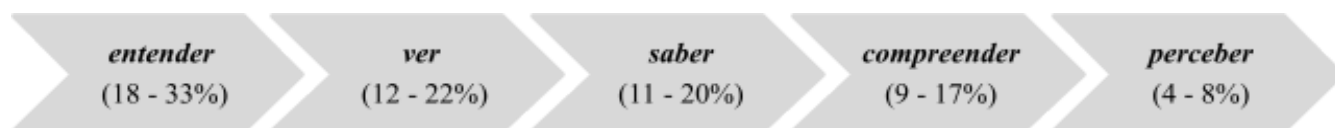
Instrução	Compreender		Entender		Perceber		Saber		Ver		TOTAL	
	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%
Sem instrução	0/17	0%	4/17	23,5%	0/17	0%	4/17	23,5%	9/17	53%	17/253	7%
Instrução básica	14/78	18%	30/78	38%	6/78	8%	15/78	19%	13/78	17%	78/253	31%
Instrução média	11/104	10%	34/104	33%	0/104	0%	33/104	32%	26/104	25%	104/253	41%
Instrução avançada	9/54	17%	18/54	33%	4/54	8%	11/54	20%	12/54	22%	54/253	21%
TOTAL	34/253	13%	86/253	34%	10/253	4%	63/253	25%	60/253	24%	253/253	100%

Fonte: própria

Assim como na subseção anterior, referente à variável fase da vida, vamos discutir cada um dos níveis de instrução separadamente, levando em consideração a hierarquia de produção dos RAD. Antes, porém, vale frisar que, ao observar os resultados da tabela 7 verticalmente, notamos que os RAD derivados de *compreender* e de *perceber* têm uma certa tendência aos níveis básico (*compreender* - 18% e *perceber* - 8%) e avançado (*compreender* - 17% e *perceber* - 8%) de instrução. Além disso, os RAD provenientes dessas duas fontes verbais não são usados por personagens sem instrução. O RAD *viu?*, em contrapartida, tende a se destacar na fala de personagens sem instrução (53%).

Partindo para a discussão dos resultados referentes a cada um dos níveis de instrução, temos a seguinte ordem de produção:

Sem instrução:**Instrução básica:****Instrução média:**

Instrução avançada:

Personagens sem instrução, como mencionado há pouco, não fazem uso de RAD derivados de *compreender* e de *perceber*; personagens com instrução média também não produzem RAD oriundos de *perceber*. De acordo com os esquemas apresentados, personagens sem instrução produziram mais os RAD derivados de *ver* (9 – 53%), seguido por *entender* e por *saber*, que foram usados na mesma proporção (4 – 23,5%). Personagens com instruções básica, média e avançada, por outro lado, foram os que mais produziram RAD derivados de *entender* e, diferentemente de personagens sem instrução, utilizaram RAD derivados de *compreender*.

Os RAD advindos de *entender* e de *compreender*, comparados aos demais, possuem mais material fonológico, aparentam ser semanticamente complexos, por estarem mais apegados ao sentido do verbo fonte e menos abstratizados, e apresentam maior variabilidade formal do que *ver*, por exemplo, que é uma forma curta e sem variação no *corpus*; por essa razão, é provável que os autores tenham associado os RAD derivados de *entender/compreender* a personagens com algum tipo de formação. Entretanto, personagens sem instrução também fazem uso dos RAD derivados de *entender*, demonstrando que, possivelmente, os RAD originados dessa fonte verbal não são tão influenciados pela variável instrução. Em contrapartida, os autores das peças parecem atribuir um caráter mais culto aos RAD provenientes de *compreender/perceber*, que não são usados por personagens sem instrução, e um caráter mais coloquial ao RAD *viu?*, que é o mais produzido por esses personagens.

4.3.1.4 Classe social

No quadro 17, sintetizamos os critérios considerados para a variável classe social.

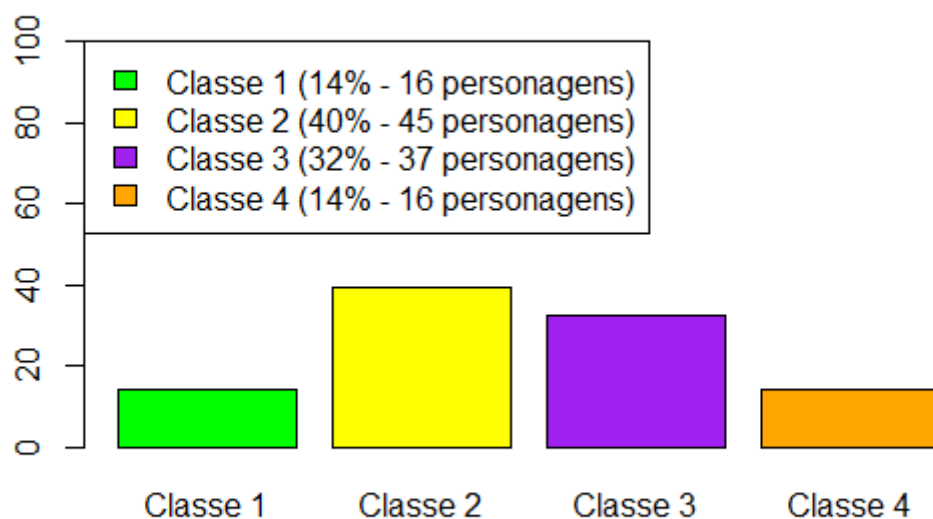
Quadro 17 – Critérios da variável classe social

Classe social	Classe 1	alta sociedade / alta burguesia / classe dominante
	Classe 2	classes médias / pequena burguesia / classes emergentes
	Classe 3	classe subalterna / proletariado / camponeses
	Classe 4	oprimidos / marginais
	NA	não se aplica

Fonte: própria

O próximo gráfico apresenta a distribuição dos personagens em consonância com a classe social à qual pertencem.

Figura 11 - Distribuição dos personagens de acordo com a classe social⁴⁹



Fonte: própria

Os personagens das peças se distribuem, predominantemente, entre as classes 2 (45 – 40%) e 3 (37 – 32%). A quantidade de personagens pertencente às classes 1 e 4 é a mesma: 16 – 14%. Abaixo, podemos verificar a correlação da classe social com os RAD em estudo.

Tabela 8 - Distribuição dos RAD de acordo com a classe social dos personagens

Classe social	Compreender		Entender		Perceber		Saber		Ver		TOTAL	
	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%
-												
Classe 1	6/34	18%	18/34	53%	2/34	6%	3/34	9%	5/34	14%	34/305	11%
Classe 2	19/115	17%	44/115	38%	7/115	6%	24/115	21%	21/115	18%	115/305	38%
Classe 3	11/115	9,5%	24/115	21%	1/115	1%	30/115	26%	49/115	42,5%	115/305	38%
Classe 4	3/41	7%	17/41	42%	0/41	0%	14/41	34%	7/41	17%	41/305	13%
TOTAL	39/305	13%	103/305	34%	10/305	3%	71/305	23%	82/305	27%	305/305	100%

Fonte: própria

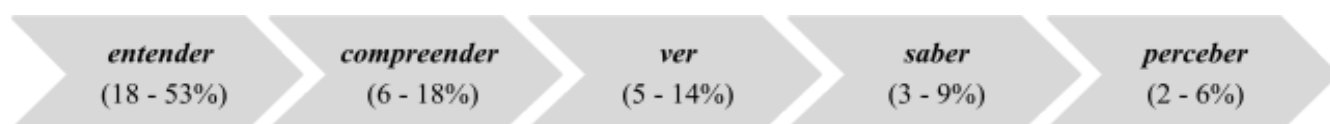
Ao olharmos para a tabela 8 no sentido vertical, notamos algumas particularidades. Primeiramente, os RAD derivados de *compreender* tendem a ser usados por personagens de classes 1 (18%) e 2 (17%); os RAD provenientes de *perceber*, além de demonstrarem certa tendência de uso pelas classes 1 e 2, que apresentam a mesma proporção (6%), não são usados pela classe 4, fator que os distingue de *compreender*, que foi usado por alguns personagens dessa classe (7%). Os RAD originados de *saber* e de *ver*, por sua vez, quando comparados aos

⁴⁹ Nove personagens foram neutralizados (NA).

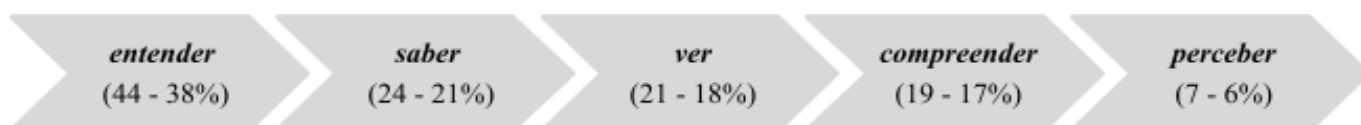
outros, apresentam resultados contrastantes: os RAD derivados de *saber* tendem a ser usados por personagens de classe 4 (34%) e o RAD *viu?*, por personagens de classe 3 (42,5%). Os RAD oriundos de *entender*, diferindo dos demais, tendem a ser utilizados por todos os segmentos sociais; embora manifeste certa tendência de uso por personagens de classe 1 (53%), notamos que esta é uma variante produtiva na fala de todas as classes e aparenta ser menos influenciada por aspectos sociais. Podemos, por meio desses resultados – e a partir das discussões desenvolvidas na seção concernente ao nível de instrução dos personagens –, inferir que haja uma certa distinção social entre *compreender/perceber* e *saber/ver*, mas seria necessário aprofundar essa análise para fazermos afirmações mais precisas.

Diante disso, observamos as seguintes ordens de produção:

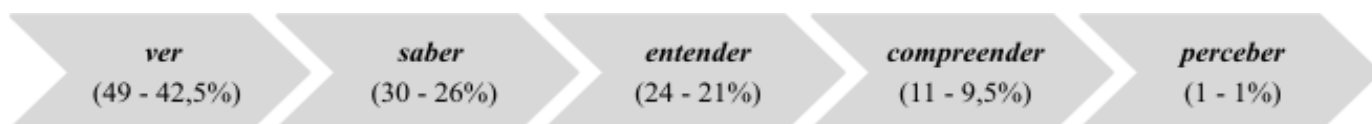
Classe 1:



Classe 2:



Classe 3:



Classe 4:



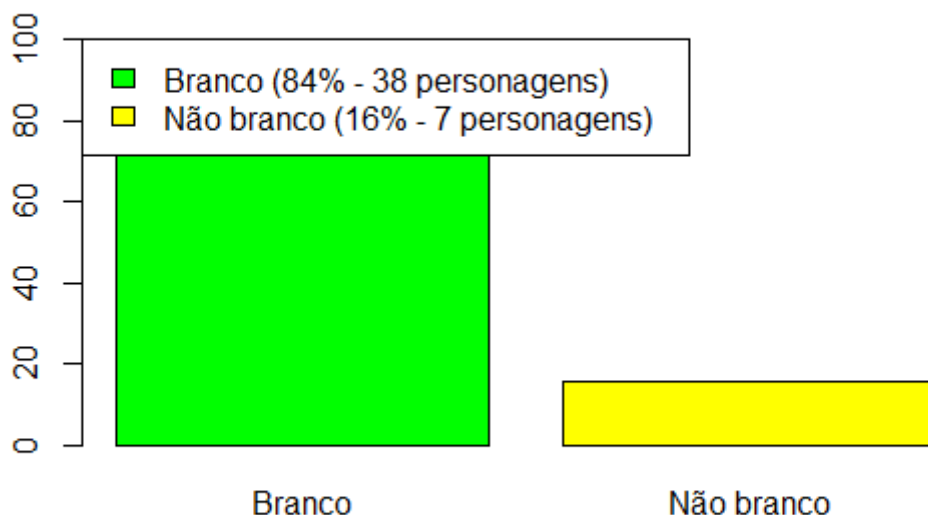
Personagens de classes 1, 2 e 4 fazem mais uso dos RAD originados de *entender*, corroborando a nossa hipótese de que esse RAD seja menos influenciado por fatores sociais, enquanto personagens de classe 3 produzem mais o RAD *viu?*. Vale frisar, ainda, que, diferentemente dos personagens pertencentes às classes 2, 3 e 4, as quais têm como segunda maior produção os RAD provenientes de *saber*, personagens de classe 1 produzem mais os RAD derivados de *compreender* (6 – 18%), reforçando a ideia de que pode haver uma certa

distinção social relacionada ao uso desse RAD. O que se percebe, então, é que personagens de classes altas (classes 1 e 2) podem usar os mesmos RAD que personagens de classes baixas (classe 4), mas, ainda assim, há certas tendências de uso que, conforme demonstramos anteriormente, não podem ser ignoradas.

4.3.1.5 Etnia

Anteriormente, na discussão dessa variável, declaramos que informações concernentes à raça dos personagens nem sempre estão presentes nas peças e, muitas vezes, não é possível realizar inferências a respeito dessa categoria. Em decorrência disso, no tocante à etnia, os personagens, quase que de maneira geral, tiveram de ser neutralizados e rotulados como NA.

Figura 12 - Distribuição dos personagens de acordo com a etnia⁵⁰



Fonte: própria

A figura demonstra que, dos 123 personagens, apenas 45 tiveram suas etnias especificadas, a saber: 38 – 84% brancos e 7 – 16% não brancos. Ainda assim, correlacionamos essa variável ao uso dos RAD, por ter sido considerada significativa pelo teste de qui-quadrado (cf. Quadro 14).

Tabela 9 - Distribuição dos RAD de acordo com a etnia dos personagens

Etnia	Compreender		Entender		Perceber		Saber		Ver		TOTAL	
	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%
Branco	15/87	17%	43/87	50%	6/87	7%	14/87	16%	9/87	10%	87/106	82%
Não branco	0/19	0%	6/19	32%	0/19	0%	5/19	26%	8/19	42%	19/106	18%
TOTAL	15/106	14%	49/106	46%	6/106	6%	19/106	18%	17/106	16%	106/106	100%

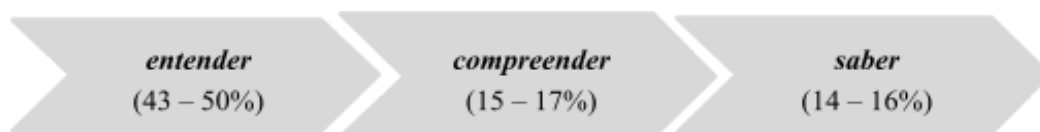
Fonte: própria

⁵⁰ Setenta e oito personagens foram neutralizados (NA).

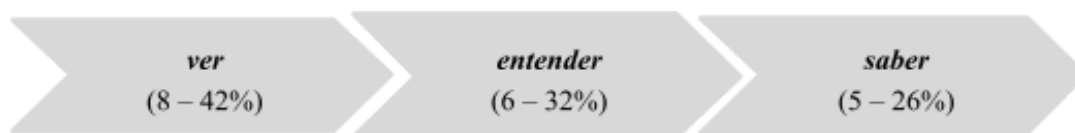
Mais uma vez, ao observarmos os resultados verticalmente, percebemos certas tendências de uso. Personagens brancos tendem a usar os RAD derivados de *compreender* (17%) e de *perceber* (7%), os quais não são utilizados por personagens não brancos. Por outro lado, *saber* e *ver* contrastam, novamente, com *compreender* e *perceber*, pois a tendência é que sejam produzidos por personagens não brancos (*saber* - 26% e *ver* - 42%).

Levando em consideração os três RAD mais usados por personagens brancos e não brancos, obtemos a seguinte distribuição:

Branco:



Não brancos:



Personagens brancos produzem mais os RAD derivados de *entender* (43 – 50%), ao passo que não brancos produzem mais o RAD *viu?* (8 – 42%). No entanto, personagens não brancos têm como segunda maior produção RAD advindos de *entender* (6 – 32%), demonstrando, mais uma vez, que essa variante pode ser produzida por personagens com diferentes perfis sociais; personagens brancos, por outro lado, têm como segunda maior produção os RAD originados de *compreender* (15 – 17%). Nesse caso, é válido retomar a discussão desenvolvida na seção destinada ao nível de instrução dos personagens: é possível que os RAD derivados de *compreender* – por terem mais material fonológico, apresentarem mais variabilidade e carregarem, de maneira acentuada, traços semânticos e morfológicos do verbo fonte – sejam percebidos como mais complexos, sendo associados a personagens com algum grau de instrução e, conseqüentemente, brancos.

Considerando-se o contexto do século XX, sabe-se que, por muito tempo, as questões raciais definiram quem teria ou não acesso à educação (BELTRÃO; NOVELLINO, 2002; BELTRÃO, 2005). Segundo Beltrão (2005, p. 78),

As desigualdades raciais são também desigualdades sociais e ficam flagrantes quando examinamos os dados relativos à escolaridade, desagregando-os pelas diferentes categorias de cor ou raça que compõem a população brasileira. Quanto maior o nível educacional, maior as

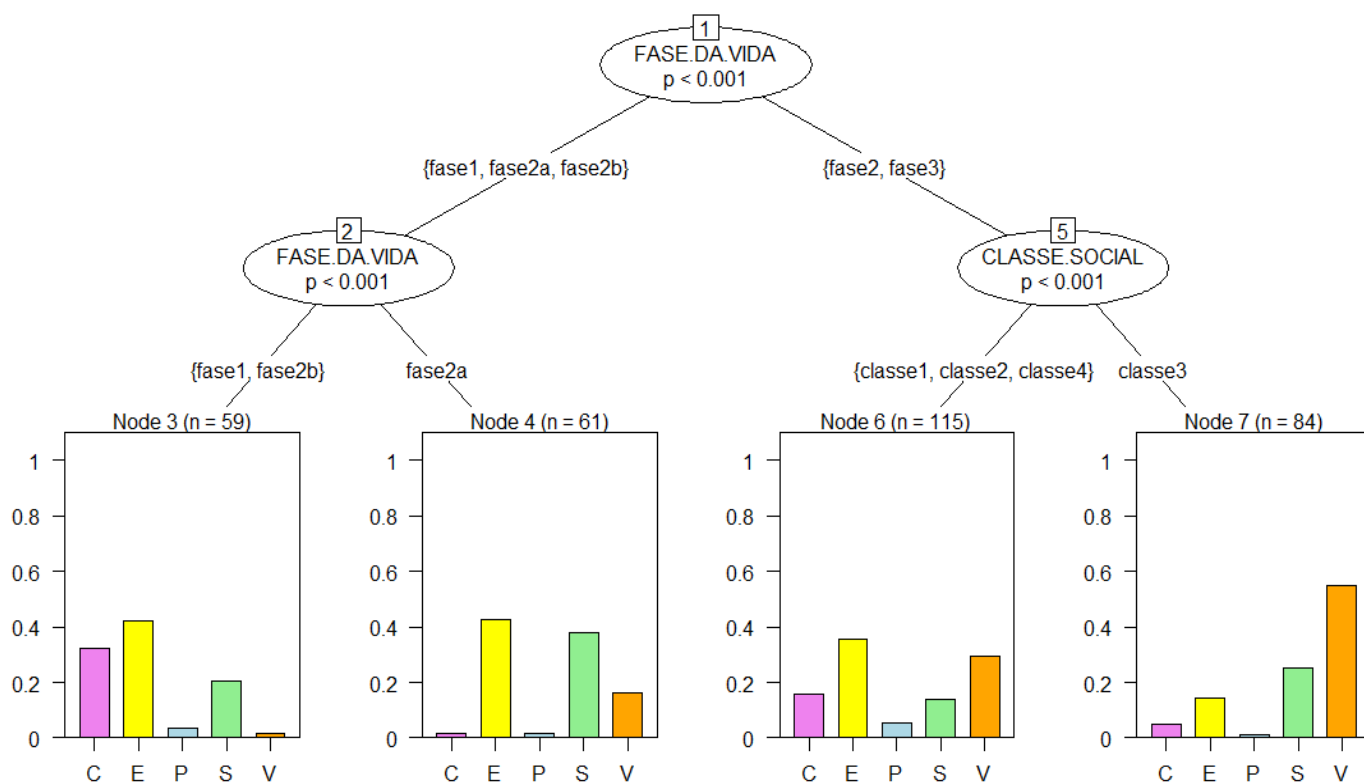
disparidades entre pretos, pardos e indígenas, de um lado, e brancos e amarelos, de outro.

À vista disso, realizamos, na próxima seção, o cruzamento entre a etnia e o nível de instrução dos personagens com o objetivo de avaliar se, de fato, há relações entre as duas variáveis.

4.3.1.6 Correlacionando as macrocategorias sociais

Tendo em vista as discussões desenvolvidas nas seções anteriores, decidimos realizar alguns cruzamentos entre as variáveis utilizadas para categorizar o perfil social dos personagens, isto é, gênero/sexo, fase da vida, instrução, classe social e etnia. Nesse cruzamento, optamos por amalgamar os dois períodos da variável fase da vida (1950-1963 e 1964-1970). No primeiro período, não há personagens de fase 1 e, no segundo período, não há personagens de fase 3; por esse motivo, no gráfico, os personagens de fase 1 pertencem às peças produzidas entre 1964 e 1970, e os personagens de fase 3 pertencem às peças escritas entre 1950 e 1963.

Figura 13 - Diagrama arbóreo: macrocategorias sociais – gênero/sexo, fase da vida, instrução, classe social e etnia



Fonte: própria

Ao incluir as cinco variáveis no modelo, o programa retorna um diagrama, destacando as variáveis mais significativas. Por meio desse gráfico, percebemos que a fase da vida é a variável mais relevante no que diz respeito ao uso dos RAD. Além disso, o padrão de RAD produzidos por personagens de fases 1, 2a e 2b difere do de personagens de fases 2 e 3. Personagens de fases 1 e 2b usam mais os RAD originados de *entender* e de *compreender*, e personagens de fase 2a utilizam mais *entender* e *saber*. As ocorrências de personagens de fases 2 e 3 estão intimamente relacionadas à classe social. Conforme podemos ver no gráfico, personagens de classes 1, 2 e 4 fazem mais uso de RAD derivados de *entender*, ao passo que personagens de classe 3 usam mais os RAD provenientes de *ver*.

Nas seções anteriores, ao olharmos verticalmente para as tabelas, notamos que *entender* aparenta ser o RAD menos segmentado no tocante às macrocategorias sociais. Ainda assim, o fato de os usos linguísticos de personagens de classe 4, a classe mais baixa, se equipararem aos de personagens de classes mais elevadas (classes 1 e 2) chamou nossa atenção; por essa razão, recorreremos à categorização do perfil social de personagens pertencentes à classe 4 para verificar quem são esses personagens e quais são os outros critérios usados em sua categorização.

Quadro 18 – Categorização de personagens de classe 4

PERSONAGENS	GÊNERO/SEXO	FASE DA VIDA	INSTRUÇÃO	CLASSE SOCIAL	ETNIA
Bichado	masculino	fase 2a	NA	classe4	NA
Chicó	masculino	fase 2	sem instrução	classe4	branco
Estudante	masculino	fase 2a	instrução avançada	classe4	NA
Garcia	masculino	fase 2a	instrução média	classe4	NA
Geni	feminino	fase 2b	instrução básica	classe4	NA
Homem	masculino	fase 2a	NA	classe4	NA
João Grilo	masculino	fase 2	sem instrução	classe4	não branco
Kátia	feminino	fase 2a	instrução básica	classe4	NA
Lula	masculino	fase 2	instrução média	classe4	branco
Mariana	feminino	fase 2a	instrução básica	classe4	NA
Maria-vai	feminino	fase 2a	NA	classe4	NA
Marieta	feminino	fase 2	NA	classe4	NA
Mulher	feminino	fase 2a	NA	classe4	NA
Nhanhá	feminino	fase 2b	NA	classe4	NA
Padre Maximilian	masculino	fase 2b	instrução avançada	classe4	NA
Tonho	masculino	fase 2a	instrução básica	classe4	NA

Fonte: própria

Por intermédio desse quadro, notamos que uma quantidade significativa de personagens de classe 4 possui algum tipo de instrução, seja básica, média ou avançada. A existência de personagens com instrução avançada em classes consideradas marginalizadas não condiz com o *status* social daquele período; contudo, ao reunirmos mais informações

sobre os personagens que possuem instrução avançada (*Estudante e Padre Maximilian*), descobrimos que ambos pertencem à mesma peça, são prisioneiros sentenciados à morte e, por isso, apesar de terem cursado o ensino superior, estão entre os personagens de classe mais baixa. A partir disso, podemos supor que a semelhança de produção de RAD entre a classe 4 e as classes mais abastadas ocorra em função desses personagens. Abaixo, segue a distribuição de RAD produzidos por personagens de classe 4.

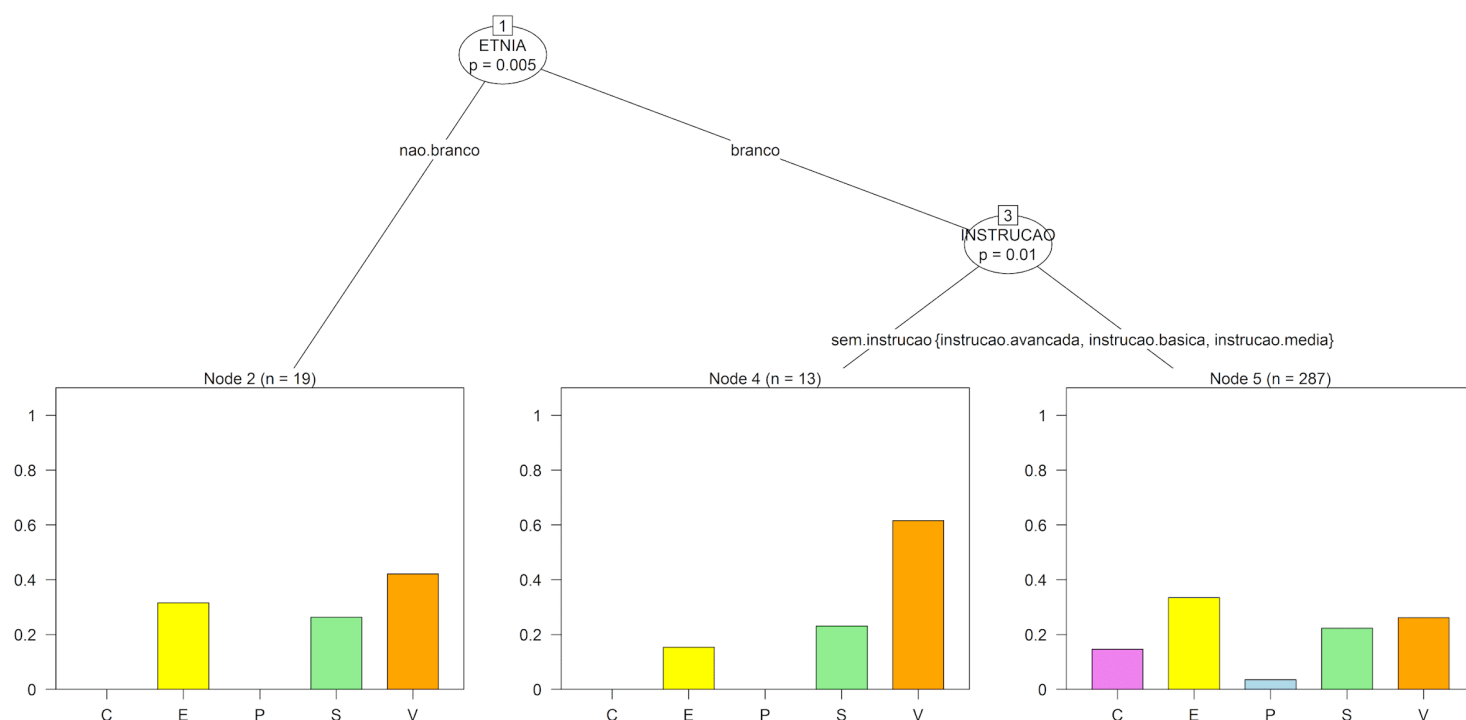
Tabela 10 – Distribuição dos RAD produzidos por personagens de classe 4

Instrução	Compreender		Entender		Perceber		Saber		Ver		TOTAL	
	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%
-												
Sem instrução	0/5	0%	3/5	60%	0/5	0%	1/5	20%	1/5	20%	5/28	18%
Instrução básica	0/11	0%	4/11	36%	0/11	0%	6/11	55%	1/11	9%	11/28	39%
Instrução média	0/2	0%	1/2	50%	0/2	0%	1/2	50%	0/2	0%	2/28	7%
Instrução avançada	3/10	30%	2/10	20%	0/10	0%	5/10	50%	0/10	0%	10/28	36%
TOTAL	3/28	11%	10/28	36%	0/28	0%	13/28	46%	2/28	7%	28/28	100%

Fonte: própria

Diferentemente do que esperávamos, ao observar a tabela 10, percebemos que personagens de classe 4 com instrução avançada produzem mais os RAD derivados de *saber* (5 - 50%). Observando a coluna de *entender* no sentido vertical, notamos que, na verdade, os RAD derivados de *entender* tendem a ser produzidos por personagens de classe 4 sem instrução (60%). Em suma, esses usos sugerem uma possível tendência, segundo a qual a produção de RAD não está correlacionada ao nível de instrução dos personagens, uma vez que não são os mais instruídos que fazem usos semelhantes aos dos personagens de classes altas, mas sim aqueles com menos instrução. Todavia, seria necessário ter um número maior de dados para chegar a uma conclusão.

Nessa mesma direção, realizamos o cruzamento das variáveis instrução e etnia.

Figura 14 - Diagrama arbóreo: etnia e instrução

Fonte: própria

O diagrama evidencia que a distribuição está sendo condicionada pela etnia, que se mostrou mais significativa do que o nível de instrução dos personagens. O que mais se destaca nesse gráfico é o fato de os RAD mais usados por personagens não brancos serem semelhantes aos produzidos por personagens brancos sem instrução (*ver, saber e entender*). Tal resultado leva-nos a questionar se personagens brancos analfabetos fariam do mesmo modo que personagens não brancos, os quais, nem sempre, tinham acesso à educação naquele período (cf. BELTRÃO; NOVELLINO, 2002; BELTRÃO, 2005). A fim de verificar o nível de instrução dos personagens não brancos, montamos um quadro, baseando-nos na categorização realizada previamente.

Quadro 19 – Categorização de personagens não brancos

PERSONAGENS	GÊNERO/SEXO	FASE DA VIDA	INSTRUÇÃO	CLASSE SOCIAL	ETNIA
Benedito	masculino	fase 2	NA	classe 3	não branco
Bonitão	masculino	fase 2	instrução básica	classe 3	não branco
Cancão	masculino	fase 2	instrução básica	classe 3	não branco
Dedé Cospe-rima	masculino	fase 2	instrução média	classe 3	não branco
João Grilo	masculino	fase 2	sem instrução	classe 4	não branco
Jorge	masculino	fase 2	instrução básica	classe 3	não branco
Marina	feminino	fase 2	instrução básica	classe 3	não branco

Fonte: própria

Dentre os personagens não brancos, apenas uma personagem é do gênero/sexo feminino; todos pertencem à fase 2 – são, portanto, jovens ou adultos – e às classes 3 e 4, ou seja, as classes mais baixas, que correspondem aos subalternos ou marginalizados. O fator interessante é que, com exceção de dois personagens, que foram categorizados como *sem instrução* e NA, todos possuem algum tipo de formação, básica ou média. A partir disso, podemos supor que o uso dos RAD não está diretamente atrelado a essas macrocategorias – que, certamente, são cruciais para a construção dos personagens –, mas sim à percepção dos autores em relação aos usos linguísticos de indivíduos reais. É provável, então, que, para os autores, os RAD derivados de *ver*, *saber* e *entender* seriam predominantes na fala de personagens não brancos e de personagens brancos sem instrução.

Personagens brancos com instruções básica, média e avançada, diferentemente dos demais, fazem uso de RAD derivados de *compreender* e de *perceber* e produzem, com mais recorrência, os RAD derivados de *entender*. No entanto, observando o gráfico, vemos que os RAD originados de *ver* e de *saber* também são usados de maneira significativa por esses personagens, ainda que em proporção menor que personagens não brancos e brancos sem instrução.

Por meio desses resultados, percebemos que nossa hipótese se confirma parcialmente, pois a produção de RAD não é diretamente influenciada por essas macrocategorias, resultado que condiz com as análises de estudos anteriores (SILVA; MACEDO, 1996; VALLE, 2001, 2014), mas essas variáveis, com exceção de gênero/sexo, são estatisticamente significativas (cf. Quadro 14), demonstrando que, nas peças teatrais, essas categorias norteiam o autor em suas escolhas linguísticas – principalmente porque são usadas na criação dos personagens.

4.3.2 Relação entre os personagens/interlocutores

Inicialmente, esperávamos que as relações entre os personagens e seus interlocutores motivassem a escolha de determinados RAD; por isso, verificamos o grau de intimidade e de simetria dessas relações para avaliar se a produção de RAD seria influenciada por esses contextos. Nossa hipótese era que os RAD seriam mais produzidos em contextos de intimidade e de simetria entre os personagens; todavia, ainda que a produção de RAD tenha sido bastante frequente nos referidos contextos, conforme exposto nas tabelas 11 e 12, os testes de qui-quadrado revelaram que essas variáveis não são significativas em relação ao fenômeno.

Tabela 11 – Distribuição dos RAD de acordo com as relações entre os personagens/interlocutores: íntimas e não íntimas

Tipo de relação	Compreender		Entender		Perceber		Saber		Ver		TOTAL	
	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%
-												
Íntima	20/171	12%	54/171	31%	8/171	5%	43/171	25%	46/171	27%	171/319	54%
Não íntima	22/148	15%	50/148	34%	2/148	1%	29/148	20%	45/148	30%	148/319	46%
TOTAL	42/319	13%	104/319	33%	10/319	3%	72/319	23%	91/319	28%	319/319	100%

Fonte: própria

Tabela 12 - Distribuição dos RAD de acordo com as relações entre os personagens/interlocutores: assimétricas e simétricas

Tipo de relação	Compreender		Entender		Perceber		Saber		Ver		TOTAL	
	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%
-												
Assimétrica	18/142	12%	41/142	29%	4/142	3%	28/142	20%	51/142	36%	142/319	45%
Simétrica	24/177	13,5%	63/177	35,5%	6/177	3%	44/177	25%	40/177	23%	177/319	55%
TOTAL	42/319	13%	104/319	33%	10/319	3%	72/319	23%	91/319	28%	319/319	100%

Fonte: própria

Apesar disso, não é possível afirmar que a produção dos RAD não possui relação alguma com contextos situacionais, pois há, ainda, outros fatores que podem ser levados em consideração em análises futuras. Além da relação entre os interlocutores, podemos investigar fatores como assunto, intenções, entre outros aspectos.

4.3.3 Ambientação

Dentre as peças selecionadas, algumas apresentam a ambientação de maneira explícita e outras revelam essa informação ao longo da trama; no entanto, há peças nas quais não há indicações a respeito da localidade geográfica. Assim sendo, as ambientações indicadas em algumas das peças são as seguintes: Interior de São Paulo; Nordeste; Rio de Janeiro; Salvador; São Paulo e Taperoá - PB. A tabela 13 exibe a distribuição das peças teatrais de acordo com as regiões ou localidades nas quais são ambientadas.

Tabela 13 – Ambientação das peças de teatro

Ambientação	Distribuição das peças	
	Nº/T	%
-		
Interior de São Paulo	1/39	2,5%
Nordeste	6/39	15%
Rio de Janeiro	11/39	28%
Salvador	1/39	2,5%
São Paulo	7/39	18%
Taperoá	3/39	8%
NA	10/39	26%
TOTAL	39/39	100%

Fonte: própria

A partir da tabela 13, observamos que a maior parte das peças é ambientada no Rio de Janeiro (11 – 28%), em São Paulo (7 – 18%) e no nordeste brasileiro (6 – 15%). Há, também, uma quantidade significativa de peças que não trazem informações a respeito da região onde a trama se passa (10 – 26%). Diante disso, na correlação da ambientação com o uso dos RAD, buscando por resultados mais consistentes, optamos por amalgamar essas regiões, utilizando os rótulos *nordeste* e *sudeste*.

Tabela 14 - Distribuição dos RAD de acordo com a ambientação das peças

Ambientação	Compreender		Entender		Perceber		Saber		Ver		TOTAL	
	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%
-												
Nordeste	0/80	0%	12/80	15%	0/80	0%	10/80	12,5%	58/80	72,5%	80/242	33%
Sudeste	22/162	13%	63/162	39%	8/162	5%	37/162	23%	32/162	20%	162/242	67%
TOTAL	22/242	9%	75/242	31%	8/242	3,3%	47/242	19,5%	90/242	37,2%	242/242	100%

Fonte: própria

Na região nordeste, destacam-se os usos do RAD *viu?* (58 – 72,5%), e, na região sudeste, sobressaem-se as ocorrências dos RAD derivados de *entender* (63 – 39%) e de *saber* (37 – 23%). A partir desses resultados, então, supõe-se que o RAD *viu?* seja um elemento mais difundido na fala nordestina, enquanto os RAD provenientes de *entender* e de *saber* podem ser mais difundidos no sudeste do país. É válido ressaltar, ainda, que não há produção de RAD derivados de *compreender* e de *perceber* na região nordeste, demonstrando que esses marcadores podem ser característicos do sudeste brasileiro. Desse modo, nossa hipótese de que algumas formas seriam mais frequentes em certas regiões se confirma nos casos de *compreender* e de *perceber*, que ocorrem apenas no sudeste; somado a isso, esperávamos que os RAD oriundos de *entender* e de *saber* fossem os menos segmentados em relação à região, ocorrendo tanto no nordeste quanto no sudeste, o que de fato ocorre, mas *viu?* também se

mostra recorrente em ambas as localidades, embora seja significativamente mais produzido em peças ambientadas no nordeste.

Em vista disso, decidimos controlar o local de nascimento dos autores para verificar se a produção de RAD está relacionada à identidade local dos dramaturgos ou se tais usos têm relação com a percepção dos autores no que tange à fala de indivíduos de determinadas regiões.

Quadro 20 – Local de nascimento dos autores

Autor(a)	Local de nascimento
Abílio Pereira de Almeida	São Paulo
Antônio Bivar	São Paulo
Ariano Suassuna	Paraíba
Augusto Pinto Boal	Rio de Janeiro
Consuelo de Castro	Minas Gerais
Dias Gomes	Bahia
Francisco Pereira da Silva	Piauí
Hilda Hilst	São Paulo
José Vicente de Paula	Minas Gerais
Leilah Assumpção	São Paulo
Mário Prata	Minas Gerais
Nelson Rodrigues	Pernambuco
Oduvaldo Vianna Filho	Rio de Janeiro
Plínio Marcos	São Paulo

Fonte: própria

Considerando o número de autores que nasceu em um determinado local, obtivemos a seguinte relação: i) São Paulo – 5 autores; ii) Minas Gerais – 3 autores; iii) Rio de Janeiro – 2 autores; iv) Bahia – 1 autor; v) Paraíba – 1 autor; vi) Pernambuco – 1 autor; e vii) Piauí – 1 autor. Na correlação dos RAD com o local de nascimento dos autores, preferimos amalgamar essas regiões, utilizando os rótulos *nordeste* e *sudeste*.

Tabela 15 - Distribuição dos RAD de acordo com o local de nascimento dos autores

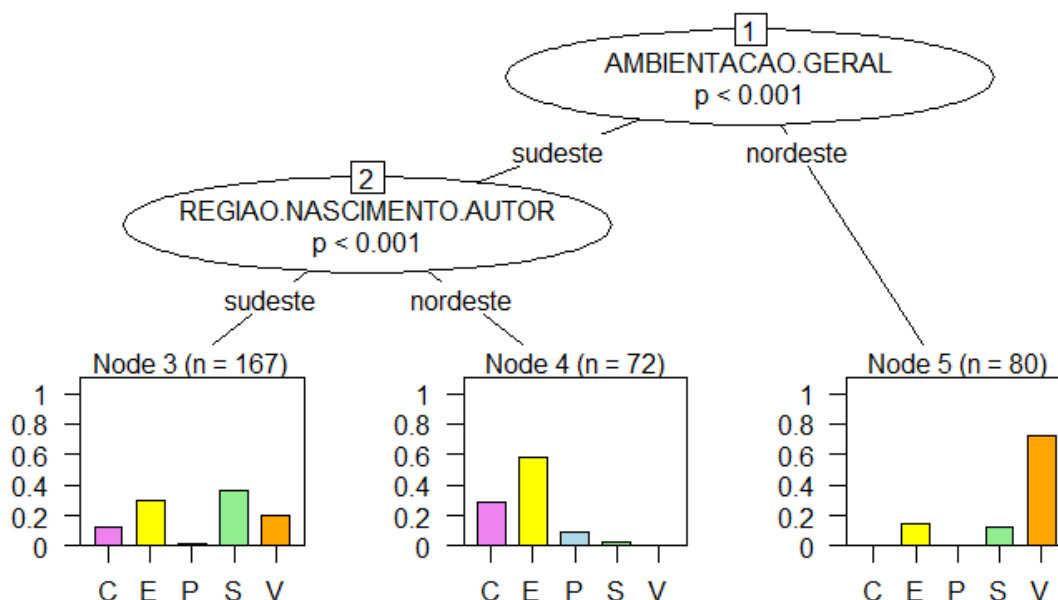
Local	Compreender		Entender		Perceber		Saber		Ver		TOTAL	
-	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%
Nordeste	21/152	14%	54/152	35%	7/152	5%	12/152	8%	58/152	38%	152/319	48%
Sudeste	21/167	12,5%	50/167	30%	3/167	2%	60/167	35,5%	33/167	20%	167/319	52%
TOTAL	42/319	13%	104/319	33%	10/319	3%	72/319	23%	91/319	28%	319/319	100%

Fonte: própria

Ao observar a tabela, notamos que os RAD derivados de *compreender*, *entender* e *perceber* são usados, em proporções semelhantes, por autores nascidos no sudeste e no nordeste. Os RAD provenientes de *ver* são mais produzidos por autores nascidos na região nordeste, e os de *saber* são usados, majoritariamente, por autores nascidos na região sudeste. Apesar de os resultados serem interessantes, percebemos que o local de nascimento dos autores não motiva, de fato, a produção dos RAD nas peças. Nelson Rodrigues, por exemplo, nasceu no nordeste do país (Recife - Pernambuco), mas mudou-se, muito cedo, para o Rio de Janeiro, região na qual grande parte de suas obras são ambientadas, retratando a sociedade carioca e seus costumes. Ariano Suassuna, por outro lado, nasceu no nordeste e é conhecido pela produção de peças que trazem à tona a cultura nordestina. Logo, é provável que a ambientação das peças seja, para os autores, um fator mais relevante do que o local de nascimento deles em si, pois é esperado que uma peça que se passa na Bahia retrate os usos linguísticos dessa região do mesmo modo que uma peça que se passa em São Paulo caracterize o linguajar dessa localidade – a não ser que as obras tratem sobre temas como migração ou casos semelhantes, situações que podem ocasionar distinções linguísticas.

Diante dessas considerações, verificamos a relação entre a ambientação das peças e o local de nascimento dos autores das obras, o que vem representado na figura 15.

Figura 15 – Diagrama arbóreo: ambientação e local de nascimento dos autores



Fonte: própria

O diagrama evidencia que a ambientação das peças é mais atuante sobre o uso dos RAD do que o local de nascimento dos autores. Nas peças ambientadas no nordeste, destaca-se o RAD *viu?* e não há dados derivados de *compreender* e de *perceber*. Em

contrapartida, as peças ambientadas no sudeste têm alguma relação com o local de nascimento dos dramaturgos e possuem dois padrões possíveis: autores nascidos no sudeste, com peças ambientadas nessa região, produzem mais os RAD derivados de *saber* e de *entender*, e autores nascidos no nordeste, cujas peças são ambientadas no sudeste do país, produzem em maior quantidade os RAD provenientes de *entender* e de *compreender*. Por meio desses resultados, é possível afirmar que a ambientação das peças é mais relevante no tocante à construção dos personagens e à representação de seus usos linguísticos do que o local de nascimento dos autores.

4.3.4 Autoria

O controle da variável autoria revela-se essencial por alguns motivos: primeiro, diferentemente do personagem criado, o autor constitui um usuário real da língua; segundo, por ser um usuário real, o autor tem suas preferências linguísticas e estilísticas; e, por fim, o modo como os sujeitos fictícios são retratados, linguística e socialmente, é construído a partir da percepção de um falante real, isto é, o autor. À vista disso, na tabela 16, apresentamos a distribuição dos RAD de acordo com a autoria das peças teatrais que integram o *corpus*.

Tabela 16 - Distribuição dos RAD de acordo com a autoria

Autor(a) ⁵¹	Compreender		Entender		Perceber		Saber		Ver		TOTAL	
	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%	Nº/T	%
-												
Abílio	1/5	20%	1/5	20%	0/5	0%	3/5	60%	0/5	0%	5/319	1,6%
Antônio	0/11	0%	4/11	36%	0/11	0%	6/11	55%	1/11	9%	11/319	3,4%
Ariano	1/83	1,2%	17/83	20,4%	0/83	0%	7/83	8,4%	58/83	70%	83/319	26%
Augusto	0/5	0%	0/5	0%	0/5	0%	5/5	100%	0/5	0%	5/319	1,6%
Consuelo	0/4	0%	3/4	75%	0/4	0%	1/4	25%	0/4	0%	4/319	1,3%
Dias	0/6	0%	3/6	50%	0/6	0%	3/6	50%	0/6	0%	6/319	1,9%
Francisco	0/5	0%	4/5	80%	0/5	0%	1/5	20%	0/5	0%	5/319	1,6%
Hilda	19/44	43%	11/44	25%	0/44	0%	14/44	32%	0/44	0%	44/319	14%
José	1/29	3%	13/29	45%	2/29	7%	13/29	45%	0/29	0%	29/319	9%
Leilah	0/9	0%	3/9	33%	0/9	0%	0/9	0%	6/9	67%	9/319	3%
Mário	0/13	0%	3/13	23%	1/13	8%	5/13	38%	4/13	31%	13/319	4%
Nelson	20/58	34%	30/58	52%	7/58	12%	1/58	2%	0/58	0%	58/319	18%
Oduvaldo	0/40	0%	5/40	12,5%	0/40	0%	13/40	32,5%	22/40	55%	40/319	12,5%
Plínio	0/7	0%	7/7	100%	0/7	0%	0/7	0%	0/7	0%	7/319	2,1%
TOTAL	42/319	13%	104/319	33%	10/319	3%	72/319	23%	91/319	28%	319/319	100%

Fonte: própria

⁵¹ Abílio Pereira de Almeida; Antônio Bivar; Ariano Suassuna; Augusto Boal; Consuelo de Castro; Dias Gomes; Francisco Pereira da Silva; Hilda Hilst; José Vicente de Paula; Leilah Assumpção; Mário Prata; Nelson Rodrigues; Oduvaldo Vianna Filho e Plínio Marcos.

Os dramaturgos destacados na tabela são os que mais utilizam RAD em suas peças: Ariano Suassuna (83 – 26%); Nelson Rodrigues (58 – 18%); Hilda Hilst (44 – 14%) e Oduvaldo Vianna Filho (40 – 12,5%). Em relação à quantidade de peças desses dramaturgos em nosso *corpus*, temos a seguinte distribuição: Ariano Suassuna: 8 peças; Nelson Rodrigues: 8 peças; Hilda Hilst: 5 peças e Oduvaldo Vianna Filho: 1 peça. À primeira vista, poderíamos afirmar que a frequência dos RAD nas peças de Ariano Suassuna e de Nelson Rodrigues é apenas um reflexo da quantidade de obras que integram o *corpus*; no entanto, ao levar em consideração o caso de Oduvaldo Vianna Filho, percebemos que, embora ele seja autor de apenas uma das peças, sua produção de RAD foi quase semelhante à de Hilda Hilst, que possui cinco peças no *corpus*.

Antes de discutirmos as preferências desses autores no que concerne ao uso dos RAD, a tabela 17 sistematiza os 123 personagens que constituem a nossa análise de acordo com a autoria.

Tabela 17 - Distribuição de personagens de acordo com a autoria

Autor(a)	Quantidade de personagens	
	Nº/T	%
-		
Abílio Pereira de Almeida	3/123	2,5%
Antônio Bivar	4/123	3,2%
Ariano Suassuna	30/123	24,4%
Augusto Pinto Boal	4/123	3,2%
Consuelo de Castro	3/123	2,5%
Dias Gomes	5/123	4%
Francisco Pereira da Silva	4/123	3,2%
Hilda Hilst	16/123	13%
José Vicente de Paula	8/123	6,5%
Leilah Assumpção	2/123	1,6%
Mário Prata	4/123	3,2%
Nelson Rodrigues	28/123	23%
Oduvaldo Vianna Filho	8/123	6,5%
Plínio Marcos	4/123	3,2%
TOTAL	123/123	100%

Fonte: própria

Por meio dessa tabela, é possível verificar que a maior parte dos personagens que usam RAD pertence às peças teatrais dos seguintes autores: Ariano Suassuna (30 – 24,4%); Nelson Rodrigues (28 – 23%); Hilda Hilst (16 – 13%); Oduvaldo Vianna Filho (8 – 6,5%) e

José Vicente de Paula (8 – 6,5%). Essa ordem corresponde à dos dramaturgos que mais utilizam RAD em suas peças; por isso, iremos focalizar as preferências de RAD dos quatro autores que mais produziram esses elementos em seus textos (cf. Tabela 16).

Nas peças de Ariano Suassuna, vemos que há mais ocorrências do RAD *viu?* (58 – 70%). A maior parte das peças do autor é representativa do nordeste brasileiro; logo, *viu?*, na percepção do escritor, pode ter associação aos usos linguísticos de falantes dessa região e, por essa razão, esse marcador foi usado com frequência nos diálogos de seus personagens, como exemplificado em (54) e (55).

(54)

PADRE ANTÔNIO - O motivo foi diferente, **viu?** Eu me mantive na linha: perdi o nome, mas não recebi nenhum outro em lugar dele. (*A pena e a lei* – Ariano Suassuna, 1959).

(55)

DODÓ - Deixe de brincadeira, **viu?** Cadê Margarida? Onde está Caroba? (*O santo e a porca* – Ariano Suassuna, 1957).

Tais resultados corroboram os da tabela 14, na seção sobre ambientação, que mostra o RAD *viu?* com maior recorrência no nordeste.

Nelson Rodrigues, em contrapartida, produziu mais os RAD derivados de *entender* (30 – 52%) e de *compreender* (20 – 34%). O autor também foi responsável pela produção de sete das dez ocorrências dos RAD advindos de *perceber*. Os exemplos em (56), (57) e (58) ilustram esses usos.

(56)

ZULMIRA - A mulher de maiô está nua. **Compreendeu?** Nua no meio da rua, nua no meio dos homens! (*A falecida* – Nelson Rodrigues, 1953).

(57)

HERCULANO - O senhor acredita que isso que aconteceu, essa monstruosidade, que isso possa alterar, **entende?** Mudar, enfim, a personalidade do meu filho? (*Toda nudez será castigada* – Nelson Rodrigues, 1965).

(58)

MÉDICO - Dobre a língua! Já lhe disse que não quero intimidades durante o serviço. Aqui me chame de doutor, **percebeu?** E vê se não me dá peso! (*Perdoa-me por me traíres* - Nelson Rodrigues, 1957).

Na tabela 14, na seção concernente à ambientação, vimos que os RAD derivados de *perceber* ocorrem mais em peças ambientadas no sudeste. Afunilando a discussão, é provável que os RAD que se originaram de *perceber* sejam comuns no linguajar carioca, uma vez que as peças de Nelson Rodrigues que compõem o *corpus* são ambientadas no Rio de Janeiro; seria necessário, porém, reunir obras de diferentes autores, ambientadas nessa região, para corroborar essa hipótese.

No caso de Hilda Hilst, os RAD mais utilizados pela autora provêm do verbo *compreender* (19 – 43%), como vemos em (59) e (60). Os RAD derivados de *compreender* parecem estar bastante apegados aos traços semânticos do verbo de origem e não são tão recorrentes no *corpus*, característica que, segundo a GTI, não se aplica aos RAD, elementos muito frequentes na fala.

(59)

ESCUDEIRO 2 - É uma coisa lógica, muito lógica, **compreendeu?** Mas por enquanto é preciso ter muita paciência. E nisso o senhor escudeiro mór é um mestre, aliás em tudo... [...]. (*O novo sistema* – Hilda Hilst, 1968).

(60)

MONSENHOR - [...] As alunas e as postulantes são muito jovens... podem fantasiar a seu respeito, podem querer partilhar desse seu... brilhantismo. Isso não é o desejo dos superiores. **Você compreende?** Todas estão perturbadas. Modificaram-se. (*A empresa ou A possessa* – Hilda Hilst, 1967).

As peças da autora que compõem o nosso *corpus* não trazem indicações referentes à localidade em que se passam, há apenas informações mais específicas, como cela, casa ou praça; por esse motivo, não é possível estabelecer correlações entre a produção dos RAD provenientes de *compreender* e a ambientação das peças de Hilda Hilst.

Quanto ao dramaturgo Oduvaldo Vianna Filho, há dois RAD que se sobressaem pela recorrência: *ver* (22 – 55%) e *saber* (13 – 32,5%). A única peça do autor no *corpus* é caracterizada por contextos de maior informalidade – dado que os personagens ou pertencem ao mesmo time de futebol ou são conhecidos de longa data – e é repleta de marcas da

conversação natural (CASTILHO, 2019), traços que podem ser verificados no diálogo reproduzido a seguir:

(61)

EUNÁPIO – Mestre Durva! (*Outro tom.*) Tá nervoso, Cafuné?

CAFUNÉ – Tô de pé porque tô acostumado a equilibrá.

EUNÁPIO – Ai, meu Maranhão! Como vai o pé?

MARANHÃO – O pé é que não vai...

EUNÁPIO – Mas cê joga, né?

BILA – Claro!

CAFUNÉ – Bem que cê num queria, né, babão?

BILA – Cê é tão bobo... Inteiro bobo.

(*Chapetuba Futebol Clube* – Oduvaldo Vianna Filho, 1959).

O uso dos RAD originados de *saber* e de *ver* demonstram que, sob a percepção do autor, esses marcadores estariam mais relacionados a situações coloquiais, menos formais e com menor monitoramento – o que fortalece as nossas hipóteses a respeito da distinção social entre *compreender/perceber* e *saber/ver*. É válido destacar que, na peça de Oduvaldo Vianna Filho, todas as ocorrências de *saber* aparecem na terceira pessoa do singular do presente do indicativo – até mesmo a forma composta *num sabe?* –, que é uma característica comum desse RAD. *Viu?*, por seu turno, é invariável em nosso *corpus*. Essas características já haviam sido identificadas por Urbano (2019) em cuja amostra *sabe?* ocorreu somente no presente do indicativo e *viu?* apenas no pretérito perfeito. Zanin e Gonçalves (2022, p. 65), nessa mesma direção, explicam que os MD não estão “sujeitos à flexão gramatical, típica da categoria verbal de que se originam. Na função de MD, as formas verbais se fixam na terceira pessoa, no modo indicativo e no tempo de presente ou de pretérito perfeito”. Vejamos os exemplos abaixo:

(62)

ZITO - Tá boa. O sorriso é desse tamanho no rostinho. Tem uma enfermeira cuidando delas **sabe?** Vai ficá até as oito, só por que eu sô de Chapetuba. (*Chapetuba Futebol Clube* – Oduvaldo Vianna Filho, 1959).

(63)

PAULINHO – [...] Eu nasci aqui, viu, Maranha? Quem tem Chapetuba no coração qué ganhá! Mas lá no campo. É. Aqui pode fazer fita... usá barba. Jogá é no campo! (*Chapetuba Futebol Clube* – Oduvaldo Vianna Filho, 1959).

Nogueira (2018), em um estudo sobre a gramaticalização do verbo *ver* no português do interior paulista – que, a propósito, é o lugar onde a peça de Oduvaldo Vianna Filho é ambientada – declara que, como MD, *ver* atua

[...] na camada do ato discursivo, cuja função não é mais retomar o sentido original de ver algo (ou perceber algo, em um sentido mais abstrato), mas sim de checar a atenção do ouvinte, em outras palavras, perde o valor lexical e passa a ter função pragmático-discursiva, servindo exclusivamente para a situação comunicativa. (NOGUEIRA, 2018, p. 434)

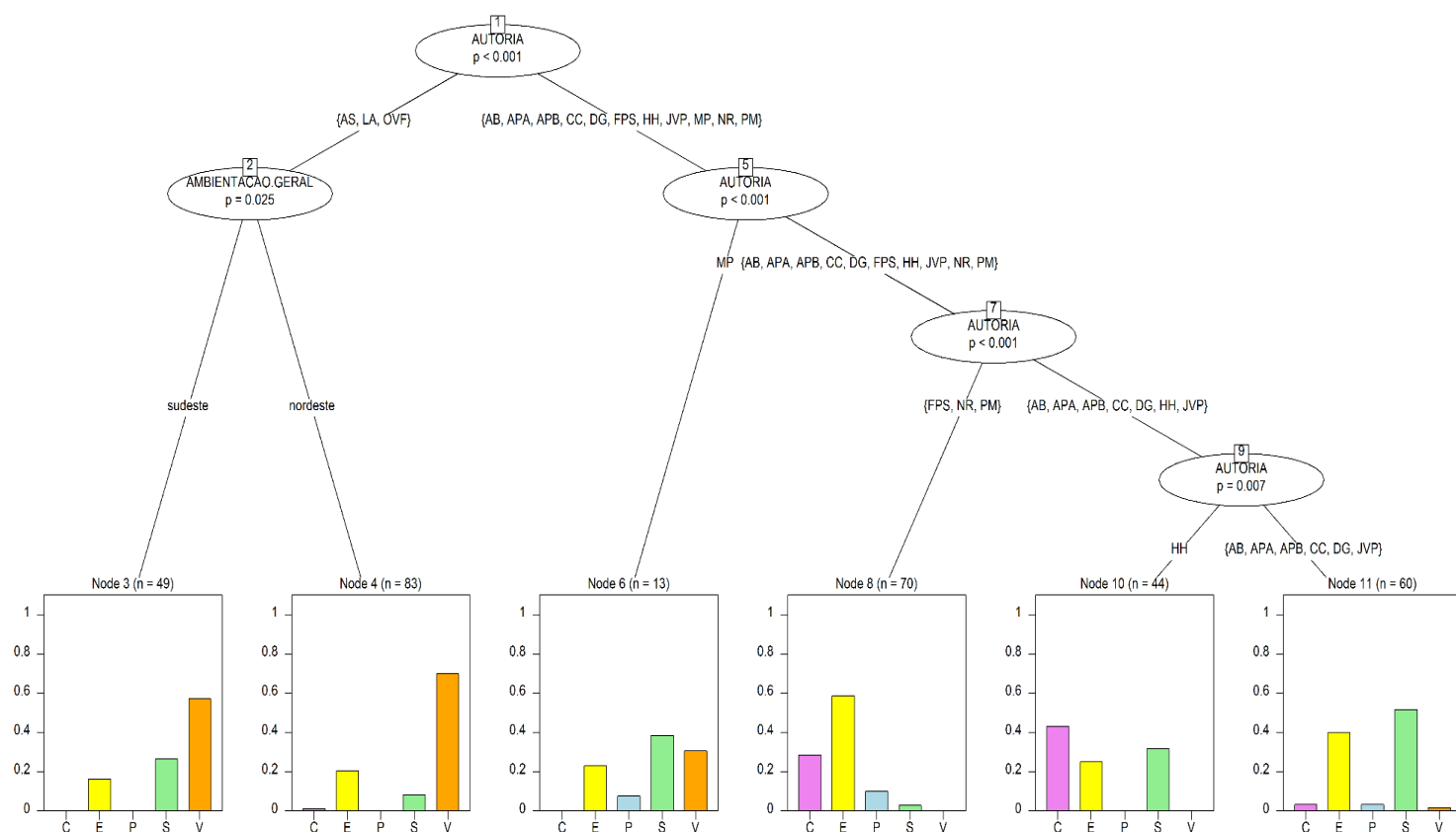
Nesse sentido, podemos afirmar que os RAD advindos de *ver*, e também de *saber*, são formas recorrentes e cristalizadas, com pouca ou nenhuma variabilidade, e possuem sentidos mais abstratos, traços comuns aos MD.

Em vista desses resultados, nossa hipótese de que alguns RAD seriam mais utilizados nas peças de determinados autores se confirma, pois os dramaturgos, como usuários reais da língua, optam por determinados RAD em detrimento de outros. Está claro, também, que fatores regionais parecem ter alguma relação com essas escolhas, pois, conforme discutimos anteriormente, uma peça ambientada no nordeste dificilmente vai apresentar traços linguísticos tipicamente paulistas ou fluminenses, a não ser que haja migração de personagens de um local para outro – nesses casos, é provável que os autores retratem as distinções linguísticas.

Adiante, fazemos alguns cruzamentos que podem corroborar tais hipóteses.

4.3.5 Cruzamentos

O cruzamento entre as variáveis ambientação e autoria propiciou o seguinte resultado:

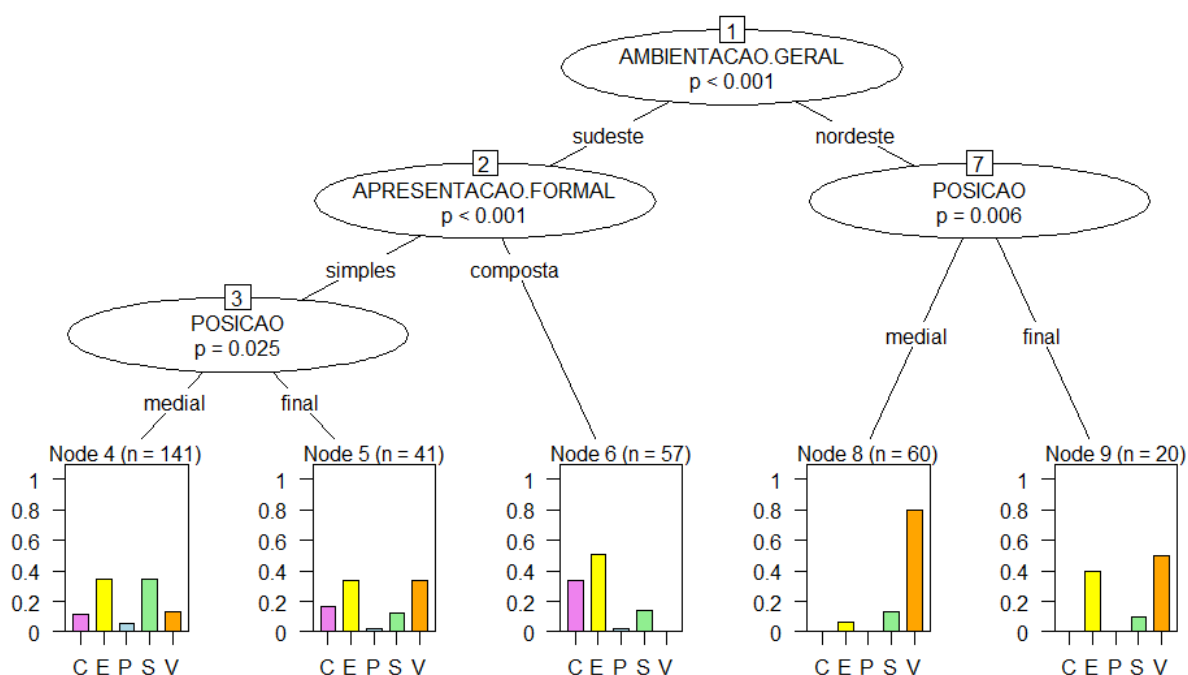
Figura 16 - Diagrama arbóreo: ambientação e autoria⁵²

Fonte: própria

Não obstante haja diversos subgrupos a serem explorados nesse diagrama, neste momento, focalizaremos apenas os autores diretamente atrelados à ambientação. Dessa maneira, a autoria revela-se a variável mais atuante, e a ambientação está diretamente relacionada à produção de três autores: Ariano Suassuna (AS), Leilah Assumpção (LA) e Oduvaldo Vianna Filho (OVf). Tanto no nordeste quanto no sudeste, sobressaem-se RAD derivados de *ver*, *entender* e *saber*, com um uso proporcionalmente maior de *ver* e *saber* nas peças ambientadas no sudeste e de *ver* e *entender* nas peças ambientadas no nordeste.

Ao cruzar a variável ambientação às variáveis de natureza linguística, obtemos o seguinte resultado:

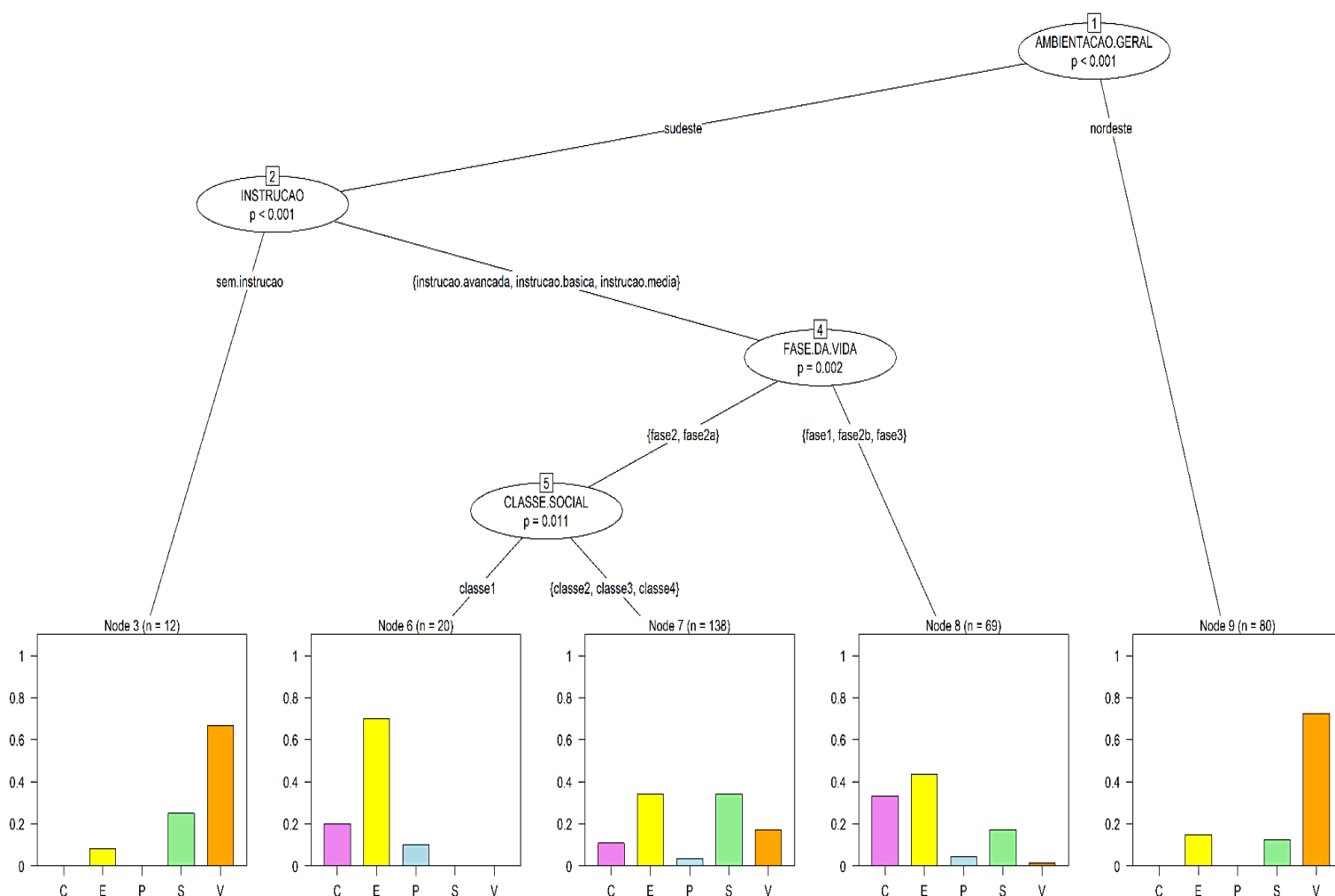
⁵² APA = Abílio Pereira de Almeida; AB = Antônio Bivar; AS = Ariano Suassuna; APB = Augusto Pinto Boal; CC = Consuelo de Castro; DG = Dias Gomes; FPS = Francisco Pereira da Silva; HH = Hilda Hilst; JVP = José Vicente de Paula; LA = Leilah Assumpção; MP = Mário Prata; NR = Nelson Rodrigues; OVf = Oduvaldo Vianna Filho e PM = Plínio Marcos.

Figura 17 - Diagrama arbóreo: ambientação e variáveis de natureza linguística

Fonte: própria

Nesse caso, a ambientação mostrou-se mais atuante sobre o fenômeno. Peças ambientadas no nordeste estão diretamente relacionadas à posição dos RAD: tanto em posição medial quanto em posição final, a produção de *viu?* é mais recorrente; ademais, na posição final, o uso de RAD derivados de *entender* também é frequente. Por outro lado, as peças ambientadas no sudeste mantêm relação com a apresentação formal dos RAD, que podem ser simples ou compostos. Os RAD derivados de *entender* e de *compreender* são os que mais apresentam formas compostas. Os RAD cuja forma é simples, por sua vez, estão relacionados à posição no turno: em posição medial, destacam-se *entender* e *saber* e, em posição final, sobressaem-se os RAD advindos de *entender* e de *ver*.

Correlacionamos, também, a variável ambientação ao perfil social dos personagens.

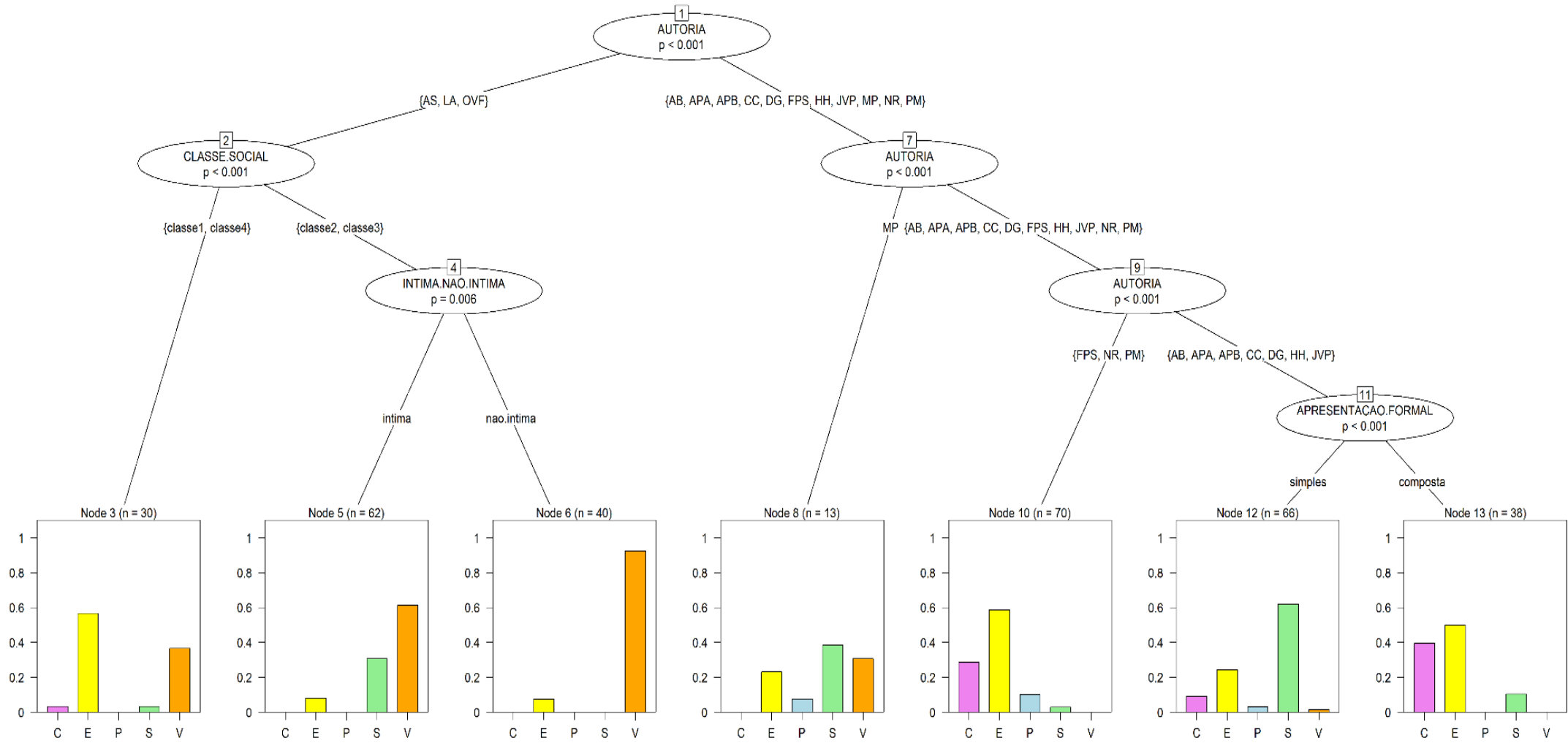
Figura 18 - Diagrama arbóreo: ambientação e perfil social dos personagens**Fonte:** própria

O gráfico demonstra que a ambientação é mais atuante do que as demais variáveis incluídas no modelo. As peças ambientadas no sudeste se diferenciam, primeiramente, segundo o grau de instrução dos personagens: os personagens sem instrução, conforme vimos previamente, não fazem uso de RAD provenientes de *compreender* e de *perceber*, e produzem mais os RAD derivados de *ver*. Personagens com instruções básica, média e avançada se distinguem em função da fase da vida. Aqueles que pertencem às fases 1, 2b e 3 produzem mais os RAD oriundos de *entender* e de *compreender*. Os personagens pertencentes às fases 2 e 2a, por outro lado, mantêm uma íntima relação com a classe social: personagens de classe 1 utilizam mais os RAD derivados de *entender* e concentram os RAD provenientes de *perceber*, enquanto os personagens das demais classes (2, 3 e 4) produzem, em proporções semelhantes, RAD provenientes de *entender* e de *saber* e apresentam um leque mais amplo de variantes.

Antes de partir para o próximo cruzamento, vale destacar a estabilidade da distribuição dos RAD provenientes de *entender*, *saber* e *ver* em peças ambientadas no nordeste (cf. Figuras 16, 17 e 18); algo que se mantém independentemente da interação com as demais variáveis.

Quando cruzamos todas as variáveis independentes, linguísticas e extralinguísticas, percebemos que a autoria é a variável mais atuante sobre os RAD em estudo (cf. Figura 19). Os RAD produzidos por Ariano Suassuna (AS), Leilah Assumpção (LA) e Oduvaldo Vianna Filho (OVF) têm relação com a classe social dos personagens. Nas peças desses autores, personagens de classes 1 e 4 produzem mais os RAD derivados de *entender* e de *ver*, enquanto personagens de classes 2 e 3, além de produzirem mais os RAD derivados de *ver*, apresentam ligação com as relações de maior ou menor intimidade: em contextos mais íntimos, a produção de RAD provenientes de *ver* e de *saber* é bastante significativa e, em contextos menos íntimos, não há produção de RAD derivados de *saber*, mas observa-se uma alta recorrência de *viu?*. Além disso, há, em ambos os contextos, pouquíssimas ocorrências de RAD originados de *entender* e não há dados de *compreender* e de *perceber*.

Em relação aos demais autores, Mário Prata (MP) destaca-se pelo uso de RAD oriundos de *saber* e de *ver*. Mais uma vez, o diagrama se ramifica, revelando que Francisco Pereira da Silva (FPS), Nelson Rodrigues (NR) e Plínio Marcos (PM) produzem mais os RAD originados de *entender*. Os dados produzidos pelos demais autores, a saber: Antônio Bivar (AB); Abílio Pereira de Almeida (APA); Augusto Pinto Boal (APB), Consuelo de Castro (CC), Dias Gomes (DG), Hilda Hilst (HH) e José Vicente de Paula (JVP), estão intimamente relacionados à apresentação formal dos RAD; dentre os RAD com forma simples, destacam-se os derivados de *saber* e, dentre os RAD com forma composta, sobressaem-se os originados de *entender* e de *compreender*.

Figura 19 - Diagrama arbóreo: correlação entre os RAD e as variáveis independentes

Fonte: própria

O fato de a autoria ter sido considerada a variável mais atuante sobre o fenômeno demonstra que, no tocante a variáveis discursivas, as escolhas dos autores são mais significativas do que o perfil social dos personagens, mesmo que esse perfil seja crucial para a construção dos “modos de falar” desses personagens. De acordo com Ubersfeld (2005[1996], p. 163), o discurso teatral é desenvolvido da seguinte forma: “X (autor) ordena a Y (ator) que diga (enunciado)”, ou seja, o autor determina o que será dito pelo personagem e como ele deverá dizê-lo. Entretanto, adiante, a autora explica que,

Na medida em que o discurso teatral é discurso de um sujeito-*scriptor*, ele é discurso de um sujeito imediatamente destituído de seu *Eu*, de um sujeito que se nega como tal, que se afirma como quem fala pela voz de um outro, de muitos outros, como quem fala sem ser sujeito: o discurso teatral é *discurso sem sujeito*. A função do *scriptor* é organizar as condições de emissão de uma fala pela qual, ao mesmo tempo, nega ser responsável. (UBERSFELD, 2005[1996], p. 168, grifos da autora).

Os dramaturgos, então, assumem diferentes vozes, dando vida aos personagens criados. Algumas vezes, o discurso teatral é usado para atribuir determinados estatutos sociais aos personagens, como nos casos em que um criado distingue-se linguisticamente de seus padrões. Nas palavras de Ubersfeld (2005[1996], p. 173, grifo da autora), “a diferença linguística nunca é considerada no teatro uma diferença *específica*, mas a designação daquele que está fora do grupo, em posição de inferioridade” ou, complementando a declaração da autora, de superioridade. De fato, há, em algumas peças, diferenciação da fala de personagens de classes sociais mais baixas ou de outras etnias, que não seja a branca. Além disso, a autora esclarece que,

[...] a rigor considera-se que toda personagem de teatro fala, antes de tudo, a linguagem da classe social a que pertence. Tal “evidência” precisa ser atenuada: é acaso a camada social a que se supõe pertencer ou aquela à qual pertence o *scriptor* [...]? Não seria antes a linguagem *imaginária* que o *scriptor* empresta à camada social à qual pertence sua personagem? (UBERSFELD, 2005[1996], p. 174, grifos da autora).

Em vista disso, percebe-se que as macrocategorias sociais – não apenas classe social, mas também gênero/sexo, fase da vida, instrução e etnia – dependem da vontade do autor, isto é, cabe ao dramaturgo decidir se irá diferenciar os usos linguísticos de seus personagens com base nessas categorias.

Por fim, neste trabalho, podemos concluir que a produção de RAD está mais associada à variável autoria do que às macrocategorias sociais. Esse resultado condiz, em certa medida, com os estudos de terceira onda da sociolinguística (ECKERT, 2012; OLIVEIRA;

ROCKENBACH; GUTIERRES, 2022), nos quais a variação não é tomada como o reflexo de macrocategorias sociais, mas é vista como “um sistema complexo de significados sociais em potencial dentro de uma comunidade, que pode ser manipulado localmente pelos indivíduos na construção de estilos identitários” (COELHO *et al.*, 2019, p. 69). Diante disso, o estilo diz respeito à construção da *persona*, e a identidade de um indivíduo não se reduz a categorias como escolaridade, sexo, idade etc., mas é dinâmica, mutável, plural e construída por meio de relações/práticas sociais. Logo, os usos linguísticos de pessoas pertencentes às mesmas macrocategorias sociais podem ser muito distintos (OUSHIRO, 2019).

Nessa direção, Valle e Görski (2019), considerando a relação dos RAD com as categorias sociais, com a identidade local e com a construção da *persona*, pontuam que: i) os RAD não são sensíveis às categorias sociais, pois o funcionamento estilístico desses elementos não é determinado por esses aspectos, que não seriam suficientes para descrever o uso dos MD; ii) os RAD carregam marcas de identidade local e de significado ideológico; e iii) os RAD têm um papel crucial na construção de *personas*. Para sustentar essa última afirmação, Valle e Görski (2019, p. 1255) elucidam que “[...] alguns RADs têm o potencial de caracterizar identidades e/ou indivíduos, sendo usados na construção de personagens em algumas novelas e também para compor a caracterização de alguns humoristas”. De maneira geral, as autoras sugerem que a escolha desses elementos é motivada por fatores estilísticos, identitários e por preferências particulares.

Sendo assim, nas peças teatrais, as macrocategorias sociais são fulcrais para a construção dos personagens, os quais têm características sociais diversas, porém, em concordância com Valle e Görski (2019), a produção dos RAD em estudo não é imediatamente motivada por essas categorias, mas sim pelas preferências particulares dos usuários da língua, neste caso, os autores. Silva e Macedo (1996), diante do fato de que fenômenos de natureza discursiva são menos influenciados por variáveis sociais, sugerem que, possivelmente, isso se dá porque as pressões normativas impostas pela escola não teriam alcançado, ainda, o nível discursivo e, por isso, essas formas seriam menos sistematizadas. Freitag (2007), por sua vez, em uma discussão mais recente, afirma que os MD são estigmatizados pela tradição gramatical; segundo ela, “os marcadores discursivos não são reconhecidos como uma categoria nas gramáticas normativas, assim como verbo, pronomes e conjunção, por exemplo” (FREITAG, 2007, p. 28). Consequentemente, esses elementos não constam nos currículos e nos materiais escolares. Dessa maneira, as afirmações das autoras dialogam, pois, se os MD não são tratados como categoria gramatical e tampouco constam

nos materiais usados nas escolas, a hipótese de Silva e Macedo (1996) se confirma: o discurso não foi sistematizado.

4.4 Sintetizando...

Neste capítulo, desenvolvemos a análise do fenômeno e avaliamos o comportamento das variáveis independentes em relação aos RAD em estudo. Inicialmente, apresentamos uma visão geral dos resultados para, na sequência, discutirmos cada uma das variáveis controladas e sua correlação com o fenômeno sob análise. Além disso, realizamos alguns cruzamentos com a finalidade de aprofundar e de refinar a discussão dos resultados.

Ao controlarmos as variáveis de natureza linguística, verificamos que os RAD analisados, em sua maioria, têm formas simples, atuam em posição medial de turno e são caracterizados pela ausência de respostas após seu uso.

Em relação às variáveis de natureza extralinguística, observamos que as macrocategorias sociais (gênero/sexo, fase da vida, instrução, classe social e etnia) não são as variáveis que mais motivam a produção dos RAD; ainda assim, essas categorias não deixam de ser importantes para a construção do perfil social dos personagens. Ademais, a produção de RAD no *corpus* analisado não foi influenciada pela relação entre os personagens/interlocutores, conforme evidenciado pelos testes de qui-quadrado (cf. Quadro 14). A variável ambientação, por seu turno, demonstrou que o uso dos RAD pode ser motivado por fatores regionais, uma vez que os autores buscam representar usuários reais da língua; verificamos que *compreender* e *perceber* são usados apenas em peças ambientadas no sudeste, enquanto *entender*, *saber* e *ver* são utilizados em ambas as localidades, nordeste e sudeste. Ao realizarmos o cruzamento de todas as variáveis, linguísticas e extralinguísticas, constatamos que a variável mais significativa é a autoria, demonstrando que a percepção dos autores é um fator relevante no tocante à produção dos RAD no *corpus* investigado.

5 CONCLUSÃO

O objetivo geral deste trabalho foi investigar o uso variável dos RAD derivados dos verbos de cognição *compreender*, *entender*, *perceber*, *saber* e *ver*. Para alcançar esse propósito geral, verificamos quais forças linguísticas e sociais poderiam estar correlacionadas ao uso dos RAD em estudo e discurremos sobre o processo de gramaticalização desses elementos, buscando entender de que modo essas formas passaram de verbos plenos a RAD, tornando-se intercambiáveis em um mesmo domínio funcional.

Articulando os pressupostos teórico-metodológicos da TVML, do Sociofuncionalismo e da Sociolinguística Histórica, partimos da concepção de língua como sistema heterogêneo, passível de variação e mudança linguísticas – as quais podem ser identificadas em registros escritos, como as peças teatrais. Relembramos, também, os principais eventos que marcaram a história do teatro brasileiro no século XX e defendemos o uso das peças teatrais como *corpus* para os estudos (socio)linguísticos. Na sequência, voltamos nossa atenção para os MD, que são elementos típicos da fala, caracterizando, assim, a espontaneidade e a dinamicidade da língua em uso. Somado a isso, sintetizamos as abordagens de Schiffrin (1987, 2001, 2006) e de Risso, Silva e Urbano (2019), as quais fundamentam este trabalho, discurremos a respeito dos RAD, pontuando as principais características desse subconjunto, e discutimos o processo de gramaticalização desses elementos. Dando continuidade à pesquisa, detalhamos o caminho percorrido até a obtenção dos resultados: descrevemos o *corpus* – constituído por 39 peças teatrais brasileiras produzidas entre as décadas de 1950 e 1970 –, explicamos como foram realizados a coleta e o tratamento dos dados e especificamos cada uma das variáveis independentes e as hipóteses a respeito delas.

Como mencionado na seção dedicada à discussão dos pressupostos teórico-metodológicos da Sociolinguística Variacionista, dos cinco problemas empíricos postulados por Weinreich, Labov e Herzog (2006[1968]), o presente trabalho lidou com o *problema da restrição*, porque avaliamos os determinantes linguísticos e extralinguísticos envolvidos na escolha dos RAD em estudo, e com o *problema do encaixamento*, dado que nos propusemos a explicar de que modo os RAD derivados dos verbos de cognição *compreender*, *entender*, *perceber*, *saber* e *ver* estão inseridos na matriz social do período histórico em investigação (1950-1970). A análise e a discussão dos resultados demonstraram que: i) é possível dispensar um tratamento variável aos RAD, posto que atuam no mesmo domínio funcional e possuem naturezas e sentidos semelhantes; e ii) há aspectos linguísticos e sociais envolvidos na produção desses elementos.

Constatamos **319 ocorrências** dos RAD em estudo no *corpus*, distribuídas da seguinte forma: *entender* (104 – 33%), *ver* (91 - 28%), *saber* (72 - 23%), *compreender* (42 – 13%) e *perceber* (10 – 3%).

Na discussão concernente à variável *apresentação formal*, observamos que, com exceção de *ver*, os RAD em estudo apresentam variabilidade de forma. Contudo, verificamos que, mesmo havendo variação, os RAD mais recorrentes no *corpus* são formas simples e fixas, que, possivelmente, estão mais gramaticalizadas, porque manifestam traços identificadores dos MD prototípicos (cf. Tabela 1).

Quanto à *posição*, os RAD ocorrem mais em posição medial, operando no desenvolvimento e na sustentação do turno, bem como no gerenciamento da interação. Considerando que a apresentação formal dos RAD poderia ter alguma relação com a posição ocupada por esses elementos no turno conversacional, conjecturamos que formas compostas, por terem mais material linguístico, ocorreriam, preferencialmente, em posição final. Ao contrário do que havíamos previsto, as formas compostas, assim como as simples, surgem mais em posição medial de turno, porém notamos que, de fato, há uma certa tendência de as formas compostas ocuparem a posição final.

A variável *presença/ausência de respostas* demonstrou que os RAD em estudo, em sua maioria, não são acompanhados de manifestações verbais por parte do ouvinte, o que condiz com os resultados de estudos anteriores (URBANO, 1999a; VALLE, 2001, 2014; TRAPP, 2014). Esse resultado, portanto, reforça o caráter dos RAD como perguntas retóricas, para as quais não se espera uma resposta, como também sinaliza a perda da ilocução interrogativa desses elementos (FREITAG, 2009).

Resumidamente, os RAD possuem formas mais simples e reduzidas, atuam, geralmente, em posição medial de turno e não costumam ser sucedidos por respostas. Levando em consideração o núcleo-piloto de MD proposto pela GTI, concluímos que, de maneira geral, as características dos RAD em estudo condizem com os traços mais estáveis dos MD, revelando que as formas sob análise já não atuam como verbos plenos, mas encontram-se em diferentes estágios do processo de gramaticalização. Com base nos princípios de Hopper (1991), exibidos na seção referente à gramaticalização dos RAD, constatamos que: i) os RAD advindos de *compreender*, *entender*, *perceber*, *saber* e *ver* variam e coexistem em um mesmo domínio funcional, o de requisição de apoio discursivo, e são formas intercambiáveis, que podem ser substituídas umas pelas outras sem perda de sentido proposicional (**princípio da estratificação**); ii) em certos contextos, é possível que a variação deixe de ocorrer, propiciando a fixação de uma das formas em coexistência – a exemplo disso,

podemos mencionar as formas apresentadas na tabela 1, que foram as mais frequentes no *corpus*, ou seja, mesmo havendo outras variantes, há uma rotinização de determinadas formas, que acabam sendo mais usadas pelos falantes (**princípio da especialização**); iii) os verbos dos quais os RAD em estudo se originam permanecem em uso no sistema linguístico, desempenhando outras funções, conforme demonstramos ao longo do trabalho (**princípio da divergência**); iv) em algumas formas, principalmente naquelas que aparentam estar em estágios incipientes ou intermediários de gramaticalização, há manutenção de certos traços semânticos dos verbos de origem (**princípio da persistência**); e v) à medida que se gramaticalizam, os RAD perdem as características de verbos plenos e assumem propriedades particulares aos MD, atuando exclusivamente na organização textual-interativa dos textos falados (**princípio da descategorização**).

As macrocategorias sociais (*gênero/sexo, fase da vida, instrução, classe social e etnia*), apesar de não influenciarem diretamente a produção dos RAD, são essenciais para a construção do perfil social dos personagens e para a elaboração de seus “modos de falar”. Dessa forma, consideramos que, no processo de criação dos personagens e dos diálogos, os autores são norteados por essas categorias, que, em sua quase totalidade, se mostraram significativas pelos testes de qui-quadrado (cf. Quadro 14). Ainda em relação a essas variáveis, ao analisarmos as tabelas verticalmente, detectamos uma possível distinção social entre *compreender/perceber* e *saber/ver*. Identificamos, mediante as proporções apresentadas nas tabelas, a seguinte tendência:

Compreender e Perceber

- **Instrução:** básica ou avançada;
- **Classe social:** classes 1 ou 2;
- **Etnia:** brancos.

Saber e Ver

- **Instrução:** média ou sem instrução;
- **Classe social:** classes 3 ou 4;
- **Etnia:** não brancos.

Os RAD derivados de *entender*, por seu turno, aparentam ser os menos influenciados por esses aspectos sociais, pois são utilizados nas falas de personagens com diferentes perfis sociais. Esses resultados, embora sejam interessantes, requerem análises mais profundas e um

número maior de dados, de modo que possamos realizar afirmações precisas e consistentes. Estas, então, são atividades para possíveis estudos futuros.

Quanto à *relação entre os personagens/interlocutores*, ainda que as relações simétricas e íntimas, nas quais esperávamos maior produção de RAD, tenham sido mais frequentes do que as relações assimétricas e não íntimas, essa variável não foi considerada significativa pelos testes de qui-quadrado (cf. Quadro 14), indicando, que, nas peças teatrais analisadas, a produção de RAD não é motivada por esses fatores.

Por meio da variável *ambientação*, observamos que alguns RAD são utilizados somente em uma das regiões consideradas (nordeste ou sudeste), como é caso de *compreender* e de *perceber*, que foram produzidos apenas em peças ambientadas no sudeste, e outros, apesar da predominância em uma das regiões, são usados em ambas as localidades, tais como *entender*, *saber* e *ver*.

Finalmente, a *autoria* revelou ser a variável mais significativa em nossa análise. Constatamos que os dramaturgos fazem uso de RAD específicos, que estão relacionados à ambientação das tramas e, menos acentuadamente, ao perfil social dos personagens, variáveis que se mostraram relevantes para a construção e representação dos personagens. Diante disso, depreendemos que a percepção dos dramaturgos, usuários reais da língua, é imbuída de traços sociais, históricos, culturais e, também, linguísticos, pois, para construir seus personagens, os autores precisam definir quais características intelectuais, físicas, sociais, regionais e linguísticas essas figuras irão assumir. Nessa direção, nossos resultados condizem, em certa medida, com os de estudos anteriores (SILVA; MACEDO, 1996; VALLE, 2001, 2014), os quais já haviam constatado que a escolha e o uso dos RAD estão mais relacionados a fatores contextuais e atitudinais do que às macrocategorias sociais; nas peças de teatro, tais fatores conduzem os autores na criação dos personagens e dos diálogos.

Futuramente, é possível realizar um estudo diacrônico, utilizando materiais escritos como fonte de análise, a fim de investigar mais a fundo a trajetória de mudança dos RAD derivados de verbos de cognição. Além das peças teatrais, pode-se verificar outros gêneros, como roteiros de cinema, que são similares às peças por se estruturarem a partir do diálogo, constituindo, assim, materiais produtivos para a investigação de fenômenos linguísticos. É válido, também, considerar, além de forças linguísticas e sociais, forças estilísticas e identitárias que possam motivar a produção dos RAD.

Este trabalho, então, integra o conjunto de estudos que busca responder questões concernentes aos MD – e aos RAD, mais especificamente –, grupo extremamente amorfo e heterogêneo, que, a cada investigação, ocasiona o surgimento de novas ramificações e de

caminhos diversos a serem percorridos. Lidar com variáveis discursivas não é uma tarefa simples, ainda mais quando essa atividade envolve a análise de materiais escritos; todavia, o trabalho com os MD, e, em especial, com os RAD, não se esgota facilmente, antes, quanto mais exploramos esse conjunto, mais fascinante ele se torna. Silva e Macedo (1996, p. 14), comparando os estudos dos MD a uma floresta repleta de árvores distintas e de difícil penetração, declaram que “há que se fazer cortes, traçar caminhos e avançar com cautela, voltar sobre os passos e contornar obstáculos” até que sejamos capazes de “chegar a uma clareira de onde se possa ter uma melhor visão”. Esperamos, portanto, que esta pesquisa contribua para a ampliação da visão desse fenômeno.

REFERÊNCIAS

- ALKMIM, T. M. Sociolinguística (parte 1). In: MUSSALIM, F.; BENTES, A. C. (orgs.). **Introdução à linguística: domínios e fronteiras** - vol. 1. São Paulo: Cortez, 2001. p. 21-47.
- ALMEIDA, M. A.; PINTO, L. G. O uso das peças teatrais na análise sócio-histórica do português brasileiro. **Revista da ABRALIN**, v. 19, n. 2, p. 1-5, 19 out. 2020.
- BELL, A. Language Style as Audience Design. **Cambridge University Press**. v. 13, n. 2, p. 145-204, jun. 1984.
- BELTRÃO, K. I.; NOVELLINO, M. S. **Alfabetização por raça e sexo no Brasil: evolução no período 1940-2000**. Rio de Janeiro: Ence/IBGE, 2002 (Texto para Discussão, 1). Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/livros/liv1425.pdf>. Acesso em: 16 jan. 2023.
- BELTRÃO, K. I. **Raça e fronteiras sociais: lendo nas entrelinhas do centenário hiato de raças no Brasil**. In: SOARES, S. *et al.* (orgs.). **Os mecanismos de discriminação racial nas escolas brasileiras**. Rio de Janeiro: Ipea, 2005. p. 41-87.
- BENTES, A. C.; FERREIRA-SILVA, B.; MARIANO, R. D. Atenuação e impolidez como estratégias estilísticas em contexto de entrevista televisiva. **Cadernos de Letras da UFF**, v. 23, n. 47, p. 285-314, 30 dez. 2013.
- BERGS, A. The Uniformitarian Principle and the Risk of Anachronisms in Language and Social History. In: HERNÁNDEZ-CAMPOY, J. M.; CONDE-SILVESTRE, J.C. (eds.). **The Handbook of Historical Sociolinguistics**. 1st ed. Malden, MA/Oxford: Wiley-Blackwell, 2012. p. 80-98.
- BERLINCK, R. A.; BARBOSA, J. B.; MARINE, T. C. Reflexões teórico-metodológicas sobre fontes para o estudo histórico da língua. **Revista da ABRALIN**, v. 7, p. 53-79, 2008.
- BERLINCK, R. A. **Por que (ainda) usar peças de teatro em estudos sócio-históricos da língua**. Palestra na Mesa-redonda "Sociolinguística Histórica no Brasil: Caminhos e desafios. Abralín Ao Vivo, 18/07/2020. Disponível em: <http://aovivo.abralin.org/lives/sociolinguistica-historica-no-brasil>.
- BERLINCK, R. A.; BRANDÃO, S. M. Por uma Sociolinguística Histórica: análise multidimensional de cartas pessoais e peças teatrais brasileiras. In: BIAZOLLI, C. C.; BERLINCK, R. A. (orgs.). **Gêneros textuais-discursivos no estudo de processos de variação e mudança**. Campinas, SP: Pontes Editores, 2021. p. 233-270.
- BLAKEMORE, D. **Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers**. Cambridge: Cambridge University Press, 2002.
- BORTONI-RICARDO, S. M. **Manual de Sociolinguística**. São Paulo: Contexto, 2014.
- BRANDÃO, S. M. **Alternância verbal em construções condicionais: um fenômeno variável?** 2018. 145 f. Dissertação (Mestrado em Linguística e Língua Portuguesa) - Universidade Estadual Paulista, Faculdade de Ciências e Letras, Araraquara, 2018.
- BRANDÃO, S. M. **Mudança do quadro verbal brasileiro em sentenças condicionais: contribuições à Sociolinguística Histórica**. 2022. 323 p. Tese (Doutorado em Linguística e

Língua Portuguesa) - Universidade Estadual Paulista, Faculdade de Ciências e Letras, Araraquara, 2022.

BROWN, R.; GILMAN, A. The pronouns of power and solidarity. In: SEBEOK, T. A. (ed.). **Style in Language**. MIT Press, 1960. p. 252-281.

BUSATTO, C. Um olhar sobre o feminino na dramaturgia brasileira do início do século XX. **Revista Letras**, Curitiba, n. 60, p. 223-245, jul./dez. 2003.

BYBEE, J. **Language, Usage and Cognition**. Cambridge: Cambridge University Press, 2010.

BYBEE, J. **Língua, uso e cognição**. Tradução: Maria Angélica Furtado da Cunha; Sebastião Carlos Leite Gonçalves. São Paulo: Cortez, 2016. Título original: *Language, usage and cognition*.

CAMACHO, R. G. Sociolinguística (parte 2). In: MUSSALIM, F.; BENTES, A. C. (orgs.). **Introdução à linguística: domínios e fronteiras** - vol. 1. São Paulo: Cortez, 2001. p. 49-75.

CASTILHO, A. T. de. **Nova gramática do português brasileiro**. São Paulo: Contexto, 2019.

CEBULSKI, M. C. **Introdução à História do Teatro no Ocidente: dos gregos aos nossos dias**. Guarapuava: Editora Unicentro, 2013.

COELHO, I. L.; NUNES DE SOUZA, C. M. Uma proposta metodológica para o tratamento da variação estilística em textos escritos. In: GÖRSKI, E. M.; COELHO, I. L.; NUNES DE SOUZA, C. M. (orgs.). **Variação Estilística: reflexões teórico-metodológicas e propostas de análise**. Florianópolis: Insular, 2014. p. 163-199.

COELHO, I. L. *et al.* **Para conhecer Sociolinguística**. São Paulo: Contexto, 2019.

CONDE SILVESTRE, J. C. **Sociolinguística histórica**. Madrid: Gredos, 2007.

COSTA, S. P. T. da S.; FURTADO DA CUNHA, M. A. A construção com verbos de cognição no português brasileiro: um estudo preliminar. **Gragoatá**, Niterói, v. 21, n. 40, p. 250-272, 2016.

CUNHA, A. F. Funcionalismo. In: MARTELOTTA, M. E. (org.) **Manual de Linguística**, 2. ed. São Paulo: Contexto, 2011. p. 157-176.

ECKERT, P. Three waves of variation study: The emergence of meaning in the study of Sociolinguistic Variation. **Annual Review of Anthropology**, v. 41, p. 87-100, 2012.

FARACO, C. A. **Linguística Histórica: uma introdução ao estudo da história das línguas**. São Paulo: Parábola Editorial, 2006.

FISCHER, K. (ed.). **Approaches to Discourse Particles**. Amsterdam: Elsevier, 2006.

FRASER, B. Pragmatic markers. **Pragmatics**, v. 6, n. 2, p. 167-190, 1996.

FRASER, B. What are discourse markers? **Journal of Pragmatics**, London, v. 31, n. 7, p. 931-952, July 1999.

FRASER, B. Towards a theory of discourse markers. In: FISCHER, K. (ed.). **Approaches to Discourse Particles**. Amsterdam: Elsevier, 2006. p. 189-204.

- FRASER, B. An account of Discourse Markers. **International Review of Pragmatics**, v. 1, n. 2, p. 293-320, 2009.
- FREITAG, R. M. K. Marcadores discursivos não são vícios de linguagem! **Interdisciplinar**, v. 4, n. 4, p. 22-43, 2007.
- FREITAG, R. M. K. Marcadores discursivos interacionais: análise contrastiva entre duas variedades do português falado no Brasil. In: **Anais do Círculo de Estudos Linguísticos do Sul - CELSUL**, 2008.
- FREITAG, R. M. K. Estratégias gramaticalizadas de interação na fala e na escrita: marcadores discursivos revisitados. **ReVEL**, v. 7, n. 13, p. 1-15, 2009.
- FREITAG, R. M. K.; SEVERO, C. G. **Mulheres, linguagem e poder**: estudos de gênero na Sociolinguística brasileira. São Paulo: Edgard Blucher Ltda., 2015.
- FREITAG, R. M. K.; SILVA, R. B. da; EVANGELISTA, F. R. S. Marcadores discursivos interacionais: diferentes metodologias, diferentes resultados. **Diacrítica**, 31(1), p. 55-75, 2017.
- GALEMBECK, P. T. O turno conversacional. In: PRETI, D. (org.). **Análise de textos orais**. 4. ed. São Paulo: Humanitas, 1999. p. 55-80.
- GOMES, A. C. População e sociedade. In: SCHWARCZ, L. M. (dir.). **História do Brasil nação**: 1808-2010 - Vol. 4: GOMES, A. C. (coord.). Olhando para dentro: 1930-1964. Rio de Janeiro: Objetiva, 2013. p. 41-90.
- GONÇALVES, S. C. L. Ainda em favor de uma interface entre Sociolinguística e Gramaticalização. **Work. Pap. Linguíst.**, Florianópolis, v. 22, n. esp., p. 111-136, 2021.
- GÖRSKI, E. M. *et al.* Gramaticalização/discursivização de itens de base verbal: funções e formas concorrentes. **Estudos Linguísticos**, São Paulo, v. 31, n. 1, [s.p.], 2002.
- GÖRSKI, E. M.; VALLE, C. R. M. Variação discursiva: procedimentos metodológicos para delimitação do envelope de variação. In: FREITAG, R. M. K.; SEVERO, C. G.; GÖRSKI, E. M. (orgs.). **Sociolinguística e política linguística**: olhares contemporâneos. São Paulo: Blucher, 2016. p. 79-99.
- GUERRA, A. R. **Funções textual-interativas dos marcadores discursivos**. 2007. 233 p. Dissertação (Mestrado em Linguística). Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista "Júlio de Mesquita Filho", São José do Rio Preto, 2007.
- HALLIDAY, M. A. K.; MATTHIESSEN, C. M. I. M. **An Introduction to functional grammar**. 3. ed. Great Britain: Hodder Arnold, 2004.
- HEINE, B. *et al.* **Grammaticalization**: a conceptual framework. Chicago: The University of Chicago Press, 1991.
- HILST, H. Auto da barca de Camiri. In: DUARTE, E. C. (org.). **Teatro reunido (1967-1969)**. São Paulo: Nankin Editorial, 2000[1968]. p. 194-229.
- HOPPER, P. J. On some principles of grammaticization. In: TRAUGOTT, E. C.; HEINE, B. (eds.) **Approaches to grammaticalization**. Amsterdam /Philadelphia: John Benjamins Publishing Company, 1991. v. 1, p. 17-35.

HOPPER, P. J.; TRAUGOTT, E. C. **Grammaticalization**. Cambridge: Cambridge University Press, 1993.

JAKOBSON, R. Closing statement: Linguistics and poetics. In: SEBEOK, T. A. **Style in language**. Cambridge: M.I.T. Press, 1960. p. 350-377.

JUBRAN, C. S. (org.). **A construção do texto falado**. São Paulo: Contexto, 2019 - Gramática do português culto falado no Brasil; v. I.

KLAFKE, M. F. **Heróis e Coringas no palco: o Teatro de Arena prega a resistência**. 2016. 180 f. Dissertação (Mestrado em Literatura Brasileira). Universidade Federal do Rio Grande do Sul, Instituto de Letras, Porto Alegre, 2016.

KLEIN, H. S.; LUNA, F. V. População e sociedade. In: SCHWARCZ, L. M. (dir.). **História do Brasil nação: 1808-2010 - Vol. 5: REIS, D. A. (coord.). Modernização, ditadura e democracia: 1964-2010**. Rio de Janeiro: Objetiva, 2014. p. 31-74.

LABOV, W. **Padrões sociolinguísticos**. São Paulo: Parábola Editorial, 2008 [1972].

LABOV, W. Where does the linguistic variable stop? A response to Beatriz Lavandera. **Sociolinguistic Working Papers**, Austin, Texas, Southwest Educational Development Laboratory, n. 44, p. 1-22, 1978.

LABOV, W. **Principles of Linguistic Change**. Vol. 1: Internal Factors. Cambridge, MA/Oxford: Blackwell Publishers, 1994.

LAVANDERA, B. Where does the sociolinguistic variable stop? **Language in Society**, n.7, p.171-182, 1978.

LEITE, R. M. **História do teatro no Brasil e na Bahia: das primeiras ações jesuíticas ao Pré-modernismo**. Salvador: UFBA, Escola de Teatro; Superintendência de Educação à Distância, 2022.

LOPES, C. R. S.; RUMEU, M. C. B. A identificação dos perfis socioculturais dos redatores de corpora históricos: encaminhamentos metodológicos. **Diadorim**, Rio de Janeiro, v. 20, n. especial, p. 147-168, 2018.

MARCUSCHI, L. A. **Análise da conversação**. São Paulo: Ática, 1986.

MARTELOTTA, M. E.; VOTRE, S. J.; CEZARIO, M. M. (orgs.). **Gramaticalização no português do Brasil: uma abordagem funcional**. Rio de Janeiro: Tempo Brasileiro, 1996.

MOLLICA, M. C.; BRAGA, M. L. (orgs.). **Introdução à Sociolinguística: o tratamento da variação**. 2. ed. São Paulo: Contexto, 2004.

MONTGOMERY, M. Variation and historical linguistics. In: BAYLEY, R.; LUCAS, C. **Sociolinguistic variation: theories, methods, and applications**. Cambridge: Cambridge University Press, 2007. p. 110-132.

NEVES, M. H. de M. Estudos funcionalistas no Brasil. **D.E.L.T.A.**, v. 15, n. esp., p. 71-104. 1999.

NOGUEIRA, L. C. A gramaticalização do verbo ver no português do interior paulista: uma análise discursivo-funcional. **Mosaico**, São José do Rio Preto, v. 17, n. 1, p. 421-437, 2018.

OLIVEIRA, S. G. de; ROCKENBACH, L. M.; GUTIERRES, A. As três ondas do estudo da variação: a emergência do significado no estudo da variação sociolinguística. **Organon**, Porto Alegre, v. 37, n. 73, p. 268–291, 2022.

OSÓRIO, R. G. **O sistema classificatório de “cor ou raça” do IBGE**. Brasília: IPEA, 2003.

OUSHIRO, L. Conceitos de identidade e métodos para seu estudo na Sociolinguística. **Estudos Linguísticos e Literários**, Salvador, n. 63, p. 304-325, 2019.

PAIVA, M. da C. de; PAREDES SILVA, V. L. Cumprindo uma pauta de trabalho: contribuições recentes do PEUL. **Alfa**, São Paulo, v. 56, n. 3, p. 739-770, 2012.

PAULA, J. V. O assalto. In: MORAIS, C. **O teatro de José Vicente: primeiras obras**. São Paulo: Imprensa Oficial, 2010. p. 126-202. (Coleção Aplauso - Teatro Brasil). Disponível em: <https://aplauso.imprensaoficial.com.br/edicoes/12.0.813.787/12.0.813.787.pdf> Acesso em: 11 jul. 2022.

PENHAVEL, E. Sobre as funções dos Marcadores Discursivos. **Estudos Linguísticos XXXIV**, p. 1296-1301, 2005.

PENHAVEL, E. **Marcadores discursivos e articulação tópica**. 2010. 168 p. Tese (Doutorado em Linguística). Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2010.

PENHAVEL, E. O que diferentes abordagens de marcadores discursivos têm em comum? **Revista (CON)TEXTOS Linguísticos**, Vitória, v. 6, n. 7, p. 78-98, 2012.

POTTIER, B. **Systématique des éléments de relation**. Paris: Klincksieck, 1962.

PRADO, D. A. **O teatro brasileiro moderno: 1930-1980**. São Paulo: Perspectiva, 1988.

PRADO, D. A. A Personagem no Teatro. In: CANDIDO, A. **A personagem de ficção**. 10. ed. São Paulo: Perspectiva, 2000. p. 81-101.

PRETI, D. (org.). **Análise de textos orais**. 4. ed. São Paulo: Humanitas, 1999.

PRETI, D. (org.). **Fala e escrita em questão**. São Paulo: Humanitas / FFLCH / USP, 2000.

PRETI, D. **Estudos de língua oral e escrita**. Rio de Janeiro: Lucerna, 2004.

R CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria, 2022. URL <https://www.R-project.org/>.

RIBEIRO, D. **O Povo Brasileiro: a formação e o sentido do Brasil**. São Paulo: Companhia das Letras, 1995.

RISSO, M. S. Marcadores discursivos basicamente sequenciadores. In: JUBRAN, C. S. (org.). **A construção do texto falado**. São Paulo: Contexto, 2019 - Gramática do português culto falado no Brasil; v. I. p. 391-452.

RISSO, M. S.; SILVA, G. M. O.; URBANO, H. Traços definidores dos marcadores discursivos. In: JUBRAN, C. S. (org.). **A construção do texto falado**. São Paulo: Contexto, 2019 - Gramática do português culto falado no Brasil; v. I. p. 371-390.

ROMAINE, S. **Socio-Historical Linguistics**. Its status and methodology. New York: Cambridge University Press, 1982.

ROSA, E. da. Sociolinguística histórica. **Revista de Letras**, Curitiba, v. 17, n. 21, p. 1-17, jul./dez. 2015.

ROSA, E. N. da. A língua escrita como fonte de evidências linguísticas. **Philologus**, Rio de Janeiro, v. 25, n. 73, p. 36-50, jan./abr. 2019.

ROST SNICHELOTTO, C. A. **Olha e vê: caminhos que se entrecruzam**. 2009. 408 p. Tese (Doutorado em Linguística). Universidade Federal de Santa Catarina, Centro de Comunicação e Expressão, Florianópolis, 2009.

SCHIFFRIN, D. **Discourse markers**. Cambridge: Cambridge University Press, 1987.

SCHIFFRIN, D. Discourse Markers: Language, Meaning, and Context. In: SCHIFFRIN, D.; TANNEN, D.; HAMILTON, H. E. (ed.). **The Handbook of Discourse Analysis**. Oxford: Blackwell, 2001. p. 54-75.

SCHIFFRIN, D. Discourse marker research and theory: revisiting *and*. In: FISCHER, K. (ed.). **Approaches to Discourse Particles**. Amsterdam: Elsevier, 2006. p. 315-338.

SCHNEIDER, E. W. Investigating variation and change in written documents. In: CHAMBERS, J. K. TRUGILL, P.; SCHILLING, N. (eds.), **The Handbook of Language Variation and Change**. Blackwell Reference Online, 2007. p. 1-20.

SCHWARCZ, L. M. População e sociedade. In: SCHWARCZ, L. M. (dir.). **História do Brasil nação: 1808-2010 - Vol. 3: SCHWARCZ, L. M. (coord.). A abertura para o mundo: 1889-1930**. Rio de Janeiro: Objetiva, 2012. p. 35-84.

SCHWARCZ, L. M.; STARLING, H. M. **Brasil: uma biografia**. São Paulo: Companhia das Letras, 2018.

SILVA, G. M. de O.; MACEDO, A. T. de. Análise sociolinguística de alguns marcadores conversacionais. In: MACEDO, A. T. de; RONCARATI, C.; MOLLICA, M. C. (orgs.). **Variação e Discurso**. Rio de Janeiro: Tempo Brasileiro, 1996. p. 11-49.

TARALLO, F. **A pesquisa sociolinguística**. São Paulo: Ática, 1997.

TAVARES, M. A.; GÖRSKI, E. M. **Bases teórico-metodológicas para uma interface sociofuncionalista**. Texto apresentado e discutido no XXVII Encontro Nacional da Anpoll – ENANPOLL, Rio de Janeiro - RJ, 2012.

TAVARES, M. A. Sociofuncionalismo: um duplo olhar sobre a variação e a mudança linguística. **Interdisciplinar**, v. 17, p. 27-47, 2013.

TRAPP, K. **Os marcadores discursivos sabe? e entende? na fala de informantes do município de Chapecó/SC**. 2014. 144 p. Dissertação (Mestrado em Linguística) – Programa de Pós-Graduação em Estudos Linguísticos – Universidade Federal da Fronteira Sul, Chapecó, 2014.

TRAUGOTT, E. C. From propositional to textual and expressive meanings: some semantic-pragmatic aspects of grammaticalization. In: LEHMANN, W. P.; MALKIEL, Y.

(eds.). **Perspectives on Historical Linguistics**. Amsterdam: Benjamins, 1982. p. 245-247. (Current Issues in Linguistic Theory 24).

TRAUGOTT, E. C.; HEINE, B. (eds.) **Approaches to grammaticalization**. Amsterdam /Philadelphia: John Benjamins Publishing Company, 1991.

TRAUGOTT, E. C. **The role of the development of discourse markers in a theory of grammaticalization**. Manchester: Stanford University, 1995.

UBERSFELD, A. **Para ler o teatro**. Tradução: José Simões. São Paulo, Perspectiva, 2005[1996]. Título original: Lire le théâtre.

URBANO, H. Aspectos basicamente interacionais dos marcadores discursivos. In: NEVES, M. H. de M. (org.). **Gramática do Português Falado**. São Paulo: Humanitas/FFLCH/USP; Campinas: Editora da Unicamp; v.1, 1999a, p. 195-258.

URBANO, H. Marcadores conversacionais. In: PRETI, D. (org.). **Análise de textos orais**. 4. ed. São Paulo: Humanitas, 1999b. p. 81-101.

URBANO, H. Marcadores discursivos basicamente interacionais. In: JUBRAN, C. C. A. S.; KOCH, I. G. V. (orgs). **Gramática do português culto falado no Brasil**. Campinas: Editora da UNICAMP, 2006. v. 1 – Construção do texto falado, p. 497-527.

URBANO, H. Marcadores discursivos basicamente interacionais. In: JUBRAN, C. S. (org.). **A construção do texto falado**. São Paulo: Contexto, 2019 - Gramática do português culto falado no Brasil; v. I. p. 453-481.

VALLE, C. R. M. **Sabe? ~ não tem? ~ entende?:** itens de origem verbal em variação como requisitos de apoio discursivo. 2001. 183 p. Dissertação (Mestrado em Linguística) – Universidade Federal de Santa Catarina, Florianópolis, 2001.

VALLE, C. R. M. **Multifuncionalidade, mudança e variação de marcadores discursivos derivados de verbos cognitivos:** forças semântico-pragmáticas, estilísticas e identitárias em competição. 2014. 415 p. Tese (Doutorado em Linguística). Universidade Federal de Santa Catarina, Centro de Comunicação e Expressão, Florianópolis, 2014.

VALLE, C. R. M.; GÖRSKI, E. M. A entrevista sociolinguística como locus de significados socioestilísticos: categorias macrosociológicas, identidade local e individual. **Domínios de Lingu@gem**, v. 13, n. 3, p. 1228-1265, 2019.

WEINREICH, U.; LABOV, W.; HERZOG, M. **Fundamentos empíricos para uma teoria da mudança linguística**. São Paulo: Parábola Editorial, 2006 [1968].

ZANIN, I. C. A.; GONÇALVES, S. C. L. Sabe? e entendeu? como marcadores de busca de aprovação discursiva: variantes de uma mesma variável? **PROLÍNGUA**, [s. l.], v. 17, n. 1, p. 62-77, 2022.