



Programa de Pós-Graduação em
LINGUÍSTICA

**A LÍNGUA NA TELA: DESCRIÇÃO E ANÁLISE DA VARIAÇÃO DO OBJETO
DIRETO ANAFÓRICO DE TERCEIRA PESSOA EM LEGENDAS
PROFISSIONAIS E *FANSUBS***

São Carlos
2024





UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE EDUCAÇÃO E CIÊNCIAS HUMANAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA

LÍVIA OLIVEIRA AZEVEDO

**A LÍNGUA NA TELA: DESCRIÇÃO E ANÁLISE DA VARIAÇÃO DO
OBJETO DIRETO ANAFÓRICO DE TERCEIRA PESSOA EM
LEGENDAS PROFISSIONAIS E *FANSUBS***

São Carlos – SP
Agosto de 2024

LÍVIA OLIVEIRA AZEVEDO

**A LÍNGUA NA TELA: DESCRIÇÃO E ANÁLISE DA VARIAÇÃO DO
OBJETO DIRETO ANAFÓRICO DE TERCEIRA PESSOA EM
LEGENDAS PROFISSIONAIS E *FANSUBS***

Dissertação apresentada ao Programa de Pós-Graduação em Linguística da Universidade Federal de São Carlos, como requisito à obtenção do título de Mestra em Linguística.

Linha de pesquisa: Descrição, análise e processamento automático de línguas naturais

Orientadora: Profa. Dra. Caroline Carnielli Biazolli

Bolsa: CAPES

São Carlos – SP

Agosto de 2024



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Educação e Ciências Humanas
Programa de Pós-Graduação em Linguística

Folha de Aprovação

Defesa de Dissertação de Mestrado da candidata Livia Oliveira Azevedo, realizada em 02/08/2024.

Comissão Julgadora:

Profa. Dra. Caroline Carnielli Biazolli (UFSCar)

Profa. Dra. Livia Oushiro (UNICAMP)

Prof. Dr. Ronald Beline Mendes (USP)

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Linguística.

Azevedo, Livia Oliveira

A língua na tela: descrição e análise da variação do objeto direto anafórico de terceira pessoa em legendas profissionais e fansubs / Livia Oliveira Azevedo -- 2024. 217f.

Dissertação (Mestrado) - Universidade Federal de São Carlos, campus São Carlos, São Carlos

Orientador (a): Caroline Carnielli Biazolli

Banca Examinadora: Ronald Beline Mendes, Livia Oushiro

Bibliografia

1. Objeto direto anafórico. 2. Legenda audiovisual. 3. Gêneros textuais-discursivos. I. Azevedo, Livia Oliveira. II. Título.

Ficha catalográfica desenvolvida pela Secretaria Geral de Informática (SIn)

DADOS FORNECIDOS PELO AUTOR

Bibliotecário responsável: Arildo Martins - CRB/8 7180

À Elane, que, com um abraço e algumas palavras,
se tornou lembrete constante do que eu posso ser.

AGRADECIMENTOS

Aos meus pais, Lara e Eduardo, por me cercarem de um amor inabalável, me darem apoio em todas as minhas empreitadas pessoais, artísticas e acadêmicas, e me mostrarem o valor da educação e da ciência.

À minha irmã, Raquel, por ser hoje uma companheira em tantos momentos, pacientemente me ajudando a lidar com as diversas demandas que eu arranjo e estando sempre disponível para me socorrer nos problemas tecnológicos.

Aos outros membros da minha família que, mesmo distantes, sempre se mostraram interessados no que eu fazia, deixando claro para todos o quanto se orgulhavam de mim e da minha trajetória.

À minha companheira Beatriz, por todo o afeto e carinho que me dá, por estar sempre ao meu lado, por não me deixar esmorecer e por compartilhar comigo os sonhos da vida acadêmica.

Aos amigos da Best, por serem os melhores sócios que eu poderia ter, por acompanharem meu crescimento pessoal e profissional ao longo dos anos e por entenderem os muitos momentos em que eu não pude estar com eles.

Aos amigos do Teatro, por serem tantos e tão queridos, por dividirem comigo não só os palcos, mas a vida, e por garantirem que eu não esqueceria de mim mesma nesse processo, fornecendo um afago em todos os nossos encontros.

Aos amigos que fiz nessa pós-graduação, em especial à Dinastia Carolíngia, por terem compartilhado comigo as disciplinas, as dúvidas, as aflições e as aventuras desses dois anos.

Aos amigos do SoLAR, por terem me acolhido com tanto carinho em seu grupo, e por terem me presenteado com discussões tão ricas nas nossas reuniões.

Aos colegas do projeto Pró-norma plural, por terem proporcionado tantos encontros e momentos de aprendizagem, e por me incluírem em uma equipe verdadeiramente estelar, com a qual aprendo todos os dias.

À minha amiga Carol, por ter entrado na minha vida com um tênis All Star e um sorriso, por ter aceitado dividir as mesas de bar comigo, e por ter me doado seu tempo em tantos momentos, me aconselhando, me incentivando, e torcendo por mim.

Às minhas amigas Giovana e Monique, por estarem ao meu lado há mais de 10 anos, por serem presença ativa e constante na minha vida mesmo com a distância, e por estarem sempre abertas a me ouvir e me ajudar.

Ao Programa de Pós-graduação em Linguística e seus professores, por promoverem valiosos espaços de discussão e por terem contribuído com a minha formação como linguista.

Ao professor Ronald Beline Mendes, por aceitar o convite para ser membro das bancas dos Exames de Qualificação e Defesa e pelos apontamentos certos que tanto colaboraram com a minha pesquisa.

À professora Livia Oushiro, por ter me recebido de braços abertos como aluna especial na Unicamp, por ter me ensinado tanto sobre estatística e por ter topado fazer partes das bancas dos Exames de Qualificação e Defesa, trazendo contribuições inestimáveis para que meu estudo pudesse se tornar melhor.

À professora Monique Débora Alves de Oliveira Lima, pelas trocas acerca do nosso objeto de estudo, pela gentileza em compartilhar comigo seus *scripts*, e por participar das bancas dos Exames de Qualificação e de Defesa como suplente.

À professora Rosane de Andrade Berlinck, por ter ministrado a primeira aula que tive no curso de Letras, se tornando minha primeira referência na Sociolinguística, e por ter aceitado prontamente o convite para ser membra suplente das minhas bancas dos Exames de Qualificação e Defesa.

À minha querida orientadora, Caroline Carnielli Biazolli, por ser sempre presente e humana, por ter incessantemente depositado em mim sua confiança, por ser fonte contínua de encorajamento, e por me presentear carinhosamente com seu conhecimento e apoio. Por ter topado a empreitada de me orientar, sabendo dos desafios que isso implicaria, e pelo laço que pudemos construir, ficam aqui minhas palavras de eterna gratidão. Que nós sigamos juntas e fortes.

À CAPES, pelo apoio financeiro concedido desde março de 2023, na forma de bolsa de estudos.

There's only us, there's only this
Forget regret, or life is yours to miss
No other path, no other way
No day but today.

(Finale, 1999)

RESUMO

Partindo do pressuposto de que um olhar para as intersecções entre a variação linguística e os gêneros textuais-discursivos é não só desejável, mas imprescindível para a compreensão da variação (Vieira; Lima, 2019; Biazolli; Berlinck, 2021; Lima, 2022; Vieira, em andamento), o presente trabalho possui como objetivos principais (i) discutir como se dá a realização do objeto direto anafórico de terceira pessoa em um *corpus* de legendas da série *Grey's Anatomy* e (ii) descrever a legenda audiovisual como um gênero. Tomando como base, então, os princípios teórico-metodológicos da Teoria da Variação e Mudança Linguísticas (Weinreich; Labov; Herzog, 2006[1978]; Labov, 1982, 1994, 2001, 2003, 2008[1972]), o primeiro passo desta pesquisa foi a elaboração do *corpus* de legendas que reuniu duas amostras: uma de legendas profissionais extraídas da plataforma de *streaming Amazon Prime Video* por intermédio de um *script* de usuário, e outra de legendas feitas por fãs, baixadas gratuitamente do *site* Legendas.TV e conhecidas como *fansubs*. Com o material em mãos, teve início a caracterização da legenda a partir do conceito de gênero textual-discursivo (Bakhtin, 2016[1979]; Marcuschi, 2005, 2008, 2010), levando em conta suas características estruturais, trabalhadas a partir de princípios da Tradução Audiovisual (Zabalbeascoa, 2008; Díaz-Cintas, 2010; Remael, 2010; Franco; Araújo, 2011), e situacionais, abordadas a partir da proposta de Biber e Conrad (2009). Concomitantemente, foram coletadas todas as ocorrências do fenômeno investigado e definidas as variáveis independentes do estudo, levando em conta tanto a literatura disponível sobre a realização do objeto direto anafórico de terceira pessoa (Omena, 1978; Duarte; 1986; Cyrino, 1994; Marafoni, 2004; Freire, 2005; Santana, 2016; Lima, 2022, entre outros), como as características específicas do gênero textual-discursivo em foco no trabalho. Realizada a codificação de cada ocorrência, passou-se, então, à análise estatística dos dados, na plataforma R (R Core Team, 2023). Como resultado geral, observou-se uma prevalência do clítico acusativo no *corpus*, tendo essa sido a variante mais utilizada em todos os contextos, com exceção apenas de dois: diante de antecedentes de traço não animado e como complementos de formas não finitas, que tiveram como estratégia mais produtiva o objeto nulo. O fenômeno se correlacionou não só a variáveis linguísticas, mas também a extralinguísticas, inclusive relacionadas ao gênero do qual foram coletados os dados, como número de caracteres por linha e por segundo. Unindo os resultados das análises estatísticas ao olhar qualitativo voltado às legendas, concluiu-se que as amostras analisadas neste estudo se revelam bastante homogêneas, com pouca diferença entre as legendas profissionais e as *fansubs*, o que sugere que os diferentes contextos de produção desses dois tipos de legenda não têm um impacto relevante sobre os usos linguísticos encontrados. Além disso, constatou-se que a realização do acusativo anafórico de terceira pessoa ocorre de maneira particular nas legendas, relacionando-se intimamente às especificidades estruturais e situacionais desse gênero textual-discursivo.

Palavras-chave: Objeto direto anafórico. Legenda audiovisual. Variação linguística. Gêneros textuais-discursivos.

ABSTRACT

Assuming that an examination of the intersections between linguistic variation and genres is not only desirable but essential to understand variation (Vieira; Lima, 2019; Biazolli; Berlinck, 2021; Lima, 2022; Vieira, in progress), this study aims to (i) discuss how the realization of third-person anaphoric direct object occurs in a *corpus* of subtitles for the TV show Grey's Anatomy and (ii) describe audiovisual subtitles as a genre. Based on the theoretical and methodological framework of the Theory of Variation and Linguistic Change (Weinreich; Labov; Herzog, 2006[1978]; Labov, 1982, 2001, 2003, 2008[1972]), the first step of the research was the compilation of the *corpus* of subtitles, which gathered two samples: one of professional subtitles extracted from the Amazon Prime Video streaming platform using a user script, and another of fan-made subtitles downloaded for free from the Legendas.TV website – the so-called fansubs. With this material, the characterization of the subtitles based on the concept of genre (Bakhtin, 2016[1979]; Marcuschi, 2005, 2008, 2010) began, considering structural characteristics discussed through the principles of Audiovisual Translation (Zabalbeascoa, 2008; Díaz-Cintas, 2010; Remael, 2010; Franco; Araújo, 2011), and situational characteristics, based on the proposal of Biber and Conrad (2009). Simultaneously, all occurrences of the phenomenon were collected, and the independent variables of the study were defined, considering both the literature on the phenomenon available (cf. Omena, 1978; Duarte; 1986; Cyrino, 1994; Marafoni, 2004; Freire, 2005; Santana, 2016; Lima, 2022, among others), and the specificities of the genre at focus. After each occurrence was encoded, the statistical analysis of the data was conducted using the R platform (R Core Team, 2023). As a general result, a prevalence of the accusative clitic was observed in the *corpus*, with this being the most used variant in all contexts, except in two: when referring to antecedents that were not animated and as complements of non-finite forms, both of which had the null object as the most productive strategy. The phenomenon was correlated not only to linguistic but also to extralinguistic variables, including genre-related ones such as number of characters per line and per second. Combining the results of the statistical analyses with a qualitative examination focused on the subtitles, it was concluded that the analyzed samples seem to be quite homogeneous, with little difference between professional subtitles and fansubs, which suggests that the different production contexts of these two types of subtitles do not have a relevant impact on the linguistic uses observed. Furthermore, the realization of the third-person anaphoric accusative seemed to occur in a particular way in subtitles, presenting strong connection with the structural and situational specificities of this genre.

Keywords: Anaphoric direct object. Audiovisual subtitles. Linguistic variation. Genre.

LISTA DE FIGURAS

Figura 1 – Captura de tela de uma cena do episódio 14x24 da série <i>Grey's Anatomy</i> , com uma legenda profissional.....	26
Figura 2 – Captura de tela de uma cena do episódio 14x24 da série <i>Grey's Anatomy</i> , com uma legenda feita por fãs.....	26
Figura 3 - Representação dos eixos da concepção e do meio, de acordo com Marcuschi (2010[2001]).....	47
Figura 4 - Representação das conexões entre diferentes classificações de gêneros.....	49
Figura 5 - Representação gráfica da intertextualidade em processos de tradução	100
Figura 6 - Diagrama das possibilidades de distribuição de textos audiovisuais	102
Figura 7 - Divisão entre os tipos de legenda a partir do parâmetro linguístico.....	106
Figura 8 - Captura de tela do <i>site</i> do <i>Amazon Prime Video</i> após a ativação do <i>script</i> de usuário, mostrando o botão para <i>download</i> das legendas de um episódio	128
Figura 9 - Captura de tela do <i>site</i> do <i>Amazon Prime Video</i> após a ativação do <i>script</i> de usuário, mostrando o botão para <i>download</i> das legendas de uma temporada.....	128
Figura 10 - Captura de tela do arquivo original da <i>fansub</i> para o episódio 6x10 de <i>Grey's Anatomy</i>	129
Figura 11 - Captura de tela do arquivo convertido da <i>fansub</i> para o episódio 6x10 de <i>Grey's Anatomy</i>	130

LISTA DE QUADROS

Quadro 1 - Sistema pronominal do português brasileiro.....	65
Quadro 2 - Quadro dos pronomes pessoais de terceira pessoa no PB.....	72
Quadro 3 - Estudos sociolinguísticos levantados para a pesquisa, organizados de acordo com autor, ano, título, modalidade e material analisado	74

LISTA DE GRÁFICOS

Gráfico 1 - Frequências e proporções das variantes da expressão de futuro na amostra-piloto inicial	132
Gráfico 2 – Expressão de futuro na amostra-piloto inicial por tipo de legenda	132
Gráfico 3 - Frequências e proporções de próclise e ênclise na amostra-piloto inicial	132
Gráfico 4 – Colocação pronominal na amostra-piloto inicial por tipo de legenda.....	132
Gráfico 5 - Frequências e proporções das variantes da realização do OD anafórico na amostra-piloto inicial.....	133
Gráfico 6 - Realização do OD anafórico na amostra-piloto inicial por tipo de legenda	133
Gráfico 7 - Frequências e proporções de próclise e ênclise na amostra-piloto expandida	134
Gráfico 8 - Colocação pronominal na amostra-piloto expandida por tipo de legenda	134
Gráfico 9 - Frequências e proporções das variantes da realização do OD anafórico na amostra-piloto expandida	134
Gráfico 10 - Realização do OD anafórico na amostra-piloto expandida por tipo de legenda	134
Gráfico 11 - Frequências e proporções das quatro variantes da realização do OD anafórico de terceira pessoa no <i>corpus</i> de legenda completo (N = 1474).....	158
Gráfico 12 – Proporções do OD anafórico de terceira pessoa de acordo com a natureza do antecedente (N = 1474)	159
Gráfico 13 - Frequências e proporções das quatro variantes da realização do OD anafórico de terceira pessoa com antecedentes nominais (N = 972).....	160
Gráfico 14 - Frequências e proporções das variantes do OD anafórico de terceira pessoa de acordo com o traço semântico do antecedente (N = 955)	168
Gráfico 15 - Frequências e proporções das variantes do OD anafórico de terceira pessoa de acordo com a forma verbal (N = 955)	170
Gráfico 16 - Frequências e proporções das variantes do OD anafórico de terceira pessoa de acordo com a estrutura sintática (N = 955)	171
Gráfico 17 - Frequências e proporções das variantes do OD anafórico de terceira pessoa de acordo com o tópico (N = 955).....	174
Gráfico 18 – Distribuição dos pronomes lexicais em legendas de acordo com o tópico (N = 17)	175
Gráfico 19 - Frequências e proporções das variantes do OD anafórico de terceira pessoa de acordo com a relação entre participantes (N = 955).....	176

Gráfico 20 - Proporção de utilização das variantes do acusativo anafórico de terceira pessoa de acordo com o número de CPL	178
Gráfico 21 - Distribuição dos dados em relação ao número de CPL das legendas do corpus	179
Gráfico 22 - Proporção de utilização das variantes do acusativo anafórico de terceira pessoa de acordo com o número de CPS	181
Gráfico 23 - Distribuição dos dados em relação ao número de CPS das legendas do <i>corpus</i>	182
Gráfico 24 - Frequências e proporções das variantes do OD anafórico de terceira pessoa de acordo com a função sintática do antecedente (N = 955).....	184
Gráfico 25 - Frequências e proporções das variantes do OD anafórico de terceira pessoa de acordo com o tipo de legenda (N = 955)	185

LISTA DE TABELAS

Tabela 1 - Características situacionais para análise dos gêneros	50
Tabela 2 - Número de palavras na amostra-piloto inicial, de acordo com episódio e tipo de legenda.....	131
Tabela 3 - Número de palavras na amostra-piloto expandida, de acordo com episódio e tipo de legenda.....	133
Tabela 4 - Número de palavras no <i>corpus</i> , de acordo com episódio e tipo de legenda	136
Tabela 5 - Resultado da análise de regressão logística multinomial para a realização do OD anafórico de terceira pessoa em legendas audiovisuais (N = 955).....	164
Tabela 6 – Recomendações dos guias de legendagem para os parâmetros técnicos e resultados da análise das legendas	190
Tabela 7 – Configuração das características situacionais em legendas profissionais e <i>fansubs</i> , a partir da proposta de Biber e Conrad (2009)	197

LISTA DE ABREVIATURAS E SIGLAS

CPL	Caracteres por linha
CPS	Caracteres por segundo
OD	Objeto direto
ON	Objeto nulo
PB	Português brasileiro
PE	Português europeu
S	Sujeito
SN	Sintagma nominal
TAV	Tradução audiovisual
V	Verbo
VP	Variável previsor
VR	Variável resposta

SUMÁRIO

1 INTRODUÇÃO	21
1.1 Objetivos e hipóteses	27
1.2 Metodologia	28
1.3 Organização da dissertação	30
2 PRESSUPOSTOS TEÓRICOS	32
2.1 Conceitos básicos da Sociolinguística Variacionista	32
2.2 Gêneros textuais-discursivos: definições e características	42
2.3 Trabalhando na interface: estudos variacionistas e gêneros textuais-discursivos.....	53
3 O OBJETO DIRETO ANAFÓRICO: DA TRADIÇÃO GRAMATICAL AOS ESTUDOS LINGUÍSTICOS	63
3.1 O OD anafórico e a tradição gramatical	63
3.2 O OD anafórico em estudos descritivos	68
3.3 O OD anafórico em pesquisas (sócio)linguísticas	74
4 A TRADUÇÃO E A LEGENDA AUDIOVISUAL.....	95
4.1 Conceitos básicos de tradução	95
4.2 A tradução audiovisual.....	100
4.2.1 A legenda audiovisual.....	105
4.2.1.1 As legendas profissionais	115
4.2.1.2 As <i>fansubs</i>	117
5 PROCEDIMENTOS METODOLÓGICOS.....	122
5.1 A elaboração do <i>corpus</i>	123
5.1.1 Definição da série e recortes metodológicos.....	123
5.1.2 Obtenção e tratamento das legendas	125
5.2 Das legendas ao objeto de estudo: o trabalho com a amostra-piloto	130
5.3 A composição final do <i>corpus</i>	135

5.4 Trabalhando com os dados: procedimentos de coleta, codificação e análise	136
5.4.1 Coleta dos dados.....	136
5.4.2 Codificação dos dados.....	137
5.4.3 Análises estatísticas	138
5.5 Variáveis e hipóteses	140
5.5.1 Variáveis independentes e/ou previsoras	142
5.5.1.1 Variáveis linguísticas	142
5.5.1.1.1 Natureza do antecedente	142
5.5.1.1.2 Traço semântico do antecedente	143
5.5.1.1.3 Função sintática do antecedente.....	145
5.5.1.1.4 Forma verbal	147
5.5.1.1.5 Estrutura sintática	148
5.5.1.2 Variáveis extralinguísticas	150
5.5.1.2.1 Tópico	150
5.5.1.2.2 Relação entre participantes	152
5.5.1.2.3 Tipo de legenda.....	153
5.5.1.2.4 Número de CPL	154
5.5.1.2.5 Número de CPS	155
5.6 A descrição da legenda como um gênero.....	155
6 ANÁLISE DE DADOS E DISCUSSÃO DOS RESULTADOS.....	157
6.1 O OD anafórico de terceira pessoa em legendas audiovisuais	157
6.1.1 Distribuição geral	157
6.1.2 A análise multivariada	163
6.2 A legenda audiovisual como um gênero	186
6.2.1 O gênero legenda: aspectos estruturais	188
6.2.2 O gênero legenda: aspectos situacionais	193
7 CONSIDERAÇÕES FINAIS.....	199

REFERÊNCIAS206

APÊNDICE A.....216

1 INTRODUÇÃO

Historicamente, o campo da linguística sempre foi marcado por diversas maneiras de compreender o funcionamento das línguas. Após longos anos de desenvolvimento, imperavam, até o início da segunda metade do XX, teorias como o estruturalismo e o gerativismo, para as quais a dimensão social era secundária, uma vez que tratavam a língua como um sistema homogêneo no qual a descrição e a explicação da variação nos usos não eram relevantes, sendo ela tratada como uma ocorrência desregrada e caótica.

Foi como resposta a essas ideias que surgiu, no início dos anos 60, a área da linguística conhecida como Sociolinguística Variacionista. Inaugurada com os estudos acerca das relações entre língua e sociedade realizados pelo pesquisador americano William Labov, essa corrente se pautou no princípio de que toda língua é, por essência, heterogênea, sendo marcada por variações inerentes ao sistema. Nesse contexto, são tomados como relevantes não apenas os fatores linguísticos, mas também os sociais e estilísticos que influenciam e motivam a escolha de uma ou outra variante (Weinreich; Labov; Herzog, 2006[1968]; Labov, 1982, 1994, 2001, 2003, 2008[1972]).

Para a análise linguística, portanto, torna-se essencial utilizar situações concretas de comunicação como ponto de partida. Tal necessidade, aliada ao fato de os estudos iniciais de Labov (2008[1972]) analisarem variáveis fonológicas, levou pesquisadores da área a demonstrar forte preferência por produções advindas da língua falada na composição de seus *corpora*. Por isso, estudos variacionistas têm, ao longo dos anos, focado e descrito dados da modalidade oral da língua, com o objetivo de contribuir para o desenvolvimento dos aspectos teórico-metodológicos da Sociolinguística Variacionista e empreender descrições cada vez mais precisas da língua.

Entretanto, a fala não é a única forma de interação entre os falantes que se mostra valiosa para a apreensão da variação linguística. São também legítimas as manifestações linguísticas realizadas em outras modalidades de uso da língua, que se materializam através de gêneros textuais-discursivos diversos¹, de modo que é imprescindível, para uma compreensão mais abrangente do funcionamento linguístico, que eles sejam amplamente explorados.

¹ Nesta pesquisa, gêneros textuais-discursivos são entendidos como tipos relativamente estáveis de enunciados, formados por três elementos essenciais: o conteúdo temático, o estilo e a construção composicional (Bakhtin, 2016[1979]). Reconhece-se, ainda, a questão da inexistência terminológica em relação aos gêneros, discutida por Rojo (2005), sendo aqui adotado o termo gênero textual-discursivo por se considerar, de maneira ampla, tanto os aspectos da materialidade textual como aqueles referentes ao discurso. A seção 2 traz mais considerações a respeito desse tópico.

Além disso, os gêneros textuais-discursivos estão associados a estilos específicos e se combinam com diferentes modalidades e normas linguísticas (Marcuschi, 2008, 2010), criando interfaces interessantes para a pesquisa sociolinguística. Com efeito, e tal como afirmam Biazolli e Berlinck (2021, p. 27),

[...] pensarmos nos gêneros textuais-discursivos no momento da constituição do *corpus* de uma pesquisa, como o evento comunicativo em si, nos possibilita investir na análise de dados relacionados a monitoramento e formalidade mais intensos, menos intensos e a uma zona intermediária entre esses extremos.

Esse cuidado em considerar os gêneros textuais-discursivos, porém, não deve se fazer presente apenas em um momento inicial, de composição do *corpus* de análise. Por ser fator determinante nos usos e opções linguísticas, a questão do gênero deve permear, também, todos os passos da pesquisa, da delimitação do fenômeno a ser estudado e da seleção de dados, vinculadas à escolha do *corpus*, à definição das categorias de análise e à interpretação dos resultados (Biazolli; Berlinck, 2021).

Baseada nesses pressupostos, a presente pesquisa propõe a investigação da variação linguística na realização do objeto direto (OD) anafórico de terceira pessoa em legendas audiovisuais, aqui tratadas como um gênero textual-discursivo, uma vez que se caracterizam como um evento comunicativo inserido em uma prática social discursiva, com conteúdo temático, estilo – na concepção bakhtiniana – e construção composicional específicos. Tal escolha se deve a alguns motivos, mencionados a seguir.

Quanto ao objeto de pesquisa, cabe mencionar que se trata de um fenômeno já bem caracterizado no português brasileiro (PB), que foi e é tema de uma variedade de estudos, em diferentes perspectivas. Esse interesse pela retomada anafórica do OD se dá, principalmente, por ser uma estrutura que permite diversas formas para sua realização, com comportamentos distintos nas modalidades oral e escrita da língua.

Depois da experiência pioneira de Omena (1978), Duarte (1986) destaca quatro formas para a realização do fenômeno, sendo elas: o clítico acusativo; o pronome lexical; o sintagma nominal (SN) anafórico; e a categoria vazia ou objeto nulo, ilustradas, respectivamente, pelos exemplos (1) a (4), extraídos do *corpus* desta pesquisa²:

(1) B1: Ele não aguenta?

² Em todos os exemplos do *corpus* desta pesquisa, o OD anafórico encontra-se destacado em itálico e em negrito, ao passo que seu antecedente é sublinhado.

Agora você é a imbecil.

Deixar o tecido morrer
vai matá-**lo**. (FT10E01)³

- (2) AM1: Eu... não consigo dormir.

Não consigo dormir porque
fico vendo o rosto dele o tempo todo.

Vejo **ele** nos pontos de ônibus. (PT06E01)

- (3) M1: - Alex, pegue o bebê.

A1: - Não dá tempo agora.

M1: - Alex!

Pegue **o meu filho**. (PT10E01)

- (4) A3: Sei que a dor é real

e não vou descansar

até descobrir o que é

e como melhorar.

Tem minha palavra. Os dois tem [Ø].

Certo. (FT14E01)

De maneira geral, foi observada por Duarte (1986) uma forte preferência pelo uso do objeto nulo (com 62,6% das ocorrências) em seu *corpus* de língua falada, em contraste com um menor número de ocorrências para o SN anafórico (17,1%), para o pronome lexical (15,4%) e para o clítico acusativo (4,9%), mesmo este último sendo a estratégia apontada pelos compêndios gramaticais tradicionais (cf. Rocha Lima, 2011; Cunha e Cintra, 2017; Bechara, 2009) como a mais adequada. Como possível explicação para isso, Duarte aponta o pedantismo

³ Todos os dados utilizados como exemplos estarão identificados de acordo com a amostra, a temporada e o episódio. Assim, uma letra “P” no início do código indica uma legenda profissional e um “F” uma *fansub*; o número que segue a letra “T” indica a temporada e o que vem após o “E”, o episódio. Por exemplo, o código FT06E01 indica que a ocorrência foi retirada da *fansub* para o primeiro episódio da sexta temporada.

relacionado ao uso do clítico acusativo e a estigmatização do pronome lexical na posição de objeto – ambos mencionados por seus informantes –, de modo que o uso do objeto nulo representaria uma maneira não marcada para a realização do fenômeno.

Além disso, a tendência constatada pela autora relaciona-se, também, ao que foi revelado por pesquisas diacrônicas. Tarallo (1983), por exemplo, constatou uma diminuição na retenção do OD anafórico ao longo da história da língua, com uma queda na ocorrência de clíticos, sendo a partir do século XX que a categoria vazia passa a imperar nesse constituinte da oração. Esse resultado é corroborado, posteriormente, pelo estudo de Cyrino (1994), que mostra que é durante a segunda metade do século XIX que aparecem os primeiros pronomes lexicais funcionando como OD.

Após esses trabalhos seminais, vários outros se seguiram, buscando descrever o comportamento do fenômeno em diversas variedades do PB, modalidades de uso da língua e gêneros textuais-discursivos – como, por exemplo, Freire (2000, 2005), Marafoni (2004), Arruda (2006, 2012), Costa, S. (2011), Costa, T. (2012), Lauar (2015), Pivetta (2015), Pinto (2015), Santana (2016), Rodrigues (2018), Spinelli (2018), Marques de Sousa (2021) e Lima (2022), para citar alguns.

Contudo, apesar de a realização do OD anafórico já ter sido extensivamente estudada no PB, ainda há uma série de gêneros textuais-discursivos nos quais suas ocorrências ainda não foram exploradas, em especial naqueles tratados como emergentes, como caracterizados por Marcuschi (2009). Tais gêneros, de acordo com o autor, seriam aqueles que emergiram nas últimas décadas a partir do desenvolvimento de novas tecnologias e que, muitas vezes, se dão no meio eletrônico ou digital, comumente misturando características da fala e da escrita, o que os torna interessantes para pesquisas sobre a língua devido a essa fluidez.

De fato, o trabalho com gêneros emergentes permite que o pesquisador tenha acesso a dados que foram produzidos para fins outros que não o de análise linguística, ao contrário do que acontece com aqueles advindos de entrevistas, elicitacões ou questionários, que são comumente obtidos apenas para servir de amostra em pesquisas sociolinguísticas (Oushiro, 2020). Nesse contexto, as legendas se apresentam como um material particularmente valioso para estudos como o aqui apresentado, pois é marcada por uma natureza essencialmente híbrida, integrando escrita, imagem e som, e pode ser compartilhada tanto em meios oficiais como em contextos informais, gerando um novo tipo de dado linguístico.

Ainda com relação à escolha das legendas como material de trabalho, vale lembrar que, em um mundo globalizado, filmes e séries estrangeiros são cada vez mais assistidos por espectadores brasileiros, impulsionados tanto pelo surgimento de inúmeras plataformas de

streaming, como pelos *downloads* ilegais de arquivos encontrados gratuitamente na internet. Dessa forma, seja qual for a maneira de acessar essas obras, o fato é que essas produções culturais não chegam ao espectador sozinhas: trazem consigo as dublagens e as legendas tão essenciais para que o público possa apreciá-las, de modo que o material linguístico derivado de traduções audiovisuais se faz cada vez mais presente na vida de milhões de pessoas.

Além disso, certas características desse material o tornam interessante para a análise linguística. Por exemplo, a legenda audiovisual se propõe a emular a fala original das personagens, mesmo que limitada pelo meio – gráfico – no qual toma forma, aproximando-se da fala do público para a qual é destinada, de modo que é comum que ela apresente elementos típicos da oralidade. De fato, “[...] é essencial que a diversidade linguística seja respeitada pelos profissionais ao longo do processo de tradução, utilizando variantes mais ou menos formais ou prestigiadas quando essas se fazem necessárias [...]” (Lopes; Afonso, 2021, p. 146).

Ademais, como já há certo nível de interferência na experiência de imersão de quem assiste a um conteúdo legendado (se está o tempo todo ciente do texto em tela), há uma preocupação do tradutor em não atrair mais atenção do que o necessário para esse recurso. Assim, busca-se utilizar nas legendas uma linguagem que se aproxime daquela considerada como natural pelo público, uma vez que “[...] certas escolhas associadas ao quase indefinível português padrão/formal [...] desviariam a atenção do público para a legenda e não para o conteúdo da cena propriamente dito [...]” (Lopes; Afonso, 2021, p. 147).

Desse modo, torna-se interessante observar como o processo tradutório pode interferir no emprego de certas formas linguísticas. Por um lado, há grandes empresas, como as plataformas de *streaming*, que produzem traduções para uma audiência diversificada e ampla; por outro, encontram-se grupos de legendagem voluntária compostos por fãs engajados que produzem legendas para outros fãs, as chamadas *fansubs*, disponibilizadas rapidamente após a exibição de um episódio de forma gratuita via internet (Spolidorio, 2017). São cenários distintos que podem resultar em traduções diferentes, revelando-se como um conjunto de materiais com potencial para a análise da variação linguística.

Com efeito, é possível que um mesmo original seja reelaborado de maneiras distintas quando traduzido por pessoas diferentes, isto é, pode haver variação na tradução. É precisamente isso que as Figuras 1 e 2 ilustram, com duas legendas reais para um episódio da série *Grey's Anatomy*, criadas, respectivamente, por profissionais e por fãs.

Figura 1 – Captura de tela de uma cena do episódio 14x24 da série *Grey's Anatomy*, com uma legenda profissional



Fonte: *Amazon Prime Video* (2022)

Figura 2 – Captura de tela de uma cena do episódio 14x24 da série *Grey's Anatomy*, com uma legenda feita por fãs



Fonte: Equipe InSUBs (2018)

Diante desse contraste, pode-se afirmar que a legenda consiste em um material valioso para estudos sociolinguísticos: além de apresentar variação, caracteriza-se como um gênero essencialmente multimodal, que só se concretiza em conjunto com outras linguagens. Em razão disso, firma-se a importância da análise de dados extraídos tanto de legendas profissionais como de amadoras, diferentes expoentes do gênero legenda audiovisual, posto que as diferenças aqui mencionadas, em conjunto com outros fatores, tais como a influência dos guias de legendagem disponibilizados em cada situação, podem influenciar as escolhas linguísticas dos tradutores.

Ademais, há ainda na literatura uma lacuna no que diz respeito a uma descrição precisa e refinada desse tipo de tradução audiovisual (TAV) como um gênero textual-discursivo. Apesar de já haver propostas como as de Oliveira (2008), Esperandio (2015) e Boito e Caetano (2018), para tratar as legendas dessa maneira (para mencionar algumas), nenhuma delas se debruça com afinco sobre as diferentes características situacionais que permitiriam afirmar que a legenda pode ser considerada, de fato, um gênero. Por isso, a pesquisa aqui proposta também se ocupará de uma discussão teórica acerca dessa classificação, com base nos conceitos de gênero de Bakhtin (2016[1979]), Marcuschi (2005, 2008, 2010[2001]) e Biber e Conrad (2009).

1.1 Objetivos e hipóteses

Objetiva-se averiguar, por intermédio de um estudo descritivo-comparativo, como se dá a realização do OD anafórico em legendas audiovisuais da série *Grey's Anatomy*, buscando compreender não só quais são os aspectos internos e externos que influenciam suas ocorrências, mas também se há uma diferença relevante entre as legendas profissionais, mais institucionalizadas, e as amadoras, que são voluntariamente produzidas.

Aventa-se a hipótese de que haverá uma diferença significativa entre esses dois tipos de legenda, principalmente levando em conta a imagem que os legendistas têm do público que receberá cada produção e as condições de elaboração de cada uma delas. Por exemplo, pode ser que, por serem produzidas por especialistas da língua em um ambiente mais controlado, as legendas profissionais apresentem índices maiores do uso do pronome clítico, por ser a forma indicada como correta pela norma-padrão, em contraste com uma frequência maior de formas mais próximas ao vernáculo na tradução em legendas amadoras.

No entanto, de maneira oposta, é possível, também, que uma tendência contrária seja observada, com mais ocorrências de estruturas próximas ao que recomendam as gramáticas tradicionais nas *fansubs*. Nelas, isso se daria como uma maneira de validar o trabalho executado pelos fãs-tradutores, que poderiam, recorrendo ao uso do português comumente tido como

correto, atestar a qualidade de suas legendas, enquanto a necessidade de legitimação não seria uma preocupação do legendista profissional.

Tenciona-se, também, observar o que os guias de legendagem das equipes informais e das plataformas de *streaming* (disponíveis *online*)⁴ trazem com relação à linguagem a ser utilizada, verificando se há instruções específicas no tocante aos fenômenos discutidos, já que este pode, hipoteticamente, ser um fator decisivo na escolha por uma ou outra forma linguística.

Ademais, busca-se atingir uma descrição detalhada da legenda como um gênero textual-discursivo, levando em conta todos os seus aspectos estruturais e situacionais, assim como as especificidades dos dois tipos de legenda aqui em jogo, as profissionais e as *fansubs*. Com isso, busca-se contribuir para o preenchimento da já mencionada lacuna na literatura acerca da caracterização dos diversos gêneros que circulam na sociedade, ao mesmo tempo que se respaldam as análises e as discussões sobre as interfaces entre o gênero legenda e a variação na realização do OD anafórico de terceira pessoa.

Assim, a presente pesquisa, além de descobrir mais sobre o funcionamento do PB, especialmente ao levar em conta os diversos gêneros que circulam na sociedade brasileira, motivada pela proposição de que “[...] quanto mais se sabe sobre uma língua, mais podemos descobrir sobre ela” (Labov, 1972, p. 98, tradução nossa)⁵, apresenta potencial para: (i) colaborar com os estudos sobre tradução, ao lançar um olhar diferenciado para uma das modalidades da TAV mais difundidas atualmente; (ii) contribuir com os extensivos estudos já realizados no campo da Sociolinguística, ao explorar fenômenos variáveis encontrados em um gênero textual-discursivo ainda pouco explorado por essa área; (iii) ampliar a discussão acerca de novos gêneros textuais-discursivos, com a proposta de uma descrição de um gênero relativamente novo; (iv) aprofundar as discussões acerca da interface entre o estudo da variação linguística e o dos gêneros, apontando caminhos possíveis para lidar com essa realidade ao longo de todo o percurso de uma pesquisa científica. Busca-se, em suma,

1.2 Metodologia⁶

Para a realização desta pesquisa, organizou-se um *corpus* composto de legendas para a série *Grey's Anatomy*, tanto profissionais, extraídas de uma plataforma de *streaming*, como amadoras, obtidas gratuitamente pela internet. A escolha da série se deu graças à sua

⁴ Endereços para a obtenção desses guias serão indicados mais à frente, na seção 4 desta dissertação.

⁵ “[...] the more that is known about a language, the more we can find out about it”.

⁶ Nesta subseção, apresentam-se brevemente as considerações acerca da metodologia aplicada. Na seção 5, as etapas desenvolvidas são descritas em detalhes.

popularidade⁷ e à longevidade da produção, contando com dezenove temporadas e milhares de fãs, além de ser uma série que permite, ao mesmo tempo, a observação de situações comunicativas que envolvem a vida pessoal e a vida profissional de médicos em um hospital, o que, por hipótese, poderia promover uma variedade temática interessante. Com base no trabalho com uma amostra-piloto, na qual foram observadas as ocorrências de três fenômenos variáveis diferentes (a saber, realização do OD anafórico de terceira pessoa, colocação pronominal e expressão de futuro), decidiu-se pela investigação do OD anafórico, posto que foi esta estrutura que demonstrou os resultados mais promissores para o desenvolver do restante da pesquisa, especialmente pela aparente preferência pela clítico acusativo.

Com o objeto de estudo definido, foram estabelecidas as variáveis dependentes e independentes⁸ a serem consideradas nas análises propostas. Para tanto, foram considerados não só os resultados obtidos anteriormente por pesquisadores da área, mas também aspectos específicos da legenda, observados a partir de sua caracterização, tais como o número de caracteres por linha (CPL) e por segundo (CPS), já que eles podem, por hipótese, motivar o uso de uma ou outra variante.

Seguindo os passos de uma pesquisa sociolinguística, foram, então, realizadas análises estatísticas na plataforma *R* (R Core Team, 2023), a fim de atingir uma visão mais ampla da realização do fenômeno em questão. Nesse contexto, foram inicialmente obtidas as frequências e as proporções das variantes de acordo com cada variável independente para que, mais tarde, fossem realizados os testes estatísticos. Para a maioria das variáveis, foi realizado um teste de qui-quadrado; contudo, para as duas variáveis contínuas do estudo, recorreu-se a uma regressão logística para aferir uma possível correlação. Já nas análises multivariadas, foi criado um modelo de regressão logística multinomial que pudesse incluir todas as variáveis predictoras de interesse, com vistas a investigar quais contextos seriam mais influentes na escolha de uma ou outra forma linguística.

⁷ A série em questão conta com inúmeros grupos e páginas a ela dedicados, seguidos por milhares de internautas. Para ilustrar, no momento da escrita do projeto desta pesquisa – março de 2023 –, foi possível encontrar os grupos *Grey's Anatomy* Brasil e Loucos por *Grey's Anatomy*, no *Facebook*, que contavam com 399 mil e 494 mil membros, respectivamente; a página *Grey's Anatomy* Brasil, também no *Facebook*, com 362 mil curtidas; e o perfil *Grey's Depressão* no *Instagram*, seguido por 690 mil pessoas. Além disso, foi a mais pesquisada em 23 estados e no Distrito Federal no ano de 2020, de acordo com a revista *Superinteressante*, e conta com milhares de *downloads* em suas legendas em repositórios *online*.

⁸ Neste trabalho, quando mencionadas de maneira geral, isto é, fora do contexto de um teste estatístico, as variáveis do estudo serão identificadas como dependentes e independentes, por serem esses os termos tradicionalmente empregados em estudos sociolinguísticos. Serão essas as expressões utilizadas, também, ao se referir às variáveis em análises univariadas. Por outro lado, ao considerar as análises multivariadas, tais termos serão substituídos por variável resposta e variável predictor, por ser essa a nomenclatura mais adequada ao contexto. Uma discussão acerca dessas diferenças terminológicas será empreendida na subseção 5.4.3.

Concomitantemente, investiu-se na descrição da legenda audiovisual como um gênero textual-discursivo, levando em conta tanto suas características estruturais como as situacionais. As primeiras foram abordadas não só a partir de estudos da área da TAV, mas também com base nos guias de legendagem disponíveis para as legendas que compuseram o *corpus* desta pesquisa. Já os aspectos situacionais foram trabalhados a partir da proposta de Biber e Conrad (2009), que elencam sete parâmetros para a descrição e comparação de gêneros.

Por fim, a descrição e a interpretação dos resultados foram feitas levando-se em conta duas perspectivas: a da variação e a dos gêneros, buscando identificar se os objetivos propostos foram atingidos e se as hipóteses inicialmente levantadas foram confirmadas.

1.3 Organização da dissertação

Além desta introdução, a presente dissertação é composta por sete seções, dedicadas a diferentes aspectos da investigação realizada.

Na seção 2, são apresentados os pressupostos teóricos que embasam este trabalho, tais como: (i) a Teoria da Variação e Mudança Linguísticas; (ii) as abordagens de gênero textual-discursivo apresentadas por Bakhtin (2016[1979]) e Marcuschi (2005, 2008, 2010[2001]), além das características elencadas por Biber e Conrad (2009) para a descrição e comparação de gêneros; e (iii) alguns dos estudos anteriormente propostos que aproximam as duas perspectivas acima mencionadas.

A seção 3, por sua vez, é dedicada à descrição do objeto de estudo desta pesquisa, a realização do OD anafórico de terceira pessoa, a partir de três enfoques. O primeiro, levando em conta as prescrições de gramáticas tradicionais; o segundo, sob a visão de estudos descritivos como os de Câmara Jr. (1976, 1992[1970]), Mateus *et al.* (2003[1983]) e Castilho (2020[2010]); e, finalmente, o terceiro, com base em trabalhos variacionistas sobre o fenômeno em investigação.

Já a seção 4 é voltada a discussões acerca da TAV e suas modalidades, à apresentação da legenda audiovisual a partir de suas principais características tradutórias e técnicas, e à diferenciação entre legendas profissionais e legendas feitas por fãs, tomando como base, principalmente, seus contextos de produção/recepção.

Em seguida, a seção 5 se dedica à metodologia da pesquisa, trazendo todas as decisões tomadas ao longo de sua realização. São elencados, portanto, todos os passos para a composição do *corpus* de legendas, os resultados obtidos por meio da amostra-piloto, os procedimentos e ferramentas utilizados para a coleta e codificação de dados e os testes estatísticos empregados.

Além disso, listam-se também todas as variáveis que compõem este estudo, com exemplos para cada uma delas, além de se indicar a maneira como foi elaborada a caracterização das legendas.

Na seção 6, realiza-se a descrição detalhada das legendas a partir de seus aspectos estruturais e situacionais, levando em conta as discussões desenvolvidas nas seções 2 e 4. Nesse contexto, compara-se, também, os dois tipos de legenda utilizados, buscando averiguar a existência ou não de diferenças entre eles. Ademais, são apresentados e discutidos os resultados acerca da realização do OD anafórico obtidos a partir das análises estatísticas realizadas e tecem-se comentários acerca da relação entre esses resultados e as características específicas do gênero textual-discursivo legenda.

Por fim, a seção 7 traz as conclusões deste estudo, articuladas aos debates empreendidos nas outras partes do trabalho, seguidas pelas referências bibliográficas e o apêndice.

2 PRESSUPOSTOS TEÓRICOS

Nesta seção, apresentam-se os fundamentos teóricos considerados essenciais para embasar esta pesquisa. Uma vez que, aqui, o objeto de estudo é um fenômeno linguístico variável, inicia-se a discussão com comentários acerca da Teoria da Variação e Mudança Linguísticas (Weinreich; Labov; Herzog, 2006[1968]; Labov, 1982, 1994, 2001, 2003, 2008[1972]). Em seguida, em busca de um aporte teórico que permita realizar a caracterização de legendas audiovisuais como um gênero textual-discursivo, reflete-se sobre as definições de gênero propostas por Bakhtin (2016[1979]) e Marcuschi (2005, 2008, 2010[2001]), além dos critérios para categorização e comparação de gêneros propostos por Biber e Conrad (2009). Por fim, realiza-se uma retomada dos estudos e iniciativas que defendem a importância de se observar fenômenos variáveis em conjunto com os gêneros nos quais estão inseridos, como Biazolli (2010, 2016), Vieira e Lima (2019), Biazolli e Berlinck (2021), Lima (2022) e Vieira (em andamento).

2.1 Conceitos básicos da Sociolinguística Variacionista

No estudo aqui descrito, seguem-se os pressupostos teórico-metodológicos da Teoria da Variação e Mudança Linguísticas (Weinreich; Labov; Herzog, 2006[1968]; Labov, 1982, 1994, 2001, 2003, 2008[1972]), que tem como foco o estudo da língua em seu contexto social real, nas modalidades oral ou escrita, sendo a variação um aspecto inerente à linguagem. Nessa perspectiva, a variabilidade é tratada como algo regular e sistemático, substituindo, portanto, o axioma da categoricidade, defendido no momento anterior ao surgimento da Sociolinguística Variacionista, pelo axioma da heterogeneidade ordenada. Dessa maneira, qualquer descrição de uma língua deve ser empreendida incluindo seus fenômenos variáveis, posto que eles fazem parte de um sistema cuja variação não é arbitrária, mas organizada e condicionada por regras que garantem a estabilidade do sistema.

Nesse sistema, de acordo com Labov (1982, 1994, 2001, 2003, 2008[1972]), são três os tipos de regras que podem ser identificadas a partir da análise linguística: as categóricas, as semicategóricas e as variáveis. Para o autor, as primeiras corresponderiam às regras invioláveis da língua, que devem ser respeitadas para que a compreensão dos enunciados ocorra da maneira desejada. São, portanto, aquelas que abarcam fenômenos para os quais não há variação, com uma forma sendo utilizada 100% das vezes. No caso de fenômenos variáveis, porém, podem ser encontradas regras semicategóricas, quando uma das variantes apresenta um percentual de realização entre 95% e 99%, ou variáveis, quando esse índice fica entre 5% e 95%.

Tais regras variáveis, opostas à concepção estruturalista de regra opcional, fazem referência à possibilidade de duas ou mais formas alternantes – as variantes linguísticas – serem encontradas no mesmo contexto e com valor de verdade idêntico, transmitindo, virtualmente, a mesma informação. Essas variantes referencialmente equivalentes são condicionadas por fatores de natureza linguística ou extralinguística e são encontradas em todos os níveis gramaticais, do fonológico ao discursivo. Ainda, frente às intersecções entre língua e sociedade, cada variante pode receber um significado social e/ou estilístico (Labov, 2008[1972]).

Com essa divisão em mente, espera-se que o objeto de estudo deste trabalho, a realização do OD anafórico de terceira pessoa, se comporte como uma regra variável, com suas variantes apresentando, sempre, índices de aplicação menores do que 95%. Com base na literatura sobre o fenômeno (cf. Omena, 1978; Duarte, 1986; Cyrino, 1994; Freire, 2005; Marques de Sousa, 2021; Lima, 2022, entre outros estudos), explorada na seção 3, tem-se a expectativa geral de encontrar altas frequências para a utilização do objeto nulo nas duas amostras do *corpus* analisadas, enquanto a variante mais desfavorecida seria o pronome lexical. Além disso, em decorrência das restrições espaciais e temporais impostas pelo gênero textual-discursivo analisado, hipotetiza-se que, em alguns contextos – como em legendas que estejam próximas ao limite de caracteres permitido –, nem todas as quatro variantes estejam em competição, mantendo-se, porém, o caráter variável do fenômeno.

Do ponto de vista da Sociolinguística Variacionista, regras variáveis como essa são, como mencionado anteriormente, inerentes ao sistema linguístico de uma comunidade de falantes, de modo que elas coexistem com os outros tipos de regra. A essa comunidade, Labov (2008[1972]) dá o nome de “comunidade de fala”, sendo ela definida como “[...] um grupo que compartilha as mesmas normas a respeito da língua [...]” (p. 188), isto é, um grupo que apresenta relativamente os mesmos traços linguísticos, diferenciando-se de outros e demonstrando as mesmas atitudes com relação à língua (Guy, 2001).

Apesar de essa noção ter sido questionada ao longo do tempo e outros pesquisadores terem proposto outros tipos de agrupamento, como as redes sociais de Milroy (1980) e as comunidades de prática de Eckert (2005), o fato é que, em geral, estudos sociolinguísticos voltam o olhar para a língua em uso quando o grau de monitoramento é o menor possível, ou seja, para o vernáculo. De fato, para Labov (2008[1972]), é esse o âmbito em que a língua é falada de maneira mais natural, possibilitando, então, o estudo empírico da variação e mudança linguísticas e sua sistematicidade, a partir de dados reais.

Nesse contexto, para além de questões linguísticas, são importantes, também, fatores relacionados à influência que os comportamentos sociais exercem sobre o uso da língua, tais

como a identidade do falante e do interlocutor e os aspectos que compõem a situação comunicativa. Esses fatores, porém, não afetam somente a língua falada, mas todo tipo de comunicação empreendida pelos falantes, em qualquer que seja a modalidade. Assim, não são menos reais os dados advindos de outros estilos ou da escrita, por exemplo, de modo que, mesmo com os princípios da Sociolinguística tendo sido formulados para a fala vernacular, eles podem ser aplicados em outros cenários.

Nesta pesquisa, a investigação desenvolvida se centra na descrição e na análise de um fenômeno variável em legendas audiovisuais, um gênero textual-discursivo que mistura características das modalidades oral e escrita, resultante de um processo de tradução. Observa-se, portanto, o comportamento da língua em um contexto de maior monitoramento, no qual, entretanto, se busca reproduzir características de uma fala tida como natural (Lopes; Afonso, 2021). Dessa forma, pressupõe-se que a realização do OD anafórico de terceira pessoa seja diretamente influenciada pelas características estruturais e situacionais que organizam o gênero legenda, além, é claro, de outros aspectos linguísticos e extralinguísticos, isto é, pelos fatores condicionantes, apresentados na discussão a seguir.

Sendo a língua dotada de heterogeneidade sistemática, muito da tarefa do sociolinguista diz respeito à tentativa de entender e explicar os processos de variação e mudança linguísticas com os quais se depara na análise de seus dados. Diante dessa necessidade e buscando embasar sua teoria de maneira empírica, Weinreich, Labov e Herzog (2006[1968]) elencam cinco problemas essenciais para essa discussão, a saber: (i) os fatores condicionantes; (ii) a transição; (iii) o encaixamento; (iv) a avaliação; e (v) a implementação.

Sobre o primeiro problema, o dos fatores condicionantes, os três autores afirmam que, em face do conjunto de possibilidades de variação e mudança na ocorrência de um fenômeno variável, é fundamental que se levante, também, quais são as restrições linguísticas e sociais que operam na escolha de uma ou outra variante (Weinreich; Labov; Herzog, 2006[1968]). Essas restrições não se configuram como categóricas, mas podem indicar certa regularidade na realização do fenômeno e/ou apontar para uma tendência geral, sendo necessário, então, que elas sejam controladas, inclusive de maneira estatística. Acerca desse problema, vale mencionar que, anos após os autores proporem essa teoria, o próprio Labov (1982) reconsideraria a maneira como o postulado foi colocado, defendendo que, na realidade, a discussão dos fatores condicionantes deveria estar acoplada à do encaixamento – explorado mais a frente –, posto que as restrições estão diretamente ligadas à maneira como cada mudança integra o sistema linguístico.

Em estudos anteriores sobre a realização do acusativo anafórico de terceira pessoa, por exemplo, alguns condicionamentos linguísticos foram consistentemente apontados como fatores relevantes para a ocorrência do fenômeno, tais quais a natureza do antecedente, o traço semântico do antecedente, a função sintática do antecedente, a forma verbal e a estrutura sintática (cf. Omena, 1978; Duarte, 1986; Cyrino, 1994; Freire, 2005; Marques de Sousa, 2021; Lima, 2022). Ao lado deles, apontou-se, também, a importância de certos fatores sociais para a escolha das variantes, como a escolaridade e a faixa etária (cf. Omena, 1978; Duarte, 1986; Freire, 2005). Mais recentemente, outro fator extralinguístico que tem se provado influenciar na realização do fenômeno é o gênero textual-discursivo no qual cada dado ocorre (Santana; Pissurno; Lima, 2019; Lima, 2022), sendo essa questão também abordada nesta pesquisa.

Já o segundo problema diz respeito à transição, ou seja, à questão de como uma mudança efetivamente acontece no sistema linguístico. Desse modo, observa-se como uma forma linguística, originalmente em variação, se propaga através do tempo, ganhando força e se fixando na língua. Nesse sentido, o processo de mudança passa, obrigatoriamente por três fases. Na primeira, a variação tem início na fala de alguns falantes, por vezes restritos a um grupo específico; em seguida, as diversas variantes coexistem por certo tempo em competição, até que uma delas comece a suplantar o uso da(s) outra(s), atingindo um número maior de falantes; por fim, a nova forma passa a ser a preferida em todos os contextos, em detrimento da(s) antiga(s), que se tornam ultrapassadas.

Ao discutirem o problema do encaixamento, Weinreich, Labov e Herzog (2006[1968]) afirmam que toda mudança deve ser considerada sob dois pontos de vista: o linguístico e o social. Assim, as mudanças devem estar encaixadas tanto no sistema linguístico, interagindo com outros processos linguísticos internos, como na estrutura social, se relacionando com mudanças nesse âmbito. Como consequência, a investigação sociolinguística que lida com esse problema deve não só trabalhar com uma massa de dados expressiva, mas, também, compreender de maneira profunda o contexto social da comunicação, precisando “[...] em que medida e em que grau de intensidade se dá a covariação entre diferenças nos padrões socioculturais e ideológicos e a variabilidade observada no processo de estruturação da língua [...]” (Lucchesi, 2004, p.176).

Para exemplificar esses dois últimos problemas, pode-se mencionar as pesquisas de Cyrino (1990a, 1990b, 1994). Com base nos estudos da autora, sabe-se que a realização do OD anafórico de terceira pessoa é um fenômeno intimamente ligado ao aumento da produtividade do objeto nulo e à queda dos clíticos no PB. Em Cyrino (1994), a autora explica que a utilização da categoria vazia em posição de OD começou a se tornar mais presente primeiramente nos

contextos em que o antecedente era oracional, nos quais, aos poucos, deixa-se de empregar o clítico neutro “o”. A partir disso, o objeto nulo começou a se expandir para contextos em que o antecedente era [-animado], até que seu uso se tornasse generalizado, em todos os contextos, sempre acompanhado de uma diminuição nos índices de uso dos clíticos. Além disso, para a autora, o surgimento do pronome lexical “ele” em posição de OD também estaria conectado ao processo descrito, uma vez que, ao se perder o clítico como uma opção de fato viável, a única alternativa que restava aos falantes era empregar a forma tônica como argumento do verbo.

O quarto problema elencado pelos autores consiste na avaliação, relacionada às atitudes subjetivas e/ou conscientes dos falantes no que diz respeito a formas linguísticas que estão em variação, posto que elas podem ser decisivas na definição do futuro de cada variante. Por exemplo, uma avaliação social positiva pode acelerar a concretização de uma mudança, ao passo que, se uma avaliação social negativa recair fortemente sobre uma forma específica, diminuem as chances de que ela se estabeleça concretamente na língua.

No caso do fenômeno em questão nesta pesquisa, considera-se, em princípio, que a expressão do acusativo anafórico passa por processos de estigmatização social, já que o uso do pronome lexical em posição de objeto, por exemplo, é alvo de comentários sociais que comumente condenam seu uso, sendo combatido pelas gramáticas tradicionais e pelo sistema de ensino. Além disso, como indicam os resultados de Duarte (1986), parece haver um índice de formalidade associado ao uso do clítico, sendo essa forma e o pronome tônico “ele” apenas parcialmente aceitos em construções específicas, em oposição ao objeto nulo, que recebe ampla aceitação, até mesmo na opinião de professores, segundo a autora.

Por fim, a implementação, último problema levantado no texto de 1968, faz referência à tentativa de desvendar como e por que uma mudança linguística começa, em épocas e localidades específicas. A partir de um ponto de vista global, então, o que se quer entender é, em última análise, a maneira como o sistema linguístico de uma certa comunidade se altera no decorrer do tempo, levando em consideração todos os aspectos que afetam esse percurso. Por isso, diante dos inúmeros fatores envolvidos em um processo complexo como o aqui abordado, o problema da implementação é, talvez, o mais difícil de ser plenamente resolvido.

Por uma questão de escopo, o trabalho aqui proposto não discutirá todos os cinco problemas listados por Weinreich, Labov e Herzog (2006[1968]). Nele, o foco recairá sobre a questão dos fatores condicionantes, dado que o interesse maior da pesquisa é compreender quais são as restrições linguísticas e extralinguísticas – estas últimas referentes a características estruturais e situacionais do gênero textual-discursivo sob investigação – que se aplicam de maneira mais expressiva na realização do OD anafórico de terceira pessoa. Apesar disso, pode

ser que, caso seja necessário, os outros problemas sejam mencionados para sugerir explicações plausíveis para os resultados encontrados.

Ponto importante a se considerar na discussão sobre a Sociolinguística Variacionista é também o eixo no qual a variação ocorre. Com isso em mente, a primeira dimensão que deve ser observada é a temporal, na qual podem ser encontradas variedades diacrônicas ou sincrônicas (Cosieriu, 1981). As primeiras correspondem àquelas que se distribuem ao longo do tempo, fazendo referência ao estado da língua em sincronias passadas. Já as segundas dizem respeito à língua no momento presente, podendo se distribuir em outros quatro eixos: o espacial, o social, o do estilo e o do canal.

Como explicado por Coelho *et al.* (2015), no eixo espacial, encontra-se a variação diatópica, também conhecida como regional ou geográfica, na qual o espaço geográfico se configura como um fator condicionante. Já no social, a variação é chamada de diastrática, e pode, de acordo com Castilho (2020[2010]), ser dividida entre variação sociocultural e individual. A sociocultural é relacionada à camada social à qual o falante é vinculado, sendo sistematizada a partir do nível socioeconômico; por outro lado, a individual é associada a características pessoais do falante, tais como sexo/gênero, faixa etária e registro. Quanto ao eixo do estilo, o nome que se dá à variação é diafásica ou estilística, estando ela ligada a questões de adequação de registro, em situações formais ou informais. Fechando essa tipologia, encontra-se, ainda, a variação diamésica, chamada de variação de canal por Castilho (2020[2010]), conectada às diferenças no emprego de formas linguísticas nas modalidades oral e escrita da língua. Nesse caso, as características da situação comunicativa são importantes, posto que, de acordo com o autor, “[...] o falante não está sozinho na construção de seus enunciados, que são de certa forma controlados pelo interlocutor, presente ou ausente [...]” (Castilho, 2020[2010], p. 12).

Como já foi mencionado anteriormente (cf. seção 1), esta pesquisa se propõe a investigar, dentre outras coisas, como se dá a relação entre fala e escrita nas legendas, por se tratar de um gênero escrito que mimetiza, em certo nível, a oralidade, assim como a possível relação existente entre a formalidade e as escolhas dos tradutores, especialmente pensando na influência das gramáticas tradicionais. Dessa maneira, não serão aqui abordados todos os tipos de variação acima descritos, mas, sim, questões relacionadas principalmente à variação diafásica e à diamésica.

Independentemente do tipo de variação encontrada pelo pesquisador, as formas variantes de um fenômeno podem ser vistas, também, a partir de outra categorização laboviana,

relacionada ao valor social a elas atribuído. Trata-se da divisão entre indicadores, marcadores e estereótipos, que será aprofundada a seguir.

Para Labov (2008[1972]), os indicadores consistem em traços linguísticos relacionados à matriz social, podendo estar relacionados a fatores como origem, idade e nível socioeconômico. Entretanto, esses traços não são sensíveis a alternâncias estilísticas e não recebem qualquer julgamento social, de modo que ficam limitados a indicar a diversificação social. Já os marcadores se apresentam como traços que, além de serem associados às categorias sociais, também são sensíveis à dimensão estilística da língua. Podem, portanto, marcar variantes que estão sujeitas a processos de avaliação subjetiva, apesar de estarem abaixo do nível da consciência. Por último, encontram-se os estereótipos, que correspondem a formas que ascendem ao nível da consciência sociolinguística, tornando-se “[...] socialmente marcadas e rotuladas pela sociedade [...]” (Labov, 2008[1972], p. 287). Por isso, algumas dessas variantes passam por uma forte estigmatização social, o que pode levar a alterações no processo de mudança, comumente em direção à extinção da forma condenada.

Na realização do acusativo anafórico, objeto deste estudo, supõe-se que três das quatro variantes em jogo⁹ se caracterizam como marcadores: o SN anafórico, o clítico acusativo e o objeto nulo, dado que essas formas parecem apresentar diferenciação estilística, sendo empregadas em maior ou menor número a depender do nível de formalidade associado ao contexto de uso. Em contrapartida, o pronome lexical se encaixa na categoria de estereótipo, uma vez que é comumente apontado como um “erro de português”, além de ser consistentemente rechaçado em testes de percepção, havendo sobre ele um forte estigma social (Duarte, 1986).

Apesar de, até agora, terem sido discutidos os fundamentos teóricos da Sociolinguística, parte da potencialidade dessa área é prover, também, um aparato metodológico para o estudo da variação e da mudança linguísticas. Nesse contexto, desde os primeiros estudos da área, há uma preocupação em definir critérios objetivos para a composição de amostras, a fim de que elas sejam tanto representativas das comunidades de fala sob investigação como suficientes para a análise do(s) fenômeno(s). Assim, para definir o número de falantes necessários para uma pesquisa acerca da língua falada, devem ser levados em conta itens como o grau de variabilidade e a natureza do fenômeno, a quantidade de variáveis extralinguísticas e o número de níveis em cada uma delas, entre outros, conforme indicado, por exemplo, em Tarallo (1986).

⁹ As variantes que compõem a variável dependente desta pesquisa são descritas em mais detalhes na seção 5, referente à metodologia.

Já para estudos que trabalhem com a língua em sua modalidade escrita, assim como o aqui proposto, também se tem reafirmado a importância de se aplicar uma metodologia bem definida durante a composição dos *corpora*. Com vistas a compor amostras homogêneas e comparáveis, tornam-se importantes, portanto, critérios como número de palavras, recorte temporal, gênero textual-discursivo, entre outros aspectos (cf. Freire, 2005; Biazolli, 2016; Vieira; Lima, 2019; Lima, 2022; Vieira, em andamento).

Outro avanço trazido pela Teoria da Variação e Mudança Linguísticas diz respeito ao tratamento quantitativo dos dados, coletados a partir dessas amostras criteriosamente elaboradas, posto que é ele que permite ao sociolinguista verificar quais estruturas são favorecidas ou não em certo contexto. Com efeito,

[...] A realização de análises quantitativas possibilita o estudo da variação linguística, permitindo ao pesquisador apreender sua sistematicidade, seu encaixamento linguístico e social e sua eventual relação com a mudança linguística. [...] Antes do advento da metodologia de quantificação, a variação linguística era considerada secundária, aleatória ou mesmo impossível de ser cientificamente apreendida. O uso de métodos estatísticos, contudo, tem permitido demonstrar o quão central a variação pode ser para o entendimento de questões como identidade, solidariedade ao grupo social, comunidade de fala, prestígio e estigma, entre tantas outras. [...] (Guy; Zilles, 2007, p. 73).

Através desses métodos, são realizadas comumente análises uni e multivariadas, a fim de averiguar quais são as variáveis que condicionam a escolha de uma ou outra variante. Para tanto, é comum a utilização de programas estatísticos como o *Goldvarb-X* (Sankoff *et al.*, 2005) ou da plataforma *R* (R Core Team, 2023), escolhida para a realização das análises deste estudo (cf. seção 5).

Por fim, para encerrar as discussões teóricas desta subseção, será brevemente retomado o debate entre Lavandera e Labov acerca da variação sintática, uma vez que, nesta pesquisa, o foco é a descrição da ocorrência de um fenômeno morfossintático nas legendas audiovisuais.

Como já explicitado anteriormente, Labov (2008[1972]) defende a ideia de que todos os tipos de variação, nos mais diversos níveis linguísticos, podem e devem ser investigados, com o objetivo de se atingir uma descrição cada vez mais acurada da língua. Contudo, os primeiros estudos inseridos na linha da Sociolinguística Variacionista voltaram-se majoritariamente ao plano fonético-fonológico, dado que variações nesse nível eram bem perceptíveis e poderiam estar ligadas, também, às dimensões lexical e gramatical da língua. Desses primeiros trabalhos, emergiu a noção de variante, estando ela conectada à ideia de formas alternantes de se dizer a mesma coisa.

A partir dessa definição, Lavandera (1978, 1984) questiona fortemente a viabilidade de se estudar a variação em outros níveis que não o fonológico, sendo que, em sua visão, construções sintáticas como as apresentadas nos exemplos (5) a (8), retirados do *corpus*, obrigatoriamente teriam significados diferentes associados a elas, de modo que elas não poderiam ser consideradas variantes de um mesmo fenômeno.

- (5) CP1: Mas pode dizer para ela assim mesmo?

Pode encontrá-*la*
e dizer para ela que eu *a* amei?

Eu *a* amei tanto. (PT06E24)

- (6) CA1: Quando eu tinha 10 anos,
estava em um restaurante

com meu pai,
comendo waffles.

Eu adorava *waffles*. (FT10E24)

- (7) A1: - Ele foi descontaminado?
PM5: - Não.

A1: - Descontamine [Ø] agora!

PM5: - Pode deixar! (PT10E24)

- (8) GC1: - Pare de operá-lo.

Deixe *ele* morrer.

Deixe-o aí para morrer. (FT06E24)

Diante desse argumento, a autora propõe a substituição da equivalência semântica pela comparabilidade funcional, ou seja, o abandono da ideia de que as variantes teriam o mesmo significado em favor da noção de que elas teriam a mesma intenção comunicativa. Nesse panorama, a própria variação exprimiria, então, uma opção do falante, que teria propósitos comunicativos em mente ao proferir uma ou outra variante.

Como resposta à crítica, Labov (1978) e Weiner e Labov (1983) propõem então um refinamento do conceito de significado, que passa a ser entendido de maneira mais restrita. Para os autores, a equivalência no caso de variantes de fenômenos sintáticos se dá não por elas possuírem o mesmo significado semântico, mas, sim, o mesmo significado referencial, apresentando, portanto, o mesmo valor de verdade.

Mesmo que se assuma tal posição, é necessário reconhecer que ainda há contextos que colocam em xeque a possibilidade de se “dizer a mesma coisa” com variantes morfossintáticas diferentes, uma vez que nem sempre há opcionalidade na escolha, isto é, nem todas as formas consideradas estão, de fato, disponíveis. Com efeito, em uma fala como a reproduzida em (9), por exemplo, dificilmente um falante do PB empregaria uma forma que não a categoria vazia, não só pelo fato de a expressão “não sei” ser relativamente cristalizada na língua, mas, também, porque o uso de outras estratégias poderia ser inviável (9a) ou atingir efeitos discursivos diferentes do que o pretendido – como nos exemplos (9b) e (9c).

(9) T1: -Acha que vai ser um discurso longo?

O1: -Eu não sei [Ø]. (PT06E13)

(9a) T1: -Acha que vai ser um discurso longo?

* O1: -Eu não sei **ele**.

(9b) T1: -Acha que vai ser um discurso longo?

O1: -Eu não **o** sei.

(9c) T1: -Acha que vai ser um discurso longo?

O1: -Eu não sei **isso**.

Além disso, como o material utilizado nesta pesquisa é a legenda, há outra dimensão que ainda deve ser considerada nessa discussão: a da relação entre o original e o traduzido. Por partir, inevitavelmente, de um roteiro, seja ele disponibilizado aos legendistas ou acessado somente pelo contato com as cenas, é essencial questionar o quanto da tradução é resultado de um decalque das estruturas da língua original. A fim de discutir brevemente essa questão, que será retomada mais à frente, observam-se os exemplos (10) e (11).

(10) AZ1: Dra. Bailey, eu entendo **isso**. (FT06E01)

(11) AZ1: Dra. Bailey, eu entendo [Ø]. (PT06E01)

Considerando que a fala original nessa cena era, em inglês, “*Dr. Bailey, I understand that*”, é fácil conjecturar que a tradução amadora apresentada em (10) provavelmente emprega o pronome “isso” por, nesse contexto, ser ele o equivalente em português do demonstrativo “*that*”. Entretanto, como ilustrado pela tradução profissional em (11), nem sempre todas as palavras do original são reproduzidas, mesmo que elas tenham correspondentes adequados na língua de chegada, especialmente quando é possível transmitir a carga semântica da frase de maneira mais breve e confortável para que o leitor faça sua leitura.

De qualquer forma, mesmo que nem todas as variantes possam ser empregadas em todos os contextos, o fato é que a variação em variáveis morfossintáticas seguirá ocorrendo e gerando comentário social ao menos em alguns deles, o que por si só justifica a investigação desse tipo de fenômeno. Nesse sentido, é importante explorar não só os fatores que condicionam esses usos, mas identificar quais cenários funcionam como restrições a uma certa estratégia e o que acontece quando o falante não tem liberdade plena para fazer uma escolha linguística. Assim, a investigação sociolinguística continua sendo possível nesses casos, aplicabilidade que é reforçada pelos inúmeros trabalhos realizados acerca de variáveis sintáticas e morfossintáticas em diversas línguas, inclusive no PB.

Apresentados os pressupostos teórico-metodológicos básicos da Sociolinguística Variacionista, que embasam o estudo do fenômeno variável aqui empreendido, faz-se necessário voltar o olhar, ainda, para outras questões teóricas importantes para esta pesquisa. Assim, a próxima subseção discorrerá sobre diferentes definições de gênero textual-discursivo, a fim de fundamentar as discussões que se seguirão nas próximas seções.

2.2 Gêneros textuais-discursivos: definições e características

É inegável que a noção de gênero textual-discursivo¹⁰, essencial para a descrição da legenda neste trabalho, é, hoje, um dos temas mais explorados por várias áreas do conhecimento, não sendo diferente no campo dos estudos linguísticos. Nesse panorama, diferentes tendências e propostas para a análise dos gêneros surgiram ao longo dos anos, todas influenciadas, em maior ou menor nível, pela definição bakhtiniana (Meurer; Bonini; Motta-Roth, 2005), apresentada a seguir:

¹⁰ Ao final desta subseção, discute-se brevemente a inexactidão terminológica que acompanha o termo “gênero”, assim como os motivos para, nesta pesquisa, adotar-se a combinação “gênero textual-discursivo”.

[...] O emprego da língua efetua-se em forma de enunciados (orais e escritos) concretos e únicos, proferidos pelos integrantes desse ou daquele campo da atividade humana. Esses enunciados refletem as condições específicas e as finalidades de cada referido campo não só por seu conteúdo (temático) e pelo estilo da linguagem, ou seja, pela seleção dos recursos lexicais, fraseológicos e gramaticais da língua, mas, acima de tudo, por sua construção composicional. Todos esses três elementos – o conteúdo temático, o estilo, a construção composicional – estão indissolúvelmente ligados *no conjunto* do enunciado e são igualmente determinados pela especificidade de um campo da comunicação. Evidentemente, cada enunciado particular é individual, mas cada campo de utilização da língua elabora seus *tipos relativamente estáveis* de enunciados, os quais denominamos *gêneros do discurso*. (Bakhtin, 2016[1979], p.11-12).

A fim de destrinchar essa citação, um primeiro ponto que deve ser abordado é o conceito de enunciado que, para o autor, se caracteriza como a unidade comunicativa primordial da língua, representando seu uso concreto. Por ser resultado de uma combinação de fatores históricos, sociais e situacionais que não se repetem, o enunciado é sempre individual e único, podendo se apresentar como o espaço propício para a manifestação do estilo. Entretanto, todos os enunciados são realizados por um indivíduo social atravessado por diversas vozes, de modo que eles apresentam uma natureza ativamente responsiva, ou seja, ocorrem em resposta a enunciados anteriores, pressupondo, também, uma resposta real ou possível do interlocutor. Para Bakhtin (2016[1979], p. 25), então, “[...] cada enunciado é um elo na corrente complexamente organizada de outros enunciados [...]”, de modo que, apesar de único e individual, todo enunciado se liga a outros através da esfera de comunicação discursiva à qual pertence.

Por isso, outro aspecto relevante para a definição proposta pelo autor é o de “campo da atividade humana”¹¹, que aponta para os diversos âmbitos nos quais um indivíduo age, ou seja, o íntimo, o cotidiano, o jornalístico, o do trabalho, o político, o artístico, entre tantos outros. Com efeito, para Bakhtin (2016[1979]), seriam inesgotáveis as possibilidades da atividade humana, uma vez que cada campo, ao longo do tempo, vai se desenvolvendo e tornando-se cada vez mais complexo. Nesse contexto, se cada campo elabora seus próprios gêneros, seriam estes também inúmeros, adaptando-se tanto às tradições anteriormente estabelecidas como às novas especificidades que podem surgir em cada atividade.

¹¹ Anteriormente, o termo “campo da atividade humana” foi também traduzido como “esfera da atividade humana”. Aqui, opta-se por empregar a primeira alternativa, já que é ela a utilizada na tradução mais recente do texto de Bakhtin (2016[1979]), consultada para a realização desta pesquisa.

Explorados brevemente esses dois conceitos fundamentais, pode-se passar, agora, à discussão do tripé bakhtiniano que opera na definição de gênero do discurso. Partindo da visão do enunciado como elo na cadeia de comunicação discursiva, Bakhtin (2016[1979], p. 46) afirma que ele manifesta a posição do falante em diferentes campos do objeto e do sentido, de forma que “[...] cada enunciado se caracteriza, antes de tudo, por certo conteúdo semântico-objetual [...]”. Esse conteúdo, também chamado de tema, seria o primeiro elemento do enunciado a determinar suas especificidades estilísticas e composicionais, por emanar diretamente da intenção discursiva do falante. Trata-se, portanto, de uma maneira de relacionar o material linguístico à experiência humana, constituída na interação verbal e resultante, também, dos inúmeros enunciados precedentes. A partir do recorte do gênero, portanto, “[...] o tema dá conta do modo como recorrentemente as pessoas têm falado sobre certos assuntos [...]” (Alves Filho, 2011, p. 45), configurando-se como um comportamento tipificado.

Já o segundo ponto do tripé, o estilo, é relacionado por Bakhtin (2016[1979]) à dimensão puramente linguística do gênero, fazendo referência aos meios linguísticos, às formas gramaticais, às palavras selecionadas para um certo enunciado. Para o autor, que cunha a máxima de que “[...] onde há estilo há gênero [...]” (Bakhtin, 2016[1979], p. 21), tal aspecto é determinado diretamente pelo tema do enunciado e pela relação do falante com o interlocutor e seus enunciados, de maneira que aspectos como o grau de proximidade entre os dois e a resposta antecipada são fatores essenciais na construção da comunicação. Além disso, o estudioso considera ser possível a materialização de um estilo individual nos enunciados, que podem, portanto, refletir marcas idiossincráticas do falante, apesar de nem todo gênero discursivo se apresentar como propício a essa manifestação. Dada a especificidade do estilo em cada tipo de enunciado, Bakhtin (2016[1979], p. 21) afirma, ainda, que alterações no estilo podem alterar não só esse parâmetro, mas levar a uma destruição ou renovação do gênero. De forma geral, pode-se dizer, então, que

[...] O estilo é indissociável de determinadas unidades temática e – o que é de especial importância – de determinadas unidades temáticas composicionais: de determinados tipos de construção do conjunto, de tipos do seu acabamento, de tipos da relação do falante com outros participantes da comunicação discursiva – com os ouvintes, os leitores, os parceiros, o discurso do outro, etc (Bakhtin, 2016[1979], p. 18).

Fechando a tríade bakhtiniana, encontra-se, por fim, a construção composicional, associada à estrutura comumente utilizada em um gênero. Sobre isso, Bakhtin (2016[1979], p. 38) afirma que “[...] todos os nossos enunciados têm *formas* relativamente estáveis e típicas de

construção do conjunto [...]”, isto é, todos os enunciados são organizados a partir do que anteriormente foi dito (ou escrito) no enquadre de um gênero, apontando para o acabamento geral de um texto. Dessa maneira, pode-se dizer que, de certo modo, “[...] a forma composicional seria o esquema geral do texto, assim como sua estruturação textual em partes [...]” (Maciel, 2015, p. 254), sendo esse elemento, portanto, indissociável do plano temático do enunciado, uma vez que cada tema demandará disposições textuais específicas.

Todos esses aspectos atuam de maneira conjunta na constituição de um gênero discursivo, sendo eles interdependentes um do outro. Desse modo, a seleção por um gênero específico não é neutra e nem pode ser feita levando em conta apenas um dos três itens mencionados. Tal como afirma Bakhtin (2016[1979], p. 37-38),

[...] essa escolha é determinada pela especificidade de um dado campo da comunicação discursiva, por considerações semântico-objetais (temáticas), pela situação concreta da comunicação discursiva, pela composição pessoal dos seus participantes, etc. Em seguida, a intenção discursiva do falante, com toda a sua individualidade e subjetividade, é aplicada e adaptada ao gênero escolhido, constitui-se e desenvolve-se em determinada forma de gênero [...].

Apresentados os três componentes essenciais de um gênero, vale pensar, também, na maneira como esses tipos relativamente estáveis de enunciados circulam na sociedade e se relacionam com a língua. Considerando que toda comunicação humana ocorre por meio de gêneros materializados em textos, é primeiramente importante reconhecer que as a própria língua é apresentada ao falante através dos gêneros, estando eles, assim, diretamente vinculados à experiência linguística. Nesse sentido, Bakhtin (2016[1979]) afirma que, desde o início, o falante recebe não só as formas da língua que são impostas a ele, mas, também, os gêneros do discurso, que são igualmente obrigatórios para que a compreensão entre os falantes ocorra de maneira eficiente.

Percebe-se, portanto, que os gêneros devem desfrutar de uma posição central nos estudos linguísticos, assim como é aqui pretendido, posto que é apenas por meio dele que a comunicação e a interação podem ocorrer.

Concepção marcada também por questões sociointeracionais é a de Marcuschi (2005, 2008, 2010[2001]), que, de maneira geral, concebe os gêneros textuais como “[...] formas textuais escritas ou orais bastante estáveis, histórica e socialmente situadas [...]” (Marcuschi, 2008, p. 155). Assim como na proposta bakhtiniana, o autor também conceitua gêneros através de três conceitos principais – as composições funcionais, os objetivos enunciativos e os estilos –, que se realizam de maneira integrada com diversas forças que regem a sociedade. São, assim,

[...] fenômenos históricos, profundamente vinculados à vida cultural e social. Fruto de trabalho coletivo, os gêneros contribuem para ordenar e estabilizar as atividades comunicativas do dia-a-dia. São entidades sócio-discursivas e formas de ação social incontornáveis em qualquer situação comunicativa [...] (Marcuschi, 2005, p. 19).

Para o autor, os gêneros estão absolutamente incorporados nas culturas das quais fazem parte, definindo-se muito mais pela maneira como funcionam nas esferas comunicativas, cognitivas e institucionais do que por traços formais relacionados à sua dimensão linguística. Nessa perspectiva, utilizar um gênero textual é, portanto, muito mais do que produzir um texto como o esperado; é, sobretudo, um modo de agir sobre o mundo e sobre o outro, com propósitos comunicativos específicos. Ao dominar um gênero, então, “[...] não dominamos uma forma linguística e sim uma forma de realizar linguisticamente objetivos específicos em situações sociais particulares [...]” (Marcuschi, 2008, p. 154).

Essas situações sociais acontecem, de acordo com o estudioso, sempre no âmbito de domínios discursivos – conceito paralelo ao de “campos de atividade humana” –, que correspondem às esferas da vida em sociedade cujas práticas projetam formas de comunicação específicas. Trata-se, portanto, de instâncias discursivas como a instrucional, a jornalística, a religiosa, e a ficcional¹², que produzem modelos comunicativos estáveis, elaborados a partir de práticas sociais distintas e que, em última análise, dão origem a vários gêneros textuais. Nesse cenário, “[...] os domínios discursivos operam como enquadres globais de superordenação comunicativa, subordinando práticas sociodiscursivas orais e escritas que resultam nos gêneros [...]” (Marcuschi, 2008, p. 194).

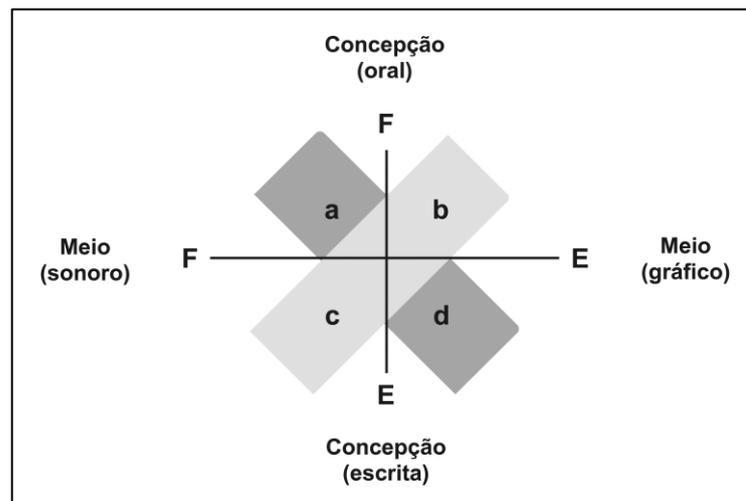
Apesar de as considerações acima serem importantes e terem a sua relevância para o presente estudo, a discussão mais interessante da proposta de Marcuschi (2008, 2010[2001]) é, sem dúvida, o debate acerca da materialização dos gêneros. Sobre isso, o linguista defende a ideia, também levantada por outros estudiosos (cf. Chafe, 1982, 1985; Biber, 1988; e Bortoni-Ricardo, 2004, 2005, 2012, para citar alguns), de que as diferenças entre as modalidades oral e escrita da língua não pressupõem uma divisão dicotômica entre elas, mas, sim, a existência de um *continuum* tipológico que vai de um polo a outro, abarcando todas as práticas de produção e de recepção de um texto. Por isso, ao passo que nos diversos domínios discursivos há gêneros que se constroem inteiramente em uma modalidade, é perfeitamente possível que outros

¹² Em uma lista não exaustiva, Marcuschi (2008, p. 195-196) menciona os seguintes domínios discursivos: instrucional, jornalístico, religioso, de saúde, comercial, industrial, jurídico, publicitário, de lazer, interpessoal, militar e ficcional.

apresentem entrecruzamentos entre fala e escrita, resultantes, principalmente, de contextos de produção e de recepção distintos.

Com vistas a descrever de forma mais segura estes últimos gêneros textuais, chamados de híbridos ou mistos, Marcuschi (2008, 2010[2001]) introduz dois termos: concepção e meio. O primeiro, pautado na distinção entre oral e escrito, diz respeito ao modo como um texto é concebido e planejado, estando ligado à produção; o segundo, organizado com base nas ideias de sonoro e gráfico, faz referência à via pela qual um material é comunicado, associando-se, portanto, à recepção. Tal postulado é representado pelo esquema na Figura 3 a seguir.

Figura 3 - Representação dos eixos da concepção e do meio, de acordo com Marcuschi (2010[2001])



Fonte: Adaptado¹³ de Marcuschi (2010, p. 39)

Como pode ser observado acima, os dois eixos propostos por Marcuschi (2008, 2010[2001]) se cruzam no plano dos gêneros, resultando em quatro combinações possíveis. De maneira geral, têm-se, primeiramente, que a fala é de concepção oral e se realiza no meio sonoro, enquanto a escrita é de concepção, naturalmente, escrita e toma forma através do meio gráfico. Assim, gêneros textuais que acumulam essas características, pertencentes aos domínios (a) e (d) na figura, podem ser considerados prototípicos de uma modalidade. É o caso de uma conversa cotidiana (a), concebida e realizada na modalidade oral, e de uma dissertação como esta (d), que, obviamente, representa um gênero ligado à modalidade escrita. Por outro lado, os gêneros que ocupam as áreas demarcadas por (b) e (c) consistem em domínios híbridos, nos quais se apresentam características de modalidades distintas. Gêneros nesses domínios podem

¹³ Neste trabalho, são indicados como adaptados os materiais que sofreram algum tipo de modificação, mesmo que apenas de formatação, como é o caso da Figura 3.

ser de concepção oral e meio gráfico (b), como um bilhete, ou de concepção escrita e meio sonoro (c), como as notícias televisivas.

Outra contribuição que Marcuschi (2005, 2009) traz, ainda que de maneira mais tímida, corresponde à discussão acerca dos gêneros emergentes. Surgidos a partir do uso cada vez mais intenso das novas tecnologias criadas nas últimas décadas, tais gêneros, comuns no meio digital ou eletrônico, são marcados por características próprias, devendo ser analisados também de maneira singular. Nessa empreitada, o autor propõe, por exemplo, um conjunto de doze parâmetros que englobam aspectos como tempo, participantes, composição, tema e estilo, cada qual com seus níveis, a fim de dar conta da complexidade que ronda esse tipo de produção¹⁴. Aspecto central e comum a todos esses gêneros, porém, é o fato de suas formas comunicativas serem dotadas de um hibridismo inerente, que coloca em xeque as relações entre fala e escrita, além de todos eles permitirem “[...] observar a maior integração entre os vários tipos de semioses, signos verbais, sons, imagens e formas em movimento [...]” (Marcuschi, 2005, p. 20).

O estudo dessa integração entre os tipos de semiose só foi se desenvolver de maneira mais aprofundada, porém, com o advento de pesquisas sobre a multimodalidade (cf. Dionisio, 2007 e Rojo; Barbosa, 2015, por exemplo), que abandonam o paradigma dual da fala e da escrita em favor de uma visão na qual diversas semioses entram em jogo na comunicação. Nesse contexto, as interações comunicativas efetuadas pelos falantes mobilizam, a um mesmo tempo, tanto recursos verbais (escritos ou orais) como visuais (estáticos ou dinâmicos), assim como no material aqui abordado, de modo que os gêneros adquirem um caráter intrinsecamente multimodal. Assim, um gênero multimodal ou multissemiótico se caracteriza por ser

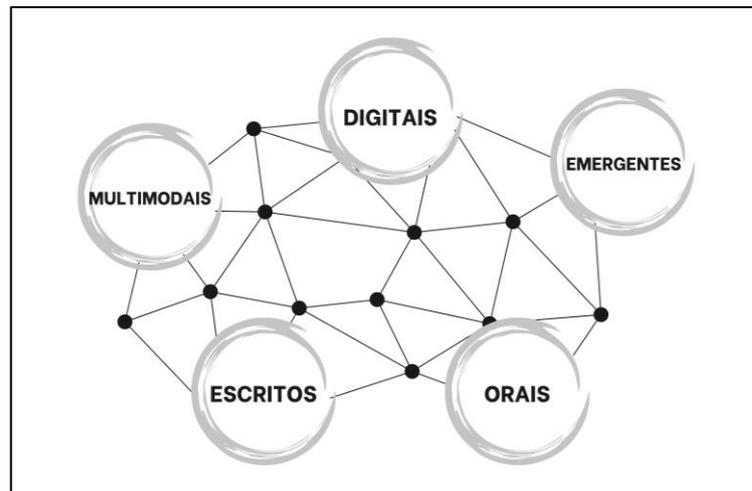
[...] aquele que recorre a mais de uma modalidade de linguagem ou a mais de um sistema de signos ou símbolos (semiose) em sua composição. Língua oral ou escrita (modalidade verbal), linguagem corporal (gestualidade, danças, performances, vestimentas – modalidade gestual), áudio (música e outros sons não verbais – modalidade sonora) e imagens estáticas e em movimento (fotos, ilustrações, grafismos, vídeos, animações – modalidades visuais) compõem hoje os textos da contemporaneidade, tanto em veículos impressos como, principalmente, nas mídias analógicas e digitais (Rojo; Barbosa, 2015, p. 108).

Da mesma maneira que têm se desfeito cada vez mais as barreiras entre diferentes sistemas semióticos, têm se dissipado, também, os limites entre os diferentes tipos de gênero,

¹⁴ Os doze parâmetros listados por Marcuschi (2009, p. 41-42) são: relação temporal, duração, extensão do texto, formato textual, participantes, relação dos participantes, troca de falantes, função, tema, estilo, canal/semioses e recuperação de mensagem. Para detalhes sobre os níveis de cada parâmetro e uma breve análise de gêneros textuais com base na proposta, ver Marcuschi (2009).

de modo que suas definições se relacionam, se aproximam e, por vezes, se sobrepõem. Gêneros digitais, por exemplo, tendem a ser multimodais por natureza; gêneros emergentes podem tomar forma tanto escrita como oral; gêneros escritos podem se tornar multimodais com a simples adição de uma imagem. O sistema que se caracteriza no âmbito dos gêneros, portanto, não é um em que eles podem ser agrupados de maneira estanque, sendo necessário considerá-los no interior da intrincada malha que os conecta (Figura 4).

Figura 4 - Representação das conexões entre diferentes classificações de gêneros



Fonte: Elaborada pela autora

Outra proposta que também se faz relevante para a pesquisa aqui descrita é a de Biber e Conrad (2009) que, ecoando outros autores, reconhecem o gênero como um aspecto fundamental da linguagem humana. Na visão dos estudiosos, os gêneros sempre cumprem um propósito comunicativo que depende tanto da cultura na qual a interação acontece como das condições específicas de produção e de recepção de um gênero. Ao analisar um gênero, então, não basta voltar o olhar apenas para os traços linguísticos empregados – características lexicais e gramaticais típicas –; deve-se, também, observar em detalhes o contexto situacional em que eles tomam forma. A partir disso, os pesquisadores elaboraram um modelo para caracterização e comparação de gêneros, cujas características situacionais podem ser observadas na Tabela 1:

Tabela 1 - Características situacionais para análise dos gêneros

Características situacionais	
I. Participantes	
A. Emissor(es) (falante ou autor)	1. único/mais de um/institucional/não identificado 2. características sociais: idade, escolaridade, profissão etc.
B. Destinatário	1. único/mais de um/não enumerado 2. si mesmo/outro
C. Há espectadores?	
II. Relações entre participantes	
A. Interatividade	
B. Papéis sociais	<i>status</i> relativo ou poder
C. Relação pessoal	amigos, colegas, estranhos etc.
D. Conhecimento compartilhado	pessoal e especializado
III. Canal	
A. Modalidade	fala/escrita/sinalização
B. Meio específico	Permanente: gravado/transcrito/impresso/escrito à mão/e-mail/etc. Discurso transitório: face a face/telefone/rádio/TV/etc.
IV. Circunstâncias de produção	
	em tempo real/planejada/roteirizada/revisada e editada
V. Cenário	
A. O tempo e lugar da comunicação são compartilhados pelos participantes?	
B. Lugar da comunicação	1. privado/público 2. cenário específico
C. Tempo	contemporâneo, período histórico
VI. Propósitos comunicativos	
A. Propósitos gerais	narrar/relatar, descrever, expor/informar/explicar, persuadir, ensinar como fazer/procedimentos, entreter, edificar, revelar-se
B. Propósitos específicos	resumir informação de várias fontes, descrever métodos, apresentar novos resultados de pesquisa, ensinar moral através da história pessoa
C. Factualidade	factual, opinião, especulativo, imaginativo

D. Expressão de posicionamento	epistêmico, pessoal, sem posicionamento aparente
--------------------------------	--

VII. Tópico

A. Domínio do tópico geral	doméstico, atividades diárias, comercial/local de trabalho, ciência, educação/acadêmico, governamental/legal/política, religião, esporte, arte/entretenimento etc.
----------------------------	--

B. Tópico específico

C. *Status* social da pessoa mencionada

Fonte: Adaptado de Biber e Conrad (2009, p. 40)

Dada a importância desses aspectos não só para a teoria proposta pelos autores, mas para a presente pesquisa, já que ela se propõe a fazer uma caracterização do gênero legenda, segue-se, agora, uma breve descrição de cada um dos fatores contextuais levantados por Biber e Conrad (2009).

O primeiro deles, os participantes, é utilizado na análise do(s) indivíduo(s) envolvidos na produção e na recepção dos textos em discussão, sendo possível dividi-los, a princípio, entre emissores e destinatários. Em alguns gêneros, tem-se a possibilidade de um texto ser produzido por uma ou mais pessoas, por uma instituição ou, então, a um indivíduo não identificado. Nos orais, é comum que a identificação desses indivíduos seja realizada prontamente, ao passo que, no caso dos textos escritos, isso nem sempre é possível. Quanto aos destinatários, pode-se endereçar um texto a apenas uma pessoa, a vários sujeitos ou a um grupo não enumerado de indivíduos. No caso de conteúdos audiovisuais, por exemplo, o público-alvo é estipulado pelos produtores, não sendo possível, porém, uma determinação exata de quem serão os consumidores finais do texto. Sobre os participantes, Biber e Conrad (2009) ainda mencionam como ponto importante a presença ou não de espectadores, isto é, pessoas que observam a elaboração do gênero, mas não se caracterizam como destinatários finais.

Em seguida, é importante pensar na relação entre os participantes, a fim de avaliar como aspectos relacionados a essa questão interferem na constituição do gênero e nas escolhas linguísticas empregadas. Pensando na interatividade, existem gêneros que permitem contato direto entre emissor e destinatário, havendo possibilidade de resposta direta e, por vezes, imediata. No outro polo, porém, há gêneros em que é impossível a realização de um diálogo entre os participantes, principalmente em casos nos quais sequer a identificação dos indivíduos pode ser efetuada. Importantes nesse contexto são, também, os papéis sociais e as relações pessoais estabelecidas entre os sujeitos envolvidos na produção e na recepção do gênero, uma

vez que relações de poder e/ou distanciamento, por exemplo, podem acarretar escolhas linguísticas distintas daqueles feitas em outros contextos. Ademais, é pertinente observar o conhecimento compartilhado entre emissor e destinatário, principalmente por esse elemento ter o poder de influenciar a escolha do tema e, sobretudo, os usos linguísticos. Se os dois participantes, por exemplo, compartilham conhecimento específico acerca de um tópico, menos informações precisam ser preenchidas concretamente para que a interação seja eficiente.

Já a terceira característica diz respeito à materialização desse gênero, sendo importante distinguir modalidade e meio específico. No âmbito da modalidade, Biber e Conrad (2009) destacam a possibilidade de a comunicação ocorrer de maneira falada, escrita ou sinalizada. Obviamente, como já foi exposto acima, as diferenças entre as modalidades são essenciais para a compreensão de um gênero, posto que elas se entrelaçam com outros elementos e afetam diretamente a maneira como a língua será empregada. No que diz respeito ao meio específico, mencionado como categoria pelos autores, destaca-se a separação feita entre meios permanentes e discursos transitórios, que pode ter claros efeitos sobre a elaboração do gênero. Por exemplo, em um romance impresso, é sempre permitido ao leitor reler trechos importantes ou que não ficaram claros em uma primeira leitura; em um filme exibido nos cinemas, por outro lado, a informação é veiculada apenas uma vez, de modo que a linguagem utilizada deve ser mais prontamente acessível.

No caso das condições de produção, Biber e Conrad (2009) agrupam parâmetros relacionados ao planejamento envolvido em cada gênero. Textos produzidos em tempo real, como uma conversa casual, recebem pouco ou nenhum planejamento devido à dinamicidade da interação, sendo, em geral, menos monitorados. Situação contrária é a encontrada no caso de textos como os do domínio acadêmico, que, além de planejados, passam por exaustivos processos de revisão, pressupondo mais monitoramento e consciência acerca de sua estruturação. Trata-se, portanto, de um critério importante, que pode dar valiosos indícios para a análise da língua empregada em cada gênero.

Por sua vez, a quinta característica elencada pelos autores faz referência à dimensão física da comunicação, considerando o tempo e o lugar em que ela acontece. Nesse contexto, o primeiro item a ser observado é se os participantes compartilham o tempo e o lugar ao longo da comunicação, já que isso está diretamente relacionado à interação e às possíveis marcas ligadas a ela. Além disso, cada uma dessas dimensões pode também contribuir à sua maneira para a construção do gênero. Um bilhete de amor é logicamente escrito de maneira completamente diversa de um discurso político, e parte dessa diferença emerge, justamente, do cenário no qual cada um é veiculado.

Igualmente indispensável é levar em conta o porquê uma comunicação ocorre, ou seja, seu propósito comunicativo. Ligado diretamente à função que um gênero assume em seu contexto de produção, o propósito comunicativo é geralmente dividido em dois níveis, o geral e o específico, sendo que, para cada um deles, diversas são as possibilidades. Os autores propõem, adicionalmente, a avaliação do tipo de informação veiculada e a existência ou não de posicionamento assumido no texto, a fim de pensar as relações estabelecidas entre esses aspectos e o gênero. Neste ponto, é importante frisar, porém, que a lista elaborada por Biber e Conrad (2009) e reproduzida acima não é exaustiva, podendo ser expandida frente à diversidade e à plasticidade que são marcas de estudos sobre os gêneros.

Como última característica, os autores mencionam, por fim, o tópico do texto, que pode também ser dividido em geral e específico. Sob o rótulo de tópicos gerais, encontram-se temas recorrentes de diversas áreas do cotidiano que, por sua vez, englobam assuntos mais particulares, relacionados às especificidades de cada comunicação. Assim como no caso dos propósitos comunicativos, a relação de tópicos proposta, especialmente dos específicos, se caracteriza como uma listagem aberta, adaptando-se aos aspectos singulares de cada gênero.

Para finalizar esta subseção, é necessário, ainda, discutir um problema essencial para a área, debatido por Rojo (2005): a inexactidão terminológica quanto ao termo gênero, uma vez que, a depender da perspectiva teórica assumida, são diferentes os complementos que o acompanham. Bakhtin (2016[1979]), focado nos aspectos discursivos relacionados à enunciação, se refere, sempre, a “gêneros discursivos” ou “gêneros do discurso”. Por sua vez, Marcuschi (2005, 2008, 2010[2001]) utiliza a combinação “gênero textual” em seus trabalhos, elaborados com preocupações mais claras sobre a materialização do texto. Entretanto, como pode ser observado através das considerações aqui trazidas, as ideias de ambos os autores se aproximam em diversos pontos, caracterizando-se, de certa maneira, como correlatas. Nesta seção, procurou-se manter a terminologia utilizada por cada autor ao apresentar suas discussões, justamente para marcar seus posicionamentos. No restante desta pesquisa, porém, opta-se por utilizar o termo “gênero textual-discursivo”, por se considerar o gênero a partir de ambas as dimensões mencionadas, observando a complementariedade entre elas.

2.3 Trabalhando na interface: estudos variacionistas e gêneros textuais-discursivos

Ao longo dos anos, diversas pesquisas variacionistas se dedicaram ao estudo de fenômenos linguísticos dos mais variados níveis, produzindo uma descrição cada vez mais abrangente e detalhada do PB. Entretanto, foi apenas em anos mais recentes que certos estudos passaram a demonstrar um interesse maior em investigar a variação e a mudança linguísticas

de maneira integrada às características dos gêneros textuais-discursivos nos quais cada forma linguística está inserida, tornando-se produtivo apresentá-los a seguir.

Um desses estudos é o de Biazolli (2010), que, ao abordar a colocação pronominal em jornais de São Paulo e de Rio Claro entre os anos de 1880 e 1920, supõe haver diferenças nas taxas de uso das variantes analisadas a depender do gênero na qual elas ocorrem e, por isso, considera o gênero textual-discursivo uma das variáveis independentes da pesquisa. Nesse trabalho, a autora propõe que a ideia de diferentes estilos contextuais de Labov (1966, 2008[1972]), presentes na entrevista sociolinguística, seja aplicada a textos escritos, sugerindo que os gêneros sejam organizados de maneira hierárquica em um *continuum* estilístico.

Nesse sentido, a proposta da estudiosa se aproxima do que foi primeiramente postulado por Romaine (2009[1982]), quando a autora norte-americana defende ser valiosa a observação de diferentes tipos de texto, dispostos em um *continuum* estilístico, em pesquisas focadas na dinamicidade de diferentes línguas. Em um dos extremos, tal *continuum* apresentaria aspectos associados a mais monitoramento e a mais formalidade, além de temas mais complexos; no outro, contaria com a presença de textos menos monitorados e menos formais, geralmente acerca de temas mais cotidianos. Pensando nessas mesmas questões, seria então possível distribuir as variantes linguísticas a partir de uma correlação com os gêneros textuais-discursivos alocados nessa escala.

A partir de sua abordagem, Biazolli (2010) observa, de maneira geral, que a alternância no uso dos clíticos pronominais é sensível ao gênero, em especial no contexto de lexias verbais simples e em ocorrências obtidas de jornais paulistanos. Em gêneros considerados mais subjetivos pela autora, por exemplo, são encontrados índices mais altos de próclise, ao passo que, em produções mais rígidas, com o uso de expressões cristalizadas, é a ênclise a estratégia preferida. Dessa maneira, os estudos da linguista chegam à conclusão de que abordar os fenômenos linguísticos a partir de sua relação com os gêneros textuais-discursivos são um frutífero caminho a ser percorrido, de modo que se torna necessário

[...] um maior aprofundamento em busca de traços que melhor definam as naturezas dos gêneros textuais, uma vez que ainda são poucas as discussões, no meio acadêmico, que retratam essa questão, a fim de que sejam somadas, cada vez mais, informações que possam auxiliar na interpretação, mais fidedigna, dos resultados [...] (Biazolli, 2010, p. 218).

Com isso em mente, Biazolli (2016) dá continuidade às suas investigações, buscando entender como o trabalho com os gêneros e com *continua*, especialmente o estilístico e o de fala-escrita, pode auxiliar no estudo e na compreensão de diferentes processos de variação

linguística. A pesquisadora se propõe a analisar, portanto, o mesmo fenômeno com o qual trabalhou anteriormente – a colocação pronominal –, mais uma vez a partir de textos jornalísticos. Dessa vez, porém, Biazolli (2016), para a composição do *corpus* da pesquisa, investe na caracterização dos gêneros do domínio jornalístico examinados, produzidos no começo do século XXI, tanto da modalidade oral (entrevista na TV e noticiário na TV) como da escrita (carta de leitor e editorial), correspondentes a duas variedades da língua: o PB e o português europeu (PE).

Ao discutir a monitoração estilística, a estudiosa sugere que estilos específicos são definidos por diversos aspectos contextuais, sendo eles determinados tanto pelas relações entre os participantes da comunicação, que podem exigir níveis maiores ou menores de monitoramento, como pelas outras características da situação comunicativa, tais como a modalidade, o assunto, as intenções do falante, entre outros. Dessa forma, o estilo de um texto estaria diretamente conectado à situação real de uso, considerando todos os elementos que a compõem, motivo pelo qual Biazolli (2016) defende que uma observação mais atenta dos gêneros e de suas características situacionais pode ser extremamente valiosa para a análise da variação estilística.

Biazolli (2016) propõe, então, a utilização de critérios objetivos para a elaboração do *continuum* estilístico, relacionando-o não só ao *continuum* de fala-escrita, mas, também, aos próprios gêneros textuais-discursivos. Desse modo, a autora descreve os quatro gêneros de sua pesquisa a partir da lista de características situacionais elaborada por Biber e Conrad (2009) – abordadas na subseção anterior –, buscando respaldo teórico para distribuí-los no *continuum*. Nesse contexto, apesar de nem todas as características elencadas pelos autores poderem ser aplicadas à análise dos gêneros com os quais Biazolli (2016) trabalha, a pesquisadora argumenta que é apenas através da observação do gênero com um todo que se pode definir a dimensão estilística. Assim,

[...] a variação estilística passa a ser mensurada a partir do maior número possível dos fatores contextuais que compõem os gêneros textuais. Os gêneros, distribuídos em um *continuum* estilístico, moldam todas as interações comunicativas de uma comunidade [...] (Biazolli, 2016, p. 111).

Com base nisso, Biazolli (2016) aproxima o *continuum* estilístico do de fala-escrita, reforçando que apenas os conceitos de meio e de concepção, propostos por Marcuschi (2008, 2010[2001]), não são suficientes, devendo-se levar em conta outros aspectos contextuais que são mobilizados na situação comunicativa. Consequentemente, a estudiosa propõe um

continuum compósito, no qual características relacionadas a maior monitoramento, a maior formalidade e à escrita, de um lado, estão atreladas, ao passo que menor monitoramento, menor formalidade e aspectos da fala, de outro, estão interligados – além de haver uma grande escala intermediária nesse *continuum*, em que essas características se misturam. Ademais, defende também que nenhum dos gêneros analisados se caracteriza como representativo dos extremos do *continuum*, de modo que outros gêneros jornalísticos poderiam ser alocados em diferentes pontos, a depender de suas características particulares.

Como resultados, Biazolli (2016) observa uma maior produtividade de formas prestigiadas pela tradição gramatical em gêneros associados mais ao extremo da escrita, em que se encontram maior monitoramento e maior formalidade, enquanto, em gêneros mais correspondentes à esfera da oralidade, são as variantes relacionadas às normas efetivas dos falantes que se configuram como preferidas. Tal resultado aponta não só para a eficiência do *continuum* proposto no entendimento da variação encontrada, mas dá indícios, também, de que

[...] fenômenos variáveis são mais bem esmiuçados se considerados também em relação aos contextos situacionais nos quais ocorrem; por isso, deve-se continuar a investigar sobre os complexos entrelaçamentos existentes entre variação, estilo, gêneros textuais, modalidades de uso da língua e normas linguísticas [...] (Biazolli, 2016, p. 319).

A terceira iniciativa que será aqui resenhada é a de Vieira e Lima (2019), que, a partir da experiência em uma disciplina de pós-graduação, organizaram uma obra acerca das intersecções entre variação, gênero e ensino de língua portuguesa, levando em conta, também, o conceito de *continuum*. Partindo da consideração de que “[...] o mapeamento de dados em meios escritos segundo estilos variados e contemplando diversos tópicos gramaticais ainda está por ser construído [...]” (Vieira; Lima, 2019, p. 8), a referida disciplina se propôs a avaliar a distribuição de oito fenômenos variáveis do PB – dos quais quatro são abordados na obra –, tomando como base um *corpus* de gêneros jornalísticos e acadêmicos elaborado precisamente para esse fim. Com isso, buscava-se contribuir com os estudos da área, além de se discutir noções de normas e usos cultos, possibilitando, ainda, a elaboração de orientações pedagógicas para o tratamento desse assunto na escola.

Pensando na correlação entre gênero e formalidade, as autoras afirmam não haver ainda parâmetros seguros para que se determine previamente a ligação entre cada gênero e um grau específico de monitoramento estilístico, reconhecendo nisso um tema de pesquisa a ser abordado. Por isso, assumem que gêneros textuais-discursivos distintos admitem graus de formalidade também diferentes, sendo esse aspecto relacionado a elementos como audiência,

suporte textual, tema, entre outros. Dessa maneira, a investigação empírica de vários fenômenos partindo somente do gênero se configura como uma importante tarefa a ser realizada, principalmente no que diz respeito à identificação das tendências no uso de uma ou outra variante linguística (Vieira; Lima, 2019).

Com base em considerações teóricas acerca de *continua* (Marcuschi, 2010[2001]; Bortoni-Ricardo, 2005), a composição do *corpus*, então, foi realizada com a preocupação de trabalhar tanto com gêneros mais prototipicamente relacionados ao extremo da escrita como com os associados à oralidade, tendo a entrevista sociolinguística como contraponto aos demais gêneros, realizados no meio gráfico¹⁵. A partir deles, foram coletados e analisados os dados correspondentes a cada um dos fenômenos linguísticos sob investigação¹⁶, a fim de que se pudesse elaborar um *continuum* de oralidade-letramento que funcionasse de forma compósita ao de monitoração estilística – proposta semelhante à de Biazolli (2016).

De maneira geral, os estudos apresentados em Vieira e Lima (2019) demonstram que os gêneros mais próximos ao polo de maior concepção de oralidade foram aqueles que apresentaram os menores índices de uso das formas tidas como prestigiosas, ao passo que os alocados no polo de maior letramento demonstraram as taxas mais altas de utilização dessas formas, havendo, em termos gerais, um aumento crescente desses índices ao longo do *continuum* proposto. As poucas exceções observadas foram encontradas em gêneros da porção intermediária do *continuum*, que misturam características da oralidade e da escrita e, por isso, acomodam, aparentemente, um espectro de variação mais dilatado. Diante disso, as autoras concluem, então, que os resultados encontrados corroboraram a hipótese principal de que as proporções de uso das variantes de prestígio – comumente identificadas como padrão – seriam sensíveis aos gêneros textuais-discursivos, aumentando conforme a modalidade e o grau de monitoração estilística.

Ainda outro trabalho que merece ser resenhado é o de Biazolli e Berlinck (2021), que discutem aspectos teórico-metodológicos com vistas a defender a investigação de processos de variação e de mudança linguísticas por meio de gêneros textuais-discursivos. No texto, as autoras iniciam o debate realizando uma retomada das ideias de Bakhtin (2016[1979]), de modo a relacioná-las diretamente a conceitos essenciais da Sociolinguística Variacionista. Partindo do

¹⁵ No polo da oralidade, foram consideradas entrevistas sociolinguísticas, anúncios e tirinhas; no da escrita, teses e dissertações, artigos científicos e editoriais; e, na porção intermediária do *continuum*, foram alocadas entrevistas impressas, notícias, cartas de leitor e crônicas.

¹⁶ Os quatro fenômenos cujos resultados são apresentados em Vieira e Lima (2019) são: a realização do acusativo anafórico de terceira pessoa, a expressão do dativo de terceira pessoa, as estratégias de relativização e a alternância entre ter e haver existenciais.

princípio bakhtiniano de que toda manifestação linguística toma forma por meio de um gênero e, por conta disso, é essencial que essa dimensão seja considerada ao realizar qualquer tipo de análise acerca da língua, as estudiosas destacam que o caráter relativamente estável dos gêneros permite a variação linguística, de modo que considerar esse aspecto, em conjunto com as condições de produção e de recepção, é especialmente relevante em pesquisas sociolinguísticas. Sobre isso, Biazolli e Berlinck (2021, p. 15) afirmam que

[...] o estudo de processos de variação e mudança linguísticas é uma das áreas da Linguística em que, talvez mais do que em outras, a realidade da língua em uso é uma condição tão essencial quanto inescapável. A língua varia e muda pela ação contínua de seus falantes, em interações verbais concretas. Desse modo, fica evidente que não é possível investigar em sua completude tais processos sem considerar como a natureza do gênero os afeta, os molda, os determina; ao mesmo tempo que essa natureza é afetada, moldada e determinada por eles [...].

Entretanto, a maior contribuição das linguistas talvez esteja na descrição detalhada que elas fazem da maneira como o gênero tem sido tratado em estudos variacionistas, considerando todas as etapas envolvidas em sua realização. Começando pela constituição do *corpus*, Biazolli e Berlinck (2021) destacam que, para estudos que se projetem para além das entrevistas sociolinguísticas, considerar os gêneros textuais-discursivos como eventos comunicativos permite a análise de dados associados a diferentes níveis de formalidade e monitoramento, sendo possível observar esses aspectos tanto em textos que materializam um mesmo gênero como em textos de diferentes gêneros. É o que se observa, por exemplo, nos estudos até aqui mencionados, que efetuam um intenso trabalho de investigação sobre a relação entre fenômenos variáveis e a distribuição de gêneros textuais-discursivos em *continua*.

Quanto ao passo da definição do fenômeno e da seleção de dados, as estudiosas enfatizam que a escolha de um gênero textual-discursivo como material a ser investigado é feita também por se hipotetizar, devido a suas características estruturais e funcionais, que ele seja um local propício para a coleta de ocorrências significativas do objeto de estudo. Nesse sentido, não seria produtivo, por exemplo, abordar formas de tratamento em gêneros que não envolvam referência a um interlocutor ou analisar a expressão do futuro em textos predominantemente focados no relato de acontecimentos passados.

Em seguida, Biazolli e Berlinck (2021) discutem a relevância de se pensar as categorias de análise a partir do gênero textual-discursivo com o qual se trabalha, reiterando o papel central que essa etapa representa em estudos variacionistas. Na concepção das autoras, considerar o gênero nesse ponto “[...] é transformar em variáveis ou grupos de fatores o gênero, tomado

como uma unidade ou entidade, e/ou elementos de sua construção composicional, de seu conteúdo temático, de suas condições de produção [...]” (Biazolli; Berlinck, 2021, p. 28), levando em conta o fato de que esses aspectos podem motivar o falante a realizar certas escolhas linguísticas. Nesta pesquisa, como será descrito em detalhes na seção 5, é precisamente isso que se tenta fazer ao propor as características típicas das legendas audiovisuais como variáveis a serem investigadas.

Com isso em mente, deve-se, então, pensar nos gêneros textuais-discursivos também no momento da quantificação dos dados e, posteriormente, da interpretação dos resultados. Ao controlar de maneira mais refinada a distribuição das variáveis de certo fenômeno linguístico de acordo com os aspectos formais e situacionais dos textos que compõem o *corpus*, é possível obter uma visão mais global da maneira como cada traço do gênero influencia em sua realização, garantindo uma maior objetividade na análise. Dessa forma, considerações acerca do gênero acabam servindo como respaldo para a discussão dos resultados, que ganham, portanto, novas possibilidades de interpretação (Biazolli; Berlinck, 2021).

Assim, é possível observar que Biazolli e Berlinck (2021) apontam diversos caminhos para o trabalho com a variação linguística e o gênero textual-discursivo de maneira integrada em todos os passos da pesquisa, sustentando a produtividade de tal abordagem. Defendem, portanto, que

[...] o estudo da variação e da mudança pode ser feito na busca de padrões gerais (explorando grandes *corpora*, com quantificação de dados), mas também na identificação de comportamentos singulares, motivados pelas condições específicas de produção e recepção do texto-discurso, o que inclui a menor ou maior agentividade do enunciador na escolha de variantes. A materialização dessas duas dimensões se dá no texto que, por sua vez, corporifica o gênero. A percepção dessa dualidade leva a uma interpretação mais qualificada do fenômeno variável [...] (Biazolli; Berlinck, 2021, p. 34).

O penúltimo trabalho a ser apresentado nesta subseção consiste na tese de Lima (2022)¹⁷, que, além de analisar a ordem dos clíticos pronominais, também se ocupa da descrição da variação do acusativo anafórico de terceira pessoa em textos jornalísticos do jornal O Globo, considerando as intersecções entre variação, gênero textual-discursivo e norma. Assim como na maioria dos estudos aqui apresentados, a pesquisa da autora também se pauta no trabalho com o *continuum* de oralidade-letramento, partindo do pressuposto de que “[...] as

¹⁷ O trabalho de Lima (2022) é retomado na próxima seção desta dissertação, dedicada à descrição do OD anafórico de terceira pessoa.

características situacionais dos gêneros textuais [...] contribuem para a compreensão da variação linguística presente nos textos escritos [...]” (Lima, 2022, p. 90).

Para testar sua hipótese de que variantes prestigiadas seriam as formas preferidas na porção de maior letramento, ao passo que, na porção de maior oralidade, o contrário seria verificado, a autora compõe um *corpus* com sete gêneros textuais-discursivos, caracterizando-os cuidadosamente a partir dos aspectos situacionais elencados por Biber e Conrad (2009), assim como o fez Biazolli (2016). A autora chega, então, a uma distribuição objetiva dos gêneros no *continuum*, permitindo que ela trace considerações mais seguras e embasadas acerca do modo como o gênero textual-discursivo e/ou o *continuum* influenciam as escolhas linguísticas realizadas pelos autores dos textos.

Após transformar os próprios gêneros e o *continuum* em variáveis independentes de seu estudo, Lima (2022) constata que as variantes do fenômeno se distribuíram de maneira distinta tanto entre os gêneros do *corpus* como entre as diferentes porções do *continuum*. Nas primeiras análises, o SN anafórico não pareceu ser sensível à variação estilística, uma vez que ocorreu de maneira irregular no que diz respeito ao gênero/*continuum*. Por outro lado, tanto o clítico acusativo como o objeto nulo apresentaram um comportamento escalar com relação a essa variável: o primeiro, aumentando do extremo de maior oralidade para o de maior letramento; o segundo, demonstrando comportamento oposto, com suas proporções diminuindo na mesma direção.

Além disso, nas duas análises multivariadas realizadas pela pesquisadora – uma com e a outra sem o SN anafórico –, a variável que representava o *continuum* foi apontada como significativa para a realização do fenômeno. De acordo com seus resultados, o clítico, isto é, a variante padrão, foi a forma mais utilizada no polo de maior letramento, ao passo que, nos outros dois pontos do *continuum*, ele foi desfavorecido em relação às outras variantes. Para explicar esses achados, Lima (2022) leva em consideração as orientações normativas que, em geral, apontam o pronome átono como a única possibilidade para a realização do OD anafórico. Assim, a autora confirma suas hipóteses, provando que a ocorrência de seu fenômeno é sensível aos diferentes gêneros textuais-discursivos, de modo a validar mais uma vez a produtividade de se realizar estudos variacionistas voltando o olhar para essa dimensão.

É precisamente em face de resultados como esse que surge, então, a última iniciativa a ser apresentada nesta subseção: o projeto “Pró-norma plural: do *continuum* fala-escrita para a norma-padrão”, encabeçado pela professora Silvia Rodrigues Vieira, da Universidade Federal do Rio de Janeiro (UFRJ). Tendo como principal objetivo contribuir para o conhecimento da modalidade escrita do PB, através de descrições do que é comumente chamado de “norma

culta”, e auxiliar na formulação de orientações normativas empiricamente fundamentadas para esse contexto, o projeto, ainda em andamento, prevê uma série de outras etapas.

A primeira delas consiste na elaboração de um *corpus* para a realização da pesquisa, composto por gêneros textuais-discursivos da modalidade escrita, referentes aos domínios jornalístico e acadêmico, de quatro capitais brasileiras, a saber: Recife, Rio de Janeiro, Salvador e São Paulo. Mantendo-se os procedimentos metodológicos utilizados em Lima (2022) para a composição das amostras, propõe-se a coleta de sete gêneros em veículos jornalísticos de ampla circulação, voltados a classes sociais privilegiadas, além de dois gêneros acadêmicos com textos da área de Comunicação Social.

A partir do material reunido, a proposta é caracterizar diversas regras variáveis de natureza morfossintática, relacionando os resultados encontrados aos gêneros do *corpus*, a fim de fornecer evidências empíricas para a postulação de um *continuum* de fala-escrita. Ademais, o projeto visa, também, estabelecer relações entre as orientações normativas em voga e as tendências observadas nos distintos usos cultos controlados pela análise quantitativa dos dados obtidos, com vistas a definir parâmetros para uma norma de referência para o PB que seja essencialmente flexível e adaptável.

Apesar de ainda estar sendo desenvolvido, o projeto já conta com alguns resultados preliminares publicados, como é o caso do estudo descrito em Biazolli *et al.* (2023)¹⁸, realizado pela subequipe de São Paulo. Nele, analisou-se a colocação pronominal em uma subamostra dos textos jornalísticos que contava com mais de 70 mil palavras – aproximadamente 10 mil para cada gênero –, focando-se, principalmente, nos dados adjungidos a lexias verbais simples, já que correspondiam à maioria das ocorrências. Ao voltar o olhar para os gêneros textuais-discursivos distribuídos nos *continua* estilístico e de fala-escrita, atestou-se, como era esperado, um aumento no uso da ênclise em gêneros considerados de maior monitoramento e formalidade, com temas mais complexos. Contudo, a mesma regularidade não pôde ser observada nos outros pontos dos *continua*, observando-se uma distribuição diversa, de modo que foi indicada a necessidade de refinar as análises para explicar o comportamento do fenômeno nesses contextos. Desse modo, esses resultados, apesar de ainda incipientes, apresentam-se como promissores, parecendo indicar uma correlação entre a realização de fenômenos variáveis e gêneros textuais-discursivos, ao menos no que diz respeito à associação de formas mais prestigiadas a gêneros tidos como mais formais e monitorados.

¹⁸ Vale mencionar que os resultados de Lima (2022) também integram o projeto “Pró-norma plural”, uma vez que seu *corpus* corresponde, hoje, à amostra de textos jornalísticos do Rio de Janeiro.

Diante do exposto, acredita-se que uma aproximação entre a Sociolinguística Variacionista e o estudo dos gêneros textuais-discursivos é não só necessária, mas imprescindível para um entendimento mais acurado de processos de variação e mudança linguísticas. Portanto, a presente pesquisa se propõe a observar a realização do OD anafórico de terceira pessoa em legendas audiovisuais, levando em conta todos os aspectos estruturais e situacionais que caracterizam esse gênero para que se possa chegar a uma descrição fidedigna do fenômeno, afinal

[...] o desconhecimento da natureza do enunciado e a relação indiferente com as peculiaridades das diversidades de gênero do discurso em qualquer campo da investigação linguística redundam em formalismo e em uma abstração exagerada, deformam a historicidade da investigação, debilitam as relações da língua com a vida [...] (Bakhtin, 2016[1979], p. 16)

3 O OBJETO DIRETO ANAFÓRICO: DA TRADIÇÃO GRAMATICAL AOS ESTUDOS LINGUÍSTICOS

Para qualquer estudo linguístico, é essencial que o fenômeno sob investigação seja compreendido a partir de diversos pontos de vista, já que é esse respaldo teórico que permite ao pesquisador chegar a conclusões bem embasadas acerca de seus resultados. Com isso em mente, a presente seção busca levantar o que tem sido dito sobre a realização do OD anafórico no português, em especial no PB, incluindo sua relação com o quadro pronominal da língua. Assim, as próximas subseções trarão um panorama das considerações de gramáticas tradicionais (GTs), de abordagens descritivas e de estudos (sócio)linguísticos referentes a esse tema.

3.1 O OD anafórico e a tradição gramatical

Não é novidade, para qualquer estudioso da linguística, que os compêndios gramaticais de orientação normativa se distanciam consideravelmente da língua que é, efetivamente, empregada pelos falantes em suas vidas cotidianas, tanto na modalidade oral como na escrita. Com efeito, é possível perceber nas GTs consultadas – Rocha Lima (2011), Cunha e Cintra (2017) e Bechara (2009)¹⁹ – uma preocupação em prescrever regras para o bem falar e escrever, o que, muitas das vezes, as leva a ignorar ou até mesmo condenar formas já consolidadas pelo uso. Entretanto, apesar dessa abordagem falha, é importante levar em conta as recomendações que essas obras oferecem, posto que elas ainda representam, em muitos contextos, aquilo que é tido como modelo de língua e podem, portanto, influenciar as escolhas linguísticas dos falantes, especialmente em situações de maior monitoração estilística.

Feitas essas considerações iniciais, pode-se passar, agora, à revisão do que as três gramáticas supracitadas trazem com relação à realização do OD anafórico. Para tal, é importante, porém, primeiro entender o que essas gramáticas tratam como pronome e como apresentam o quadro pronominal do PB, uma vez que é a partir da definição de pronomes pessoais que o acusativo anafórico é por elas tratado.

No geral, as três gramáticas apresentam definições semelhantes para a classe de palavras aqui em questão. Rocha Lima (2011, p. 156), por exemplo, define pronome com base no conceito elaborado por Said Ali, ou seja, como “[...] a palavra que denota o ente ou a ele se refere, considerando-o apenas como pessoa do discurso [...]” (Said Ali, 1985, p. 61 *apud* Rocha

¹⁹ Apesar de contarem com edições relativamente recentes, as três obras observadas tiveram suas primeiras versões publicadas em 1972, 1985 e 1961, respectivamente.

Lima, 2011, p. 156). Essa questão da denotação é mais bem explorada por Bechara (2009, p. 162) ao traçar comentários sobre a questão semântica que envolve esse tipo de forma, afirmando que

[...] os pronomes estão caracterizados porque indicam *dêixis* (“o apontar para”), isto é, estão habilitados, como verdadeiros gestos verbais, como indicadores, determinados ou indeterminados, ou de uma *dêixis* contextual a um elemento inserido no contexto [...] A *dêixis* será *anafórica* se aponta para um elemento já enunciado ou concebido, ou *catafórica*, se o elemento ainda não foi enunciado ou não está presente no discurso.

Já Cunha e Cintra (2017, p. 289) apresentam uma definição mais ligada à questão sintática, caracterizando pronomes como os termos que “[...] desempenham na oração as funções equivalentes às exercidas pelos elementos nominais [...]”, ou seja, podem representar substantivos, caso no qual seriam considerados pronomes substantivos, ou acompanhar um substantivo e delimitar seu significado, sendo, neste contexto, pronomes adjetivos.

Após essas primeiras definições, os três compêndios elencam os seis tipos de pronomes existentes no português, a saber, pessoais, demonstrativos, relativos, possessivos, indefinidos e interrogativos, e passam a se dedicar à descrição de cada um deles. Nesse contexto, definem, de maneira geral, os pronomes pessoais como aqueles que designam as três pessoas do discurso, podendo representar uma forma nominalmente expressa anteriormente quando na 3ª pessoa e variando de acordo com a função que exercem na oração e a acentuação que recebem (Cunha e Cintra, 2017, p. 290). Desse modo, consideram-se retos ou subjetivos os pronomes que figuram como sujeito na oração, e como oblíquos ou objetivos aqueles empregados em função de objeto. Quanto à tonicidade, os pronomes pessoais oblíquos podem ser átonos ou tônicos, devendo estes últimos serem sempre precedidos por alguma preposição.

Desse modo, os pronomes pessoais do PB se organizam da seguinte maneira:

Quadro 1 - Sistema pronominal do português brasileiro

PRONOMES PESSOAIS RETOS			PRONOMES PESSOAIS OBLÍQUOS	
			átonos (sem prep.)	tônicos (c/prep.)
Singular:	1ª pessoa:	<i>eu</i>	<i>me</i>	<i>mim</i>
	2ª pessoa:	<i>tu</i>	<i>te</i>	<i>ti</i>
	3ª pessoa:	<i>ele, ela</i>	<i>lhe, o, a, se</i>	<i>ele, ela, si</i>
Plural:	1ª pessoa:	<i>nós</i>	<i>nos</i>	<i>nós</i>
	2ª pessoa:	<i>vós</i>	<i>vos</i>	<i>vós</i>
	3ª pessoa:	<i>eles, elas</i>	<i>lhes, os, as se</i>	<i>eles, elas, si</i>

Fonte: Adaptado de Bechara (2009, p. 162)

Diante desse quadro, as três GTs observadas definem os pronomes pessoais oblíquos átonos como formas adequadas para expressar um OD anafórico, ou seja, um OD que retoma um termo ou oração já mencionado anteriormente. Todavia, a partir desse ponto, cada uma das obras faz comentários específicos referentes a esse tema, motivo pelo qual elas serão agora tratadas de maneira separada.

Em Rocha Lima (2011), são poucas as informações dadas acerca das maneiras de expressar o acusativo anafórico de terceira pessoa. A principal delas é, precisamente, o que já foi mencionado anteriormente: “As formas *o, a, os, as* empregam-se em substituição a um substantivo que, sem vir precedido de preposição, completa o regime de um verbo [...]”, sendo exemplos “Vi *o menino* (ou – *vi-o*)” e “Não escrevi *as cartas* (ou – não *as* escrevi)” (Rocha Lima, 2011, p. 157). Para além disso, o gramático apenas aponta as regras para conjugação de verbos com os pronomes átonos, explicando quando usar “*lo(s)/la(s)*” e “*no(s)/na(s)*” (p. 206) e explicita que os pronomes oblíquos podem desempenhar a função de sujeito de infinitivo acumulada à de objeto, quando conectados a verbos causativos ou perceptivos, como em “Mandei-*o* entrar” e “Deixe-*nos* pensar” (Rocha Lima, 2011, p. 390). Não há qualquer menção a outras formas de realizar essa retomada.

Já em Cunha e Cintra (2017), logo após caracterizarem os pronomes quanto à função, os autores chamam a atenção para as mudanças fonéticas sofridas pelas formas oblíquas de terceira pessoa a depender do verbo ao qual se ligam, explicitando em quais contextos de ênclise os pronomes objetivos “*o(s)*” e “*a(s)*” assumem as formas “*lo(s)/la(s)*” e “*no(s)/na(s)*” (Cunha e Cintra, 2017). Desse ponto em diante, a obra destina seções diferentes para tratar do emprego de pronomes pessoais retos e oblíquos.

Quanto às formas do caso reto, reconhece-se na GT em questão apenas duas possibilidades para sua manifestação: como sujeito ou predicativo do sujeito. Por isso, o uso de “ele/ela” e suas formas plurais em função acusativa é apresentado ao leitor em uma parte voltada a equívocos e incorreções no uso desses pronomes, segundo a qual,

[...] na fala vulgar e familiar do Brasil, é muito frequente o uso do pronome *ele(s), ela(s)* como objeto direto em frases do tipo:

Vi **ele**.

Encontrei **ela**.

Embora esta construção tenha raízes antigas no idioma, pois se documenta em escritores portugueses dos séculos XIII e XIV, deve ser hoje evitada. (Cunha e Cintra, 2017, p. 302)

Na parte especificamente voltada ao emprego dos pronomes oblíquos átonos, afirma-se somente que as formas “o”, “a” e seus plurais são próprias do OD, sem maiores considerações sobre o assunto. Em uma seção subsequente, os gramáticos, assim como acontece em Rocha Lima (2011), reconhecem a possibilidade de um pronome átono aparecer como sujeito de um infinitivo em construções como em “Mandei-**o sair**” (Cunha e Cintra, 2017, p. 316), isto é, quando ligado a verbos causativos ou perceptivos. Mais adiante, preveem também a possibilidade de se utilizar as formas pronominais átonas de terceira pessoa para dar ênfase ao OD, em estruturas com topicalização, tais como:

(12) **Verdades**, quem é que **as** quer? (F. Pessoa, *OP*, 530.) (Cunha e Cintra, 2017, p. 316)

(13) **O meu avô**, nunca **o** vi rezando. (J. Lins do Rego, *ME*, 83.) (Cunha e Cintra, 2017, p. 316)

Dessa maneira, pode-se perceber que, apesar de se propor a descrever “[...] o português atual na sua forma culta [...]”, sem descuidar, porém, “[...] dos fatos da linguagem coloquial [...]” (Cunha e Cintra, 2017, p. XXV), essa GT não abarca todas as formas correntes de realizar o OD anafórico de terceira pessoa. Pelo contrário, além de ignorar estratégias como o objeto nulo, condena a utilização de pronomes retos nessa função, mesmo sendo esse um uso comum na língua, inclusive nos contextos levantados pelos autores.

Das três GTs consultadas, talvez seja a de Bechara (2009) a que mais aborda as variantes do fenômeno analisado neste estudo, ainda que timidamente. Nela, o emprego dos pronomes é tratado também em uma seção do capítulo sobre essa classe morfológica, que se inicia com a

afirmação de que, a rigor, o pronome oblíquo funciona como complemento verbal. Em seguida, o autor elenca os casos em que a forma reta poderia ocorrer no lugar da oblíqua, sendo elas: (i) quando houver distância entre o verbo e o complemento nominal, e a separação acontecer por pausa; (ii) em enumerações e aposições nas quais verbo e complemento também estiverem distanciados; (iii) precedido pelos termos *todo*, *só* e alguns outros adjuntos; e (iv) quando aparece com acentuação enfática, ao final de um grupo de força (Bechara, 2009). São exemplos de cada um desses contextos os trechos (14) a (17) abaixo, respectivamente.

(14) Subiu! E viu com seus olhos. / *Ela* a rir-se que dançava [GD]. (Bechara, 2009, p. 173)

(15) Depois de muita delonga o diretor escolheu: *eu*, o Henrique e o Paulinho. (Bechara, 2009, p. 173)

(16) No latim eram quatro os pronomes demonstrativos. *Todos eles* conserva o português [PL. 1, 398] (Bechara, 2009, p. 175)

(17) Olha *ele*! [EQ] (Bechara, 2009, p. 173)

Por fim, o autor tece uma breve consideração acerca do objeto nulo, defendendo que, “[...] estando perfeitamente conhecido pela situação linguística, pode-se calar o pronome complemento do verbo; esta linguagem é correta, apesar da censura que lhe faziam os gramáticos de outrora [...]” (Bechara, 2009, p. 174). Entretanto, além de tal afirmação não dar conta da produtividade dessa variante, chama-se a atenção apenas para os contextos em que o objeto nulo ocorre após a repetição de parte de uma oração e reforça, portanto, a declaração, deixando de lado todos os outros numerosos contextos em que essa categoria vazia aparece.

Com base no exposto até aqui, pode-se perceber que, de maneira geral, as GTs demonstram ainda uma visão limitada acerca das possibilidades para realização do OD anafórico de terceira pessoa. Todas elas apontam o pronome pessoal oblíquo átono como a forma mais adequada para a expressão do OD e, na maioria dos casos, se reconhecem outras variantes, as condenam. Isso comprova o teor prescritivo que essas gramáticas demonstram, além de enfatizar o descompasso existente entre a língua realmente produzida pelo falante nos mais diversos contextos de comunicação e o padrão conservador sistematizado nessas obras. Assim, reforça-se a necessidade de voltar um olhar crítico a essas obras, uma vez que as regras nelas dispostas não conseguem dar conta da realidade da língua.

3.2 O OD anafórico em estudos descritivos

Estudos descritivos são, por princípio, voltados a uma exposição objetiva e sistemática de um sistema linguístico, sem que haja qualquer preocupação em prescrever como a língua deveria ser, mas observando como ela de fato é. É nesse contexto que se encaixam as obras destacadas nesta subseção, de Câmara Jr. (1976, 1992[1970]), Mateus *et al.* (2003[1983]) e Castilho (2020[2010]), com vistas a ampliar a visão sobre o objeto de estudo desta pesquisa.

Comentando sobre a morfologia pronominal do PB, Câmara Jr. (1976) afirma serem pronomes as formas indicativas ou dêiticas de uma língua que indicam elementos do mundo biossocial não de acordo com conceitos estabelecidos, mas com base na posição que ocupam no momento da comunicação linguística. Nesse panorama, encaixam-se os pronomes pessoais, cujo quadro é descrito pelo autor, assim como nas gramáticas normativas, apontando as formas retas e oblíquas – tônicas e átonas –, e destacando suas funções. Todavia, Câmara Jr. vai além das obras até aqui apresentadas, fazendo a ressalva de que “[...] tal quadro é, a rigor, puramente teórico; e em nenhuma região da língua portuguesa ele se realiza exatamente” (Câmara Jr., 1992[1970], p. 119).

Desse momento em diante, o estudioso se dedica a descrever as mudanças ocorridas no quadro pronominal ao longo do tempo, começando pelas alterações que atingiram a segunda pessoa. Câmara Jr. (1992[1970], p. 118) chama a atenção, então, para a entrada de “você” no sistema pronominal, salientando que ela pressupõe a manutenção das formas oblíquas adverbiais de terceira pessoa, apesar de, em registro informal, o átono “te” figurar junto ao novo pronome. Em seguida, menciona o deslocamento de “lhe” como termo correspondente a “você” e chega, enfim, à eliminação dos pronomes oblíquos de terceira pessoa em proveitos de “ele/a”. Assim, fica caracterizado um sistema que “[...] praticamente eliminou as formas de acusativo e dativo do pronome de 3ª pessoa [...]” (Câmara Jr., 1976, p. 237), de modo que as formas “ele/a” são os únicos pronomes pessoais nesse contexto, podendo funcionar como sujeitos e/ou objetos de uma oração.

Percebe-se, portanto, que as considerações de Câmara Jr (1992[1970]), embora breves, já apresentam como fato o uso dos pronomes “ele/a” como ODs anafóricos, sem coibir esse emprego, posto que se trata, puramente, de uma descrição dos fatos da língua. Apesar disso, o autor ainda não menciona as outras estratégias utilizadas pelos falantes para realizar retomadas anafóricas, deixando sua análise, nessa obra, ainda incompleta.

Por outro lado, Mateus *et al.* (2003[1983]) fazem uma exposição extensiva acerca do OD, assim como dos processos anafóricos no português. Por isso, o trabalho das autoras é aqui

apresentado, mesmo sendo uma obra focada no PE. Ademais, observar o que se diz sobre o OD anafórico em outra variedade do português pode auxiliar na compreensão do fenômeno como um todo, além de permitir o reconhecimento de diferenças entre o PB e PE, ainda que não seja esse o foco desta pesquisa.

A primeira menção que as autoras fazem à realização do OD anafórico acontece já na primeira parte da gramática, justamente em uma seção que busca comparar características gerais do PE e do PB. Ao elencar as diferenças entre as duas variedades nos níveis morfológico e sintático, Mateus *et al.* (2003[1983], p. 47) logo citam a utilização dos clíticos de terceira pessoa como um ponto de divergência, afirmando que “[...] o PB vernacular perdeu os clíticos de terceira pessoa e apresenta, ao lado do objecto directo nulo, construções com os pronomes *ele / ela e lhe* [...]”. São reconhecidas, portanto, três outras estratégias para a expressão do OD anafórico, além dos átonos: o pronome “*ele/a*”, o dativo “*lhe*” e o objeto nulo, caracterizando-os como um traço típico do PB. Como exemplos, são apresentadas as seguintes frases para o PB:

(18) as frases, ele tinha lido (-) nos livros (Mateus *et al.*, 2003[1983], p. 47)

(19) eu vi *ele* na rua (Mateus *et al.*, 2003[1983], p. 47)

(20) deixa *ela* comigo (Mateus *et al.*, 2003[1983], p. 47)

(21) quero *lhe* conhecer (Mateus *et al.*, 2003[1983], p. 47)

Posteriormente, a questão do objeto nulo é retomada em um capítulo inteiramente voltado às relações gramaticais, no qual se explora, além de outros termos essenciais da oração, o OD. Após uma breve definição, que considera serem ODs os argumentos internos diretos de predicadores verbais de dois ou três lugares, são enumeradas as propriedades típicas desses termos, sendo a primeira delas, justamente, a possibilidade de o objeto ser nulo nos seguintes contextos: (i) com verbos transitivos cujo argumento interno direto tem uma interpretação arbitrária, como em “O João leu toda a noite”; (ii) em construções de SV²⁰ nulo com verbos transitivos, tais como “O João comeu *uma tarte* e a Mari também comeu [-]”; e (iii) em construções de objeto nulo, sendo um exemplo “Tens visto *o Pedro?* / Vi [-] ontem na conferência mas não falei com ele” (Mateus *et al.*, 2003[1983], p. 285).

²⁰ Seguindo a abreviação adotada em Mateus *et al.*, 2003[1983], “SV” representa, aqui, construções de estrutura “sujeito + verbo”.

Esse último tipo de construção é retomado pelas estudiosas ao discutirem frases com tópicos marcados. Em uma subseção voltada ao OD nulo, são apresentados inicialmente exemplos em que esse termo retoma tanto SNs como objetos presentes na situação comunicativa (exemplos (22) e (23), respectivamente), podendo sempre alternar com um clítico. As autoras afirmam, ainda, que sentenças com objeto nulo partilham diversas propriedades com estruturas de topicalização, podendo ter como referência o mesmo tipo de SN.

(22) A: E o carro?

B: Parece que o Pedro espatifou [-] contra um muro. (Mateus *et al.*, 2003[1983], p. 502)

(23) A aponta para a camisola que B tem vestida.

A: Tricotaste tu [-]? (Mateus *et al.*, 2003[1983], p. 502)

Mais adiante, afirma-se que o objeto nulo é sensível a contextos de ilha sintática, sendo que a presença do clítico é imprescindível no PE, levando a construção a ser entendida como uma variável, em termos gerativistas, nessa variedade. De acordo com essa análise, o objeto nulo seria então um caso de topicalização no qual o próprio tópico marcado é uma categoria vazia que, deslocando-se à esquerda do sintagma flexionado, assumiria a função de operador sintático. Por outro lado, no PB e em outras variedades do português, é possível a ocorrência de objetos nulos em contextos de ilhas, o que sugere que, ao contrário do que é observado no PE, o objeto nulo se apresenta como uma categoria pronominal *pro*, sendo essa uma diferença marcada entre essas variedades. Desse modo, exemplos como “Gosto imenso *desse romance*, embora só tenha lido [-] uma vez” (Mateus *et al.*, 2003[1983], p. 505) seriam absolutamente gramaticais em PB, mas não em PE.

Posteriormente, Mateus *et al.* (2003[1983]) exploram a diferença entre o que chamam de anáfora do complemento nulo e objeto nulo. Explicam, então, que a primeira estrutura se dá quando a elipse corresponde a um antecedente oracional, podendo ser substituída somente pelo clítico neutro “o” ou o pronome demonstrativo “isso”; em contrapartida o objeto nulo em si retomaria preferencialmente SNs que denotam entidades, alternando tanto com os pronomes invariáveis acima mencionados como com pronomes pessoais. Apontam, também, que a anáfora do complemento nulo é lexicalmente determinada, só ocorrendo com semiauxiliares modais e alguns verbos de complementação, enquanto o objeto nulo não é determinado pelo tipo de verbo que ocorre junto ao antecedente, apesar de poder sofrer restrições em contextos

de subordinação adverbial e relativa. São exemplos dessas duas construções as frases (24) e (25) abaixo, com a anáfora e o objeto nulo, respectivamente:

(24) Ainda que queiras [-], não podes resolver esse problema com facilidade. (Mateus *et al.*, 2003[1983], p.884)

(25) Misture todos os ingredientes: o açúcar, a farinha, a manteiga e os ovos. Bata [-] bem. (Mateus *et al.*, 2003[1983], p. 887)

Além disso, as autoras abordam ainda o conceito de anáfora, definindo-a como “[...] o processo que consiste em utilizar uma forma linguística ou um vazio para remeter para algo que já foi dito anteriormente (o antecedente) [...]” (Mateus *et al.*, 2003[1983], p. 802), e elencando seus diferentes tipos. Dentre eles, destacam-se, para este estudo, a anáfora nominal, a pronominal e a realizada através de demonstrativos invariáveis.

De acordo com Mateus *et al.* (2003[1983], p. 802), a anáfora nominal se divide em dois tipos: a fiel, quando há repetição do nome, com ou sem alteração do determinante – por exemplo, “uma menina” e “essa menina” –, e a infiel, quando a retomada acontece utilizando expressões de diferentes significados, mas com mesma referência (como na troca de “o avião” por “a aeronave”). Já a anáfora pronominal ocorre quando um SN é referenciado através de pronomes foneticamente realizados ou vazios, com interpretação dependente do antecedente. Por fim, a anáfora através das formas invariáveis “o” ou “isso” corresponde aos casos em que o clítico neutro e o demonstrativo são utilizados em substituição a orações. Assim, estariam englobadas nesses tipos de anáfora todas as variantes tipicamente consideradas em estudos acerca da realização do OD anafórico, ou seja, o SN anafórico, o clítico, o pronome lexical e o objeto nulo.

Pode-se perceber, portanto, que, apesar de majoritariamente focada no PE, a obra de Mateus *et al.* (2003[1983]) reconhece, de alguma maneira, todas as variantes do fenômeno aqui em discussão, traçando análises aprofundadas sobre algumas delas. Por isso, contribui para o entendimento do fenômeno, se mostrando uma fonte valiosa para o estudo do fenômeno aqui explorado.

Castilho (2020[2010]), por outro lado, elabora uma obra completamente focada no PB, demonstrando, além de um olhar estritamente descritivo, um cuidado em abordar tanto os usos formais como os informais da língua. Nesse cenário, o linguista define os pronomes como formas que, do ponto de vista semântico-discursivo, “[...] (1) representam as pessoas do discurso, pelo caminho da dêixis, (2) permitem a retomada ou antecipação de participantes, pelo

caminho da foricidade (anáfora e catáfora) [...]” e, do ponto de vista gramatical, exibem “[...] as propriedades morfológicas de (i) caso; (ii) pessoa e número; (iii) gênero [...]” (Castilho, 2020[2010], p. 474). Divide, então, essas formas em dois grupos, com base nas propriedades de adjacência e substituição: um de pronomes possessivos, demonstrativos e quantificadores, que podem acompanhar e substituir substantivos, e outro de pronomes pessoais, que nas primeiras duas pessoas do discurso não acompanham e nem substituem outras formas e, na terceira pessoa, apenas as substitui (Castilho, 2020[2010]).

Em seguida, o autor explicita que os pronomes pessoais de primeira e de segunda pessoa são dêiticos e invariáveis, e os de terceira pessoa caracterizam-se como fóricos, apresentando flexão de gênero e número. Ademais, afirma que os pronomes pessoais são facilmente suscetíveis a mudanças, assim como atestado por estudos recentes que apontam para uma reorganização dessas formas, trazendo consequências para a estrutura do PB (Castilho, 2020[2010]). Diante disso, apresenta o quadro dos pronomes pessoais já considerando as alterações ocorridas ao longo do tempo – e já descritas anteriormente nesta seção –, de maneira que os pronomes de terceira pessoa ficam organizados da seguinte maneira:

Quadro 2 - Quadro dos pronomes pessoais de terceira pessoa no PB

PESSOA	PB FORMAL		PB INFORMAL	
	Sujeito	Complemento	Sujeito	Complemento
3ª pessoa sg.	<i>ele, ela</i>	<i>o/a, lhe, se, si, consigo</i>	<i>ele/ei, ela</i>	<i>ele, ela, lhe, Prep + ele, ela</i>
3ª pessoa pl.	<i>eles, elas</i>	<i>os/as, lhes, se, si, consigo</i>	<i>eles/eis, ela</i>	<i>eles/eis, elas, Prep + eles/eis, elas</i>

Fonte: Adaptado de Castilho (2020[2010], p. 477)

Como se pode observar, o sistema descrito por Castilho (2020[2010]) já se aproxima bem mais do português de fato utilizado no Brasil, considerando, inclusive, diferenças de registro. Como explicação para a entrada do “ele” acusativo no PB, o estudioso traz as considerações de Câmara Jr. (1972), segundo as quais o clítico, por ser comumente colocado em posição de próclise, forma uma palavra fonológica juntamente ao verbo, porém, por ser átono, passa pelo processo de aférese. Como resultado, o quadro pronominal passa a apresentar uma casa vazia, e o falante não encontra uma construção disponível para representar o OD, de modo que ou se atribui ao pronome “ele” o caso acusativo, ou se omite o pronome, gerando o objeto nulo.

Por isso, em todos os momentos em que são mencionadas as possibilidades para a realização do OD, Castilho (2020[2010]) já apresenta, ao lado do clítico, o pronome lexical “ele” e a categoria vazia. É o que ocorre na subseção dedicada ao OD, na qual se destaca que esse termo integrante da oração tem como propriedades: (i) ser proporcional aos pronomes acusativos “ele” e “o”; (ii) assumir a função de sujeito na passiva correspondente; (iii) poder ser preenchido por sintagmas nominais de núcleo pronominal ou nominal e por sentenças substantivas objetivas diretas; (iv) receber papel temático de /paciente/; e (v) poder ser omitido na sentença. Seriam exemplos, portanto, as frases (26) a (28) a seguir, retiradas de (Castilho, 2020[2010], p. 300-301)

(26) João pôs *o livro* na estante → João pôs *ele/o* pôs na estante.

(27) Disse **que viu o rapaz na rua.**

(28) Viu **Ø** na rua.

Castilho (2020[2010]) destina, ainda, uma parte de seu texto à discussão acerca das classes de preenchimento do OD anafórico e da categoria vazia, resenhando os trabalhos de Duarte (1986) e Cyrino (1994), seminais para o estudo desse tema, e discutindo os condicionamentos linguísticos e extralinguísticos que afetam essa estrutura. Posto que tais pesquisas serão abordadas na próxima subseção, voltada a estudos (sócio)linguísticos sobre o OD anafórico, basta, por ora, destacar que ambos os estudos consideram as quatro principais variantes desse fenômeno – SN anafórico, clítico, pronome lexical e objeto nulo –, e apontam, de maneira geral, para uma ampliação dos contextos que licenciam o OD nulo em detrimento do clítico acusativo, assim como para um crescente uso de pronomes lexicais em função de objeto.

Ademais, as possibilidades de expressão do OD anafórico são apontadas pelo linguista como pontos marcantes para se poder diferenciar o PB do PE, o PB do Norte e o do Sul, e o PB popular do culto. Sobre a primeira oposição, Castilho (2020[2010]) afirma que, no PB, tende a desaparecer o clítico acusativo “o”, o pronome “ele” pode funcionar como OD, e existe uma inclinação para se preencher o sujeito e apagar o objeto, ao passo que o PE mantém todos os seus pronomes átonos, emprega “ele” apenas como sujeito e elide, em geral, os sujeitos, mas expressa o OD. Quanto à segunda, salienta como marca do PB do Norte o OD expresso por “ele” e “lhe”, como em “eu não vi ele/eu não lhe vi”, enquanto, no PB do Sul, a preferência seria pela categoria vazia nessas funções, exemplificada por “eu não vi Ø, eu não conheço Ø”

(Castilho, 2020[2010]). Por fim, são características do PB popular, de acordo com o autor, a perda do pronome “o” e a omissão do OD, sendo que, quando há preenchimento dessa posição, ele é feito através de “ele” ou “lhe”; em contrapartida, ocorre no PB culto a perda de “o” na língua falada, mas mantém-se o clítico na escrita, e há emprego de objeto nulo em 70% dos casos, apesar de o pronome lexical “ele” aparecer na fala culta espontânea em proporções similares às da fala popular (Castilho, 2020[2010]).

Diante dos estudos apresentados nesta subseção, fica clara a relevância de estudos que se encarreguem da descrição linguística sem preocupação normativista, uma vez que, a partir deles, é possível obter uma visão mais abrangente das potencialidades da língua. Ademais, uma abordagem essencialmente descritiva evita purismos e conservadorismos arbitrários, validando, também na ciência, os usos que já são legitimados na realidade de todo falante.

3.3 O OD anafórico em pesquisas (sócio)linguísticas

Vários são os trabalhos que, ao longo dos anos, abordaram a variação observada na realização do OD anafórico em diferentes variedades do português, assim como fenômenos relacionados a essa construção, como a ascensão do objeto nulo e a queda dos clíticos de terceira pessoa. A partir de um ponto de vista sociolinguístico, cada um desses estudos procurou reconhecer e analisar os condicionamentos linguísticos e extralinguísticos que entram em jogo na escolha por uma ou outra variante do OD anafórico, adicionando uma nova dimensão às considerações traçadas em gramáticas e estudos descritivos em geral. Por isso, foram levantados, para o desenvolvimento desta pesquisa, os trabalhos dispostos no Quadro 3 à frente.

Quadro 3 - Estudos sociolinguísticos levantados para a pesquisa, organizados de acordo com autor, ano, título, modalidade e material analisado

Autor e ano	Título	Modalidade	Material
Omena (1978)	Pronome pessoal de terceira pessoa: suas formas variantes em função acusativa	Oral	Entrevistas sociolinguísticas com informantes do Mobral
Duarte (1986)	Variação e sintaxe: clítico acusativo, pronome lexical e categoria vazia no português do Brasil	Oral e escrita	Para a língua oral, entrevistas sociolinguísticas com falantes paulistanos, novelas, entrevistas televisivas. Para a língua escrita, textos de alunos do 2º grau, textos de jornais, revistas, propagandas e legendas de filmes

Cyrino (1994)	O objeto nulo no Português do Brasil: um estudo sintático-diacrônico	Escrita	Peças teatrais (preferencialmente comédias), cantigas, modinhas, poesia satírica e romances
Freire (2000)	Os clíticos de terceira pessoa e as estratégias para sua substituição na fala culta brasileira e lusitana	Oral	Entrevistas sociolinguísticas do projeto NURC-RJ de falantes portugueses com nível universitário
Marafoni (2004)	A realização do objeto direto anafórico: um estudo em tempo real de curta duração	Oral	Entrevistas sociolinguísticas da amostra Censo/PEUL com informantes dos anos 80 e 2000
Freire (2005)	A realização do acusativo e do dativo anafóricos de terceira pessoa na escrita brasileira e lusitana	Escrita	Textos jornalísticos do Rio de Janeiro e de Lisboa, além de quadrinhos publicados em ambos os países
Arruda (2006)	A realização do objeto direto no português brasileiro culto falado: um estudo sincrônico	Oral	Entrevistas sociolinguísticas do projeto NURC
Costa, S. (2011)	O não preenchimento do objeto anafórico na língua portuguesa: análise diacrônica do PB e do PE dos séculos XIX e XX	Escrita	Peças teatrais de autores catarinenses e lisboetas
Arruda (2012)	A realização do objeto direto anafórico em línguas românicas: um estudo sincrônico no português e no espanhol	Oral	Programas de auditório brasileiros, portugueses, argentinos e espanhóis
Costa, T. (2012)	Um estudo diacrônico das variadas realizações do objeto direto anafórico na imprensa baiana dos séculos XIX e XX	Escrita	Textos jornalísticos da imprensa baiana
Lauar (2015)	Não o vejo mais em Vitória: a substituição do clítico acusativo de terceira pessoa na fala capixaba	Oral	Entrevistas sociolinguísticas do projeto PortVix

Pivetta (2015)	Objeto direto anafórico no português brasileiro: uma discussão sobre a importância dos traços semântico-pragmáticos - animacidade/especificidade vs. gênero semântico	Escrita	Peças teatrais do século XX (as mesmas de Cyrino, 1994)
Pinto (2015)	Variação do objeto anafórico acusativo na fala de Florianópolis	Oral	Entrevistas sociolinguísticas do VARSUL e da Amostra Floripa
Santana (2016)	Diagnose e ensino de pronomes: um estudo sobre a retomada anafórica do objeto direto de terceira pessoa no Português brasileiro	Escrita	Redações de estudantes do 9º ano da rede pública de ensino
Rodrigues (2018)	O caso acusativo nos pronomes pessoais de terceira pessoa do português brasileiro e europeu	Oral e escrita	Entrevistas sociolinguísticas do <i>corpus</i> Concordância e textos jornalísticos do Rio de Janeiro e de Lisboa
Spinelli (2018)	Pronomes e sua ausência: por um tratamento unificado da omissão e da expressão de sujeitos e objetos diretos pronominais de 3ª pessoa	Oral	Entrevistas sociolinguísticas do VARSUL e do LínguaPOA
Marques de Sousa (2021)	As realizações do acusativo anafórico em variedades do português	Oral	Entrevistas sociolinguísticas com falantes do Brasil, de Portugal e de São Tomé e Príncipe
Lima (2022)	<i>Continuum</i> de gêneros textuais jornalísticos para a descrição de norma(s) culta(s): o acusativo anafórico de terceira pessoa e a ordem dos clíticos pronominais	Escrita	Textos jornalísticos do jornal O Globo

Fonte: Elaborada pela autora

Considerando os objetivos desta pesquisa, serão resenhados, a seguir, seis dos trabalhos mencionados no Quadro 3, por serem, cada qual à sua maneira, relevantes para o entendimento do fenômeno aqui em estudo, e por apresentarem resultados importantes para a discussão a ser desenvolvida nas próximas seções. São eles os estudos seminais de Omena (1978), Duarte

(1986) e Cyrino (1994), a tese de Freire (2005), e os trabalhos mais recentes encontrados sobre o assunto, isto é, Marques de Sousa (2021) e Lima (2022). Todos eles abordam uma ou mais variantes em jogo na retomada do acusativo anafórico de terceira pessoa, alguns a partir da perspectiva sincrônica e outros da diacrônica – o que permite a compreensão dos processos que ocorreram e ocorrem hoje na língua –, analisando dados advindos da modalidade oral e/ou escrita do português.

A dissertação de Omena (1978), primeiro trabalho acerca da realização do OD anafórico, parte de uma pesquisa preliminar empreendida pela autora, na qual, após serem analisadas quatro horas e meia de gravações de um universitário de Uberlândia, cinco fatores foram apontados como condicionadores do apagamento do objeto: traço [- animado] do antecedente, antecedente com função de complemento, função única do objeto, posição medial do OD na oração e distância pequena entre o objeto e o antecedente.

Em face desses resultados, assim como da ocorrência de apenas três clíticos na amostra, a pesquisadora iniciou uma nova investigação, a fim de averiguar se esse fenômeno estaria ligado a fatores extralinguísticos, especialmente à escolaridade dos falantes. Para tanto, elaborou um *corpus* composto de 24 horas de entrevista com quatro alunos do Mobral, isto é, falantes não escolarizados, sendo dois homens e duas mulheres, de idade, origem e classe social semelhantes. Com esse material em mãos, Omena (1978) coletou as ocorrências de três variantes utilizadas na expressão do OD anafórico: o clítico acusativo, o objeto nulo e o pronome lexical.

A distribuição geral dos dados apontou para uma preferência pelo objeto nulo, que correspondeu a 76% do total, enquanto os pronomes lexicais representaram os outros 24% da amostra. Como pode ser observado pelas porcentagens, não houve qualquer ocorrência de clítico na fala dos quatro informantes, o que, de acordo com a autora, pode indicar que o uso desse tipo de pronome não faz parte da gramática de falantes não escolarizados.

Quanto ao objeto nulo, a autora aponta os cinco fatores que condicionaram o apagamento em seu *corpus*: (i) traço [-animado] do antecedente; (ii) antecedente com função de complemento; (iii) função única do objeto; (iv) antecedente não reforçado; e (v) existência de mais de um candidato para o papel de antecedente. Percebe-se, portanto, que dois dos fatores elencados na pesquisa preliminar não foram novamente relevantes para o fenômeno. Os três que foram, entretanto, parecem reafirmar a influência que eles exercem na escolha de uma ou outra variante.

Dessa maneira, percebe-se que o estudo desenvolvido por Omena (1978) traz contribuições significativas para o entendimento da variação que atinge o OD anafórico. Por

ter sido o primeiro acerca desse tema, esse trabalho foi – e ainda é – retomado por diversos pesquisadores da área, seja no momento da seleção das variáveis a serem consideradas, no da formulação de hipóteses ou na análise dos resultados.

Um dos autores a fazer isso é Duarte (1986), que traça como objetivo para seu trabalho “[...] fazer uma pesquisa sociolinguística que apresente um quadro, o mais fiel possível, dos condicionamentos linguísticos, sociais e estilísticos que atuam nas formas pelas quais se realiza, na fala, o objeto direto anafórico [...]” (Duarte, 1986, p. 3). Ainda, a pesquisadora se propõe, também, a comparar os resultados obtidos com a língua escrita, para possivelmente tentar identificar mudanças mais profundas na língua.

Para realizar esse trabalho, Duarte (1986) elabora, então, um *corpus* para língua falada que conta com 40 horas de entrevistas sociolinguísticas com 45 informantes paulistanos ou residentes da cidade de São Paulo, estratificados em três graus de escolaridade e três faixas etárias, além de um grupo de jovens da 8ª série, para observar o comportamento do fenômeno em uma geração mais nova. Conta, também, com 4 horas de episódios de novelas e 4 horas de entrevistas televisivas, uma vez que esses materiais, de acordo com a autora, atingem toda a extensão do país e representam níveis diferentes de formalidade, exercendo uma força ao mesmo tempo inovadora e normalizadora. Já para a língua escrita, a autora compila uma amostra de 38 textos escritos por alunos do 2º grau, assim como textos de jornais, revistas, propagandas e legendas de filmes, apesar de esses últimos terem sido coletados de maneira menos sistemática, com vistas a observar aspectos relacionados à implementação.

Após coletar todas as ocorrências de OD anafórico com antecedente mencionado e que pudesse ser representado pelo clítico, Duarte (1986) define, então, as variantes de sua variável dependente, acrescentando uma àquelas já levantadas por Omena (1978). São elas: o clítico acusativo, o pronome lexical, o objeto nulo e o SN anafórico (incluindo SNs modificados e o uso do demonstrativo “isso”), exemplificadas, respectivamente, pelos exemplos (29) a (32):

(29) O custo da abreugrafia é muitas vezes menor, o que *a* torna capaz de ser usada com grande amplitude. (entrevista) (Duarte, 1986, p. 15)

(30) O senhor é a favor das diretas nas capitais porque acha *elas* mais democráticas? (entrevista) (Duarte, 1986, p. 15)

(31) Era um problema de tráfico também. Consumiram com o filho dele. Primeiro foi seqüestro e parece que eles não conseguiram pagar o resgate. Aí mataram (e). (fala de informante) (Duarte, 1986, p. 16)

(32) Então meu filho ficou morando no apartamento, mas ele reclamava muito do barulho e a gente foi na onda dele de vender *esse apartamento*. (fala de informante) (Duarte, 1986, p. 16)

De maneira geral, constatou-se uma preferência pelo objeto nulo, que correspondeu a 62,6% das ocorrências, seguido pelo SN anafórico (17,1%), pelo pronome lexical (15,4%) e, por fim, pelo clítico acusativo (4,9%). Como se pode observar, foram poucos os dados de clítico acusativo, mesmo no grupo de maior escolaridade, o que aponta para o desaparecimento dessa variante na fala, concomitante a um aumento no uso de objetos nulos.

Quanto aos fatores de natureza linguística, Duarte (1986) destaca, primeiramente, os aspectos morfológicos que influenciam a realização do OD anafórico, considerando, neste ponto da pesquisa, apenas as variantes clítico acusativo, pronome lexical e objeto nulo. Começando pela forma verbal, afirma que o objeto nulo parece não ser condicionado por qualquer forma, uma vez que, em todos os contextos, supera os usos das demais variantes. Entretanto, tal variável se mostra relevante para a ocorrência das outras variantes, especialmente para o pronome lexical, que é mais usado com o gerúndio, o imperativo e tempos simples, e o clítico, que nunca é usado com o gerúndio, o imperativo e os tempos compostos. Isso leva a autora a concluir que “[...] o uso do clítico pré e pós-verbal limita-se a formas já cristalizadas no português, não havendo dúvida de que é a forma verbal o fator que sustenta suas esporádicas ocorrências [...]” (Duarte, 1986, p. 23).

Passando à estrutura sintática da oração em que ocorre o OD, Duarte (1986) constatou que, em estruturas simples (SVO) e com objeto retomando um SN, a categoria vazia superou amplamente o preenchimento; por outro lado, quando o objeto era sentencial, o apagamento se mostrou quase categórico. Semelhantemente, em estruturas com OD e OI, o objeto nulo foi a estratégia mais expressiva na recuperação de um SN, enquanto, com objetos sentencias, o apagamento foi categórico. Por fim, em estruturas complexas nas quais o objeto funciona também como sujeito de uma minioração, observa-se uma tendência ao preenchimento, concorrendo para a realização desse OD o pronome lexical e a categoria vazia, em decorrência de questões semânticas envolvidas no processo.

Ainda pensando em aspectos linguísticos, Duarte (1986) analisa, também, a influência do traço semântico do antecedente, um fator de ordem semântica, na realização do OD anafórico. Sobre isso, a autora afirma que “[...] de todos os fatores linguísticos levantados, um deles parece ser extremamente importante na escolha da variante candidata à representação do objeto direto anafórico: o traço [± animado] do antecedente [...]” (Duarte, 1986, p. 26), posto

que, quando o antecedente é [+ animado], observa-se uma forte tendência ao preenchimento e, quando é [- animado], prevalece o uso do objeto nulo. Com efeito, ao cruzar os resultados para essa variável e para a estrutura sintática, a linguista encontra que é apenas nas estruturas simples que o traço [+ animado] deixa de ser decisivo para o preenchimento ou não do objeto, sendo que, em outros padrões, o mesmo traço favorece a realização fonológica do objeto tanto em orações com predicativo como nas complexas. Além disso, o traço [- animado] condiciona o apagamento em todos os contextos.

Adicionando os SNs às variantes possíveis para a expressão do OD anafórico, percebeu-se que favorecem esse uso o gerúndio, o traço [- animado] do antecedente, e a estrutura simples da frase. Tal resultado permite, de acordo com a autora, aproximar o uso das categorias vazias e dos SNs, afirmando que “[...] o uso de SNs lexicais plenos e do pronome demonstrativo apresenta-se como uma estratégia importante, só superada pelo uso da categoria vazia, com a qual partilha dos condicionamentos lingüísticos levantados [...]” (Duarte, 1986, p. 31).

Em relação aos fatores extralingüísticos, aqui abordados em menos detalhes por não figurarem entre as variáveis desta pesquisa, o que Duarte (1986) encontrou foi, em resumo, que o uso do objeto nulo ocorre preferencialmente entre todos os falantes, sendo que é no preenchimento do objeto que se encontram as maiores diferenças. Nesse caso, a pouca idade dos informantes resulta na não realização do clítico, enquanto a escolaridade mais alta leva a um aumento nos índices relacionados a essa variante, apesar de as porcentagens ainda serem bastante baixas. Quanto aos SNs anafóricos, a autora atesta que, com exceção da faixa etária e da escolaridade mais altas, os pronomes lexicais superam sempre essa forma, de modo que, “[...] os SNs plenos, além de condicionados pelos fatores sociais aqui considerados, se constituirão numa estratégia bastante significativa no exame de diferentes estilos em que se processa a fala [...]” (Duarte, 1986, p. 40).

Comparando a fala natural com a fala da mídia, Duarte (1986) encontrou uma semelhança entre as gravações com informantes e o texto das novelas, com ambos apresentando altos índices de apagamento, taxas moderadas de pronome lexical e SN, e pouquíssimas ocorrências de clíticos. Já para a fala das entrevistas televisivas, também representantes da fala da mídia, observou-se uma produtividade equilibrada entre a categoria vazia e os SNs, uso do clítico ligeiramente mais alto do que nas outras amostras (11,4%), e uma porcentagem irrisória de pronomes lexicais. Dessa maneira, a estudiosa propõe que “[...] o que difere basicamente estes dois estilos não é tanto a presença maior ou menor do clítico e sim a presença ou ausência do pronome lexical. A categoria vazia objeto e o uso de SNs são uma ‘saída’ eficaz para se evitar um ou outro [...]” (Duarte, 1986, p. 45).

Duarte (1986) realizou, ainda, testes de produção e de percepção, uma inovação para a época. Os resultados dessa análise permitiram à autora constatar que a formalidade do uso do clítico é percebida pelos falantes, que os falantes, excetuando-se os mais escolarizados, não têm consciência da categoria vazia com traço [- animado], e que, de maneira geral, a capacidade dos falantes de efetivamente utilizar o clítico acusativo é prejudicada em contextos de tempo composto, imperativo ou estrutura complexa. Além disso, a partir da opinião de professores, percebeu-se uma aceitação ampla do objeto nulo, mas uma aprovação parcial do clítico e do pronome lexical em construções específicas, sendo essas últimas variantes mais comumente percebidas em decorrência da saliência gerada pelo preenchimento da posição de OD.

Com relação aos dados da modalidade escrita, Duarte (1986) encontrou uma preferência pelos clíticos, seguidos pelos objetos nulos e, finalmente, pelos SNs anafóricos. Isso já demonstra uma diferença entre a língua fala e a escrita, já que, no PB oral, a variante mais utilizada pelos informantes foi, de longe, a categoria vazia, e as taxas encontradas para o uso do clítico foram consistentemente baixas. Apesar disso, alguns condicionadores linguísticos apontados como relevantes para a fala também foram atuantes na língua escrita, como a preferência pelo clítico especialmente no infinitivo, o favorecimento do OD nulo com gerúndio e tempos compostos e o traço semântico do antecedente. Quanto à complexidade sintática, porém, ao invés de estruturas mais complexas favorecerem o pronome lexical, como na fala, as estratégias selecionadas foram o clítico ou a categoria vazia.

Ademais, ao analisar textos jornalísticos, a autora observou que um recurso amplamente utilizado são os SNs lexicais com determinantes de várias naturezas, e que o clítico ainda se mostra resistente nesse contexto, podendo representar, inclusive, um objeto sentencial, uso já excluído da língua falada. Duarte (1986) afirma, também, que as ocorrências de objeto nulo nesse tipo de material apontam para o início da aceitação dessa variante em estilos mais formais, apesar de os poucos dados impedirem qualquer conclusão mais robusta.

Por fim, Duarte observa também que, em gêneros vistos como mais informais ou que tentam reproduzir a fala natural, é comum que ocorra a substituição do clítico por outras variantes, com vistas a obter maior naturalidade. É esse o caso, por exemplo, das legendas de filmes, em que as variantes não padrão são, de acordo com a autora, frequentes. Dessa maneira,

[...] a ocorrência da categoria vazia e do pronome lexical (este em menor escala) em textos escritos vem atestar a importância crescente do uso destas duas variantes quando se pretende imprimir certo grau de informalidade à língua escrita ou quando se busca reproduzir a língua falada. (Duarte, 1986, p. 57)

Assim, percebe-se que Duarte (1986) traz contribuições importantes para o estudo acerca da realização do OD anafórico no PB, principalmente quanto aos fatores linguísticos e extralinguísticos por ela levantados e pela introdução do SN anafórico como uma das variantes a ser observada no fenômeno. Além disso, a inclusão de dados advindos da escrita abriu espaço para mais pesquisas acerca desse objeto de estudo em materiais dessa modalidade.

É o caso do trabalho diacrônico de Cyrino (1994), que busca compreender, de maneira geral, como se deu o surgimento do objeto nulo no PB, relacionando essa mudança ao movimento de perda dos clíticos, através da análise de peças teatrais. Nesses trabalhos, é utilizada a abordagem da Sociolinguística Paramétrica²¹ – uma espécie de combinação entre a Teoria de Princípios e Parâmetros e a Teoria da Variação e Mudança Linguísticas –, de acordo com a qual toda mudança corresponde a uma alteração na fixação de um parâmetro, podendo ser levantados os condicionadores linguísticos e extralinguísticos que atuam nessa modificação.

A tese da autora surge após duas pesquisas anteriores: Cyrino (1990a) e Cyrino (1990b). Na primeira, realizada com uma amostra de peças teatrais populares dos séculos XVIII e XX, Cyrino (1990a) constata um aumento na ocorrência do objeto nulo ao longo do período, saindo de 14,2% na primeira metade do século XVIII para 81,1% na segunda metade do século XX. Tal emprego, restrito a alguns contextos em princípio, se tornou mais livre no decorrer dos séculos, e foi relacionado pela estudiosa à perda concomitante do clítico de terceira pessoa. A partir do século XIX, Cyrino (1990a) encontrou ocorrências de objeto nulo até mesmo em ilhas sintáticas, contexto que restringe a ocorrência dessa variante no PE, além de observar, também, o aparecimento do pronome lexical “ele” em função de objeto.

Na segunda pesquisa, realizada com o mesmo material, Cyrino (1990b) se propõe a investigar a mudança na posição dos clíticos em PB, “[...] a fim de verificar se houve uma mudança sintática que pudesse ser relacionada com a reanálise da categoria vazia em posição de objeto [...]” (Cyrino, 2018b[1993], p. 132). Após coletar e analisar 1000 dados, a linguista atestou (i) a perda da subida do clítico no PB para todas as pessoas do discurso; (ii) a queda da ênclise com imperativo afirmativo, infinitivo pessoal e sentenças com gerúndio, excetuando-se os casos de terceira pessoa; (iii) além de um decréscimo no uso da ênclise ao verbo principal ou auxiliar e na próclise ao auxiliar, aumentando, portanto, a utilização de próclise ao verbo principal. Ademais, a autora observou novamente os achados de Cyrino (1990a), constatando o movimento de queda do clítico acusativo e aumento da utilização do objeto nulo.

²¹ Para mais informações sobre a Sociolinguística Paramétrica, conferir Tarallo (1987) e Duarte (2016).

Diante desses resultados, novas perguntas surgiram para a pesquisadora, que a levaram a uma nova análise. Nesse estudo, Cyrino (1994) reuniu comédias dos séculos XVI a XX, assim como cantigas, modinhas, poesia satírica e romances de autores populares, buscando textos que apresentassem maior liberdade para retratar a linguagem da época, para que fosse possível acompanhar o surgimento do objeto nulo no PB, relacionando esse fenômeno à perda dos clíticos de terceira pessoa.

Coletados os 2308 dados, Cyrino (1994) observou, novamente, um aumento de objetos nulos no período, sendo que todos os séculos apresentaram dados para essa variante, havendo diferença, porém, entre o número de ocorrências para cada século. Pensando no aumento da incidência dessa estrutura, a estudiosa observou que a categoria vazia começou a ganhar força como opção para a realização do OD quando, em lugar de um clítico, fazia referência a antecedentes oracionais, como exemplificado em (33):

(33) O caso he este; dir-vo-lo-hei!” (Camões, *Fiolodemo*, p. 145) (Cyrino, 1994, p. 3)

Tendo caído esse clítico neutro, a expansão do objeto nulo continuou com a perda do clítico “o” com referentes SNs de traço [- animado], aparentemente meio século após a primeira mudança, até que a categoria vazia chegasse a substituir referentes em ilhas sintáticas e/ou com traço [+ animado], sendo a animacidade a variável mais importante para a consolidação do processo aqui descrito. A partir dos resultados, Cyrino (1994) afirma, então, que, após o início desse percurso de mudança, o que aconteceu foi uma generalização da regra que permitia a utilização do objeto nulo em lugar do clítico “o” para antecedentes sentenciais, essencialmente [- animados], atingindo todos os outros contextos em que os clíticos acusativos poderiam ser utilizados. Dessa maneira, enquanto os falantes dos séculos XVI e XVII preferiam o clítico, aqueles dos séculos seguintes já tendiam ao uso da categoria vazia.

Ademais, Cyrino (1994) levanta a hipótese de que o surgimento do pronome lexical e da ocorrência de SNs em posição de objeto também estaria relacionado à queda dos clíticos e ao aumento dos nulos. Sobre isso, explica que, uma vez que o clítico “o” com traço [- animado] desaparece e é substituído pelo objeto nulo, os clíticos masculinos [+ animados] também poderiam deixar de ser empregados. Conseqüentemente, as únicas opções para o preenchimento da posição de OD restantes ao falante seriam o uso do pronome “ele”, a retomada com o pronome demonstrativo “isso” e a repetição de um SN (com ou sem modificações), motivo pelo qual essas variantes começariam a figurar com mais força do século XIX em diante.

Em vista disso, Cyrino (1994) chega à conclusão de que, de fato, existe uma relação entre a diminuição no uso de clíticos e o aumento nas ocorrências de objetos nulos – assim

como o aparecimento das variantes pronome lexical e SN anafórico –, sendo as mudanças por ela descritas decorrentes de uma mudança paramétrica na qual o PB remarcou o parâmetro relacionado ao OD, deixando de ser uma língua com objetos pronominais, realizados através dos clíticos, para assumir as características de uma língua de objetos nulos. Nas palavras da autora,

[...] através da análise de dados diacrônicos, há motivos para se supor que a reanálise que levou ao objeto nulo do PB estaria relacionada às reanálises diacrônicas que levaram à mudança no sistema de clíticos dessa língua. Todas essas mudanças teriam tido como consequência, então, a refixação do parâmetro relacionado ao objeto nulo [...] (Cyrino, 2018b[1993], p. 139)

Percebe-se, portanto, que o trabalho de Cyrino (1994) ampliou largamente a visão que se tinha acerca da realização do OD anafórico, trazendo à tona informações diacrônicas importantes sobre o percurso linguístico que levou ao cenário de variação encontrado atualmente. O estudo, inclusive, serviu de suporte para a proposta de uma escala de referencialidade, elaborada por Cyrino, Duarte e Kato (2000), que influenciou diversos autores a pensar a referencialidade como uma variável independente possivelmente influente na expressão do fenômeno em questão.

Dentre eles, figura Freire (2005), que, tendo anteriormente analisado a língua oral, se volta à modalidade escrita do português, pensando a expressão do acusativo e do dativo anafóricos de terceira pessoa em duas variedades: o PB e o PE²². Para realizar essa pesquisa, o autor reuniu excertos de jornais que representavam gêneros tanto da concepção oral como da escrita e histórias em quadrinhos, sendo a amostra brasileira composta por textos coletados do *Jornal do Brasil*, *d'O Globo* e de gibis da *Turma da Mônica* e da *Disney*, entre 1995 e 2004.

Com esse material em mãos, Freire (2005) distribuiu os textos ao longo de um *continuum* de oralidade-letramento, de acordo com a metodologia delineada por Bortoni-Ricardo (2004) a fim de testar a hipótese de que a distribuição das variantes seria diferente a depender da porção do *continuum* em que o gênero materializado por um texto se encontrava. Dividiu, então, o *continuum* em três partes correspondentes a: (i) [+ oralidade/- letramento], onde encaixou tirinhas e histórias em quadrinhos; (ii) [+ oralidade/+ letramento], incluindo crônicas, falas transcritas em reportagens, e entrevistas; e (iii) [- oralidade/+ letramento], que abarcou editoriais, reportagens, críticas e artigos de opinião.

²² Os resultados obtidos pelo autor para o PE não serão aqui abordados por fugirem do escopo desta pesquisa, uma vez que ela versa somente sobre dados advindos do PB.

Em seguida, o pesquisador coletou todos os dados referentes às quatro variantes consideradas por ele, a saber, o clítico, o objeto nulo, o pronome lexical e o SN anafórico, chegando à seguinte distribuição geral para o PB: 47% de clíticos, 31% de ODs nulos, 15% de SNs anafóricos e 8% de pronomes lexicais. Frente a essas porcentagens, o autor destaca que, pelo fato de as três variantes alternativas ao clítico somadas representarem mais da metade do total, essas estratégias comuns na fala já estão plenamente infiltradas na língua escrita, especialmente o objeto nulo.

Continuando sua análise, Freire (2005) observa que o clítico é a única variante cujos índices aumentam ao longo do *continuum*, sendo mais frequente em textos alocados na porção de mais letramento. Movimento contrário demonstra o objeto nulo, que é mais utilizado na outra ponta do *continuum*, assim como o pronome lexical, que chega a desaparecer em gêneros representativos de eventos comunicativos que envolvem maior letramento, enquanto o uso do SN anafórico se mostra bastante estável em todas as partes do *continuum*. Destaca também que, para os textos das duas primeiras porções do *continuum*, correspondentes a eventos de concepção [+oralidade], a utilização do clítico se apresenta mais próxima às características da fala, posto que as outras variantes somadas superam as taxas de uso da variante padrão. Diante desses resultados, o autor referencia o trabalho de Averborg (2000), que demonstrou que o clítico cresce à medida que aumentam a escolaridade e o contato com a cultura letrada, assim como aponta para uma influência da tradição gramatical sobre a escrita, coibindo, por exemplo, o uso do pronome lexical em função de objeto, apesar de não chegar a ser absoluta, uma vez que a variante padrão – o clítico – aparece em competição com outras formas mesmo em gêneros que ocorrem em eventos de maior letramento.

Passando aos condicionamentos linguísticos que atuam na realização do OD anafórico, Freire (2005) se ocupa, primeiramente, da natureza do antecedente do objeto. Nesse contexto, o linguista constata que o clítico é favorecido quando o antecedente é um SN, enquanto, para antecedentes oracionais e predicativos anafóricos, as taxas de emprego dessa variante caem consideravelmente, em favor do SN anafórico e do objeto nulo. Acerca dessa última estratégia, aponta, ainda, que a categoria vazia, além de exibir altos índices nos contextos mencionados, corresponde a mais de um quarto dos dados em que o antecedente é nominal, novamente provando sua relevância como possibilidade de expressão do acusativo anafórico de terceira pessoa também em textos escritos.

De acordo com Freire (2005), é apenas com formais verbais simples não finitas, especialmente com o infinitivo, que o clítico supera a soma das outras. Já em contextos de formas verbais simples flexionadas e de formas verbais complexas marcadas com tempo, as

demais estratégias são favorecidas, destacando-se, mais uma vez, a categoria vazia. Sobre isso, o autor afirma que, “[...] em linhas gerais, a manutenção do pronome acusativo que ainda se percebe no PB está fortemente relacionada à presença de um contexto que permita ao clítico recuperar o *onset* silábico [...]” (Freire, 2005, p. 127), recuperando a hipótese de Nunes (2018[1993]), segundo a qual a mudança na direção de cliticização no PB dificultou a aquisição dos clíticos na língua atual, resultando na diminuição da produtividade dessa variante observada na fala.

Investigando a influência da transitividade verbal, Freire (2005) encontra uma preferência pelo clítico em estruturas sintáticas simples com ODs que retomam um SN, ao passo que, com estruturas que projetam um acusativo e um dativo, essa variante é desfavorecida em relação às outras. Quando se trata de estruturas simples com acusativo sentencial ou predicativo, o objeto nulo apresenta índices altos de utilização, atingindo, respectivamente, 70% e 86%. Entretanto, a constatação mais interessante do autor consiste no fato de que, nas estruturas complexas com verbos causativos ou perceptivos em que o objeto exerce também a função de sujeito de uma oração infinitiva, estão em competição o pronome lexical e o clítico acusativo, com a variante não padrão ocorrendo mesmo em textos que apresentam traços de letramento. A partir disso, o estudioso conclui, então, que “[...] tais resultados sinalizam que, na escrita brasileira, já não parece ser estigmatizado o pronome lexical que figura em estruturas sintáticas projetadas pelos verbos tradicionalmente chamados causativos e perceptivos [...]” (Freire, 2005, p. 134).

Quanto à realização do acusativo anafórico em ilhas sintáticas, Freire (2005) relata, como esperado, uma maior produtividade do preenchimento, sobretudo através do clítico. Chama a atenção, porém, para o fato de que o objeto nulo representa pouco mais de um quarto do total de ocorrências de OD anafórico em ilhas sintáticas, retomando tanto antecedentes sentenciais como nominais, podendo “[...] aparecer livremente com qualquer tipo de oração, com antecedente humano ou não-humano, sendo a mais expressiva de todas as variantes alternativas ao clítico [...]” (Freire, 2005, p. 140). Isso, de acordo com o autor, reforça o caráter pronominal da variante na língua já apontado por Cyrino (1994) em sua tese, uma vez que essa variante não sofre restrições no PB.

Consoante a outros estudos, Freire (2005) também analisou a relevância do traço semântico do antecedente para a realização do OD anafórico de terceira pessoa. Comentando, a princípio, acerca do uso do clítico, o linguista afirma que antecedentes com traço semântico [+ referencial; + animado] favorecem a retomada por essa variante na escrita, sendo ela a estratégia preferida no PB. Outra forma favorecida pelo traço [+ animado] é o pronome lexical,

que apresentou o maior índice de ocorrência entre as alternativas ao clítico. Em contrapartida, quando os antecedentes possuíam traço [+ referencial; - animado], a variante tida como padrão foi desfavorecida, sendo mais expressivo o uso da categoria vazia, que também foi a forma preferida para antecedentes de traço [- referencial]. Dessa maneira, verifica-se, de acordo com o autor, que a implementação do objeto nulo no PB obedece a uma hierarquia referencial, assim como proposto por Cyrino, Duarte e Kato (2000), com mais ocorrências dessa variante quando o traço é [- referencial] e menos quando é [+ referencial; + animado].

Por fim, foi realizada uma análise multivariada, na qual foram selecionadas como variáveis relevantes o *continuum* de oralidade/letramento, o traço semântico do antecedente e a forma verbal. Isso levou Freire (2005) a concluir que o clítico resiste no PB apenas nos seguintes contextos: (i) em eventos de maior letramento; (ii) quando retoma antecedentes de traço animado; e (iii) quando realizado através do alomorfe “lo”, já que é comum que ele ocorra com formas verbais simples não flexionadas em ênclise, sendo “[...] o pronome acusativo um item lingüístico típico do processo de letramento, visto que não costuma estar acessível no processo natural de aquisição da linguagem [...]” (Freire, 2005, p. 147).

Chegando, agora, aos trabalhos mais recentes sobre o objeto de estudo desta pesquisa, será comentado o trabalho de Marques de Sousa (2021). Nele, propõe-se a investigação acerca das realizações do acusativo anafórico na modalidade oral do PB, do PE, do português de Moçambique (PMO) e do português de São Tomé e Príncipe (PST)²³, por meio de entrevistas sociolinguísticas estratificadas segundo gênero, faixa etária e nível de escolaridade.

Antes de continuar, vale mencionar, contudo, que, diferentemente dos autores anteriormente mencionados, Marques de Sousa (2021) considera apenas três variantes para a expressão do OD anafórico: o clítico, o pronome lexical e o objeto nulo. A escolha por descartar o SN anafórico como estratégia possível surge do fato de que “[...] tal categoria nominal desencadeia a recategorização do referente, além de indicar a atitude do falante, atuando mais como uma estratégia textual-discursiva do que propriamente uma variante gramatical [...]” (Marques de Sousa, 2021, p. 1-2).

Com isso em mente, o primeiro aspecto observado pelo autor é a natureza dos antecedentes dos objetos, separando-os em antecedentes oracionais e não oracionais. Acerca do primeiro grupo, Marques de Sousa (2021) encontra o uso categórico do objeto nulo, independentemente das outras variáveis, fato que explica retomando Cyrino, Duarte e Kato

²³ Mais uma vez, serão resenhados apenas os resultados referentes ao PB, pois é essa a variedade de interesse do presente estudo.

(2000) e afirmando que, pelas orações serem constituintes de baixa referencialidade, elas acabam configurando o contexto propício para a utilização da categoria vazia.

Já com relação ao segundo grupo, o linguista encontrou uma distribuição semelhante àquela relatada em outras pesquisas: uma preferência pelo OD nulo (62,2%), seguido pelo pronome lexical (32,9%) e, finalmente, pelo clítico (5%). Para o baixo número de dados de clítico, apresenta-se, assim como em Freire (2005), as conclusões de Averborg (2000), que enfatiza que o clítico é uma variante aprendida de maneira tardia e em contato com a escola, resistindo apenas em alguns contextos relacionados à forma verbal. Por outro lado, no que toca ao índice expressivo de pronomes lexicais, Marques de Sousa (2021, p. 115) sugere que ele pode estar relacionado à “[...] gradual – mas não absoluta – perda de estigma da variante e à expansão de seus contextos de realização: ainda que mais raro, a variante pode fazer referência a elementos com traços [-humano/-animado] [...]”. Como prova dessa expansão, é apresentada, dentre outras, a seguinte passagem:

(34) Eu, em vez de lavar [**a fruta**]i, eu peguei [**aquela fruta**]i e ingeri [**ela**]i na boca. (Marques de Sousa, 2021, p. 116)

Após essas considerações, o pesquisador realizou rodadas de análises multivariadas que colocaram o objeto nulo em contraposição às demais variantes. Como resultado, os seguintes fatores foram selecionados como significativos para a realização da categoria vazia: (i) a combinação de traços semânticos do antecedente; (ii) a estrutura projetada pelo predicador; (iii) o papel semântico do antecedente; e (iv) o padrão estrutural e a função do antecedente.

No que se refere à primeira variável selecionada, Marques de Sousa (2021) observou que o objeto nulo foi favorecido quando o antecedente exibia o traço [-animado/- específico], atingindo 95,3% das ocorrências e peso relativo de 0,867, assim como quando o traço era [-animado/+ específico], caso em que correspondeu a 90,6% dos dados e exibiu peso relativo de 0,767. Em contraposição, o preenchimento da posição de OD, seja com pronome lexical ou com clítico acusativo, foi favorecido pelo traço [+ animado]. É interessante mencionar que o autor encontrou, ainda, dados que corroboram o uso de um objeto nulo com antecedente [+animado], desde que ele seja tópico do discurso anteriormente expresso, como no exemplo (35) abaixo, o que indica uma expansão, também, nos contextos em que a categoria vazia é licenciada.

(35) Essas pessoas que têm quatro, cinco filhos, não dão atenção que [**as crianças**]i deveriam. Deixam [**Ø**]i nas ruas. (Marques de Sousa, 2021, p. 122)

Com respeito à estrutura projetada pelo predicador, o estudioso constata, de antemão, o favorecimento do objeto nulo em padrões oracionais simples cujo antecedente é um SN, confirmando o que já fora previamente apontado por Duarte (1986). Além disso, descreve que contextos de minioração desfavorecem ligeiramente essa variante, com peso relativo de 0,419, apesar de a categoria vazia tender a aparecer mais se empregada em uma oração com topicalização. Acerca desse mesmo padrão, Marques de Sousa (2021) observa que, quando o objeto retoma um antecedente de traço [+ animado], porém, prefere-se o preenchimento da posição, principalmente com o pronome lexical. Essa mesma preferência pelo “ele” acusativo também foi identificada em contextos com verbos causativos, sensitivos e de permissão da oração matriz, como também já havia relatado Duarte (1986).

Quanto à relevância do papel semântico do antecedente, os resultados de Marques de Sousa (2021) indicam que os papéis de paciente e tema favorecem o objeto nulo, e os de experienciador, agente, e beneficiário/alvo/recipientes favorecem as demais variantes. Para explicar essa distribuição, o linguista afirma que os papéis de paciente e tema são frequentemente atribuídos a constituintes que exercem função de complemento ou sujeito derivado de um verbo inacusativo, enquanto os demais são comumente atribuídos a sujeitos e entidades com traço [+ humano], motivo pelo qual, nesse caso, é favorecido o uso de pronome lexical.

Sobre a questão do padrão estrutural e a função do antecedente, Marques de Sousa (2021) observa que as estruturas que mais favoreceram o objeto nulo, com peso relativo de 0,914, foram as de tópico marcado, seguidas pelas subordinadas com antecedente em função de sujeito, as coordenadas com antecedentes em função de complemento, as de tópico discursivo em contexto precedente, as subordinadas com antecedentes em função de sujeito e, enfim, as coordenadas com antecedentes sujeitos. Diante disso, o pesquisador esclarece que os resultados confirmam a proposta de Cyrino (2018a) acerca do paralelismo de estruturas, segundo a qual “[...] o objeto nulo é favorecido por antecedentes com função de complemento, principalmente se for [-animado] [...]” (Marques de Sousa, 2021, p. 127), independentemente do padrão estrutural.

Ao final da análise, Marques de Sousa (2021) conclui que os traços [- animado/- específico] de fato exercem uma força considerável sobre a utilização do objeto nulo, sendo que “[...] as restrições semânticas e sintáticas podem ser sobrepujadas, quando o antecedente está fora do período do nulo e o discurso oferece condições para a sua realização [...]” (Marques de Souza, 2021, p. 131). Assim, o estudioso confirma os condicionamentos anteriormente apontados por outros autores a partir de um novo e atualizado *corpus*, de modo a reafirmar a

relevância de certas variáveis para a realização do OD anafórico de terceira pessoa no PB oral atual.

Finalmente, será aqui explorado o trabalho de Lima (2022), o mais recente encontrado acerca do objeto de estudo desta pesquisa. Nele, a autora se propõe a descrever a variação que atinge o acusativo anafórico em textos jornalísticos, buscando observar a heterogeneidade da norma culta, a partir de um ponto de vista que considera a interface entre variação, gênero textual-discursivo e norma. Para empreender suas análises, compõe, então, um *corpus* com sete gêneros do jornal O Globo, a saber, artigo, carta de leitor, crônica, editorial, entrevista, notícia e tirinha, preferencialmente de autores do Rio de Janeiro, e que fossem de concepção oral e/ou escrita.

Um diferencial trazido por Lima (2022), em comparação aos outros trabalhos aqui mencionados, é o olhar mais cuidadoso da pesquisadora na caracterização de cada gênero selecionado para compor seu *corpus*. Indo além da separação dos gêneros entre eventos de concepção oral e/ou escrita, como proposto por Marcuschi (2010[2001], 2008), a estudiosa lança mão da teoria de Biber e Conrad (2009) acerca dos parâmetros situacionais para descrever em mais detalhes os contextos, propósitos comunicativos e aspectos que entram em jogo quando um gênero é manifestado através de um texto. Desse modo, Lima (2022) obtém uma visão mais ampla do material a sua disposição, podendo dispor os gêneros em um *continuum* de oralidade-letramento com maior segurança e traçar considerações mais aprofundadas e embasadas acerca da maneira como a língua é afetada pelo gênero na qual se materializa, assim como é pretendido nesta pesquisa.

Após a detalhada descrição dos gêneros e a coleta de todas as ocorrências para as quatro variantes consideradas pela autora – novamente, clítico acusativo, objeto nulo, SN anafórico e pronome lexical –, Lima (2022) chega aos seguintes resultados gerais: dos 538 dados encontrados, 38 corresponderam a ODs sentenciais, e 500 se referiram a antecedentes nominais. Para os oracionais, houve retomadas com clítico, categoria vazia e SN (incluindo o pronome demonstrativo “isso”), sendo a maior parte das ocorrências de objetos nulos (58%). Chama a atenção a resistência do clítico nesse contexto, uma vez que esse uso já desapareceu da fala, o que a autora relaciona com o fato de os dados serem advindos da modalidade escrita. Já para os objetos com antecedente SN, que foram o foco da análise, foram encontrados dados para as quatro variantes, porém com porcentagens bastante diferentes: 41% de SNs anafóricos, 40% de clíticos, 17% de objetos nulos e 2% de pronomes lexicais.

Explicando essa distribuição, Lima (2022) associa a alta incidência dos SNs ao fato de que essa é considerada uma estratégia importante para os processos de referência em um

texto escrito, além de ser uma forma que parece não carregar estigma, enquanto, para os clíticos, esclarece que é essa a forma preferida pelos compêndios gramaticais para expressão do OD anafórico, sendo mais associada ao maior domínio da modalidade escrita, motivo pelo qual figuraria tão fortemente entre os textos de seu *corpus*. Já quanto às duas estratégias menos utilizadas, a linguista menciona que o objeto nulo, forma inovadora já bem estabelecida no PB, apesar de não ter apresentado altas taxas de aplicação, se fez presente em todos os textos do *corpus*, ao passo que o mesmo não aconteceu com o pronome lexical, que só apareceu em tirinhas, muito provavelmente por conta da estigmatização dessa forma na modalidade escrita.

É interessante notar, ainda, que as variantes se distribuíram de maneira distinta, tanto entre os gêneros textuais em si como entre os diferentes pontos do *continuum*. Nesse contexto, o SN anafórico demonstrou comportamento irregular, tendo apresentado porcentagens abaixo de 50% nos extremos e acima disso na área central. Por outro lado, tanto o clítico acusativo como o objeto nulo ocorreram de maneira escalar, tendo o primeiro aumentado de 24% para 56% do polo de [+ oralidade] para o de [+ letramento] e, o segundo, diminuído de 31% para 8% no mesmo sentido.

Ao tratar da animacidade, Lima (2022) observa que antecedentes de traço [+ animado] foram preferencialmente retomados por variantes que realizavam o preenchimento da posição, sendo somente 5% das ocorrências correspondentes ao objeto nulo. Já no que diz respeito ao traço [- animado], o SN anafórico e o clítico também foram as formas mais expressivas, e não a categoria vazia, um resultado que parece ir contra a tendência anteriormente reportada de antecedentes com esse traço favorecerem o não preenchimento. Apesar disso, quando analisadas apenas as ocorrências de objeto nulo, percebe-se que a maioria delas ocorreu, de fato, em referência a antecedentes de traço [- animado], reforçando a influência dessa variável.

Com respeito à variável função sintática do antecedente, a autora considerou dois níveis: igual ou diferente do objeto. Como resultados, Lima (2022) atestou que, quando a função sintática foi diferente, houve uma utilização expressiva de SNs anafóricos (40%) e clíticos acusativos (45%), e índices bem menores de objetos nulos e pronomes lexicais (14% e 1%, respectivamente). Em contrapartida, quando a função sintática foi igual, houve uma diminuição na proporção de uso do clítico, que caiu para 28%, e um aumento na de uso da categoria vazia, atingindo 25% das ocorrências, o que comprova resultados anteriores que sugeriram que o objeto nulo é mais comum quando seu antecedente também é um OD.

No que tange à estrutura sintática, percebeu-se que a variante mais utilizada, tanto em estruturas com minioração como com complementos após o OD, foi o clítico acusativo. Em contraste com a distribuição geral, o SN anafórico só foi a forma preferida em contextos de

estrutura S V OD, sendo também com esse padrão que ocorreu o maior número de objetos nulos. Embora pouco expressivos, é importante notar, também, que o pronome lexical ocorreu mais quando em contexto de miniorações, o que, segundo Lima (2022), corrobora estudos anteriores que afirmaram que, quando o objeto exerce também a função de sujeito de uma oração, é essa a variante favorecida, sendo esse uso menos estigmatizado.

Para a forma verbal, o estudo de Lima (2022) contou com três níveis, resultantes de amálgamas, sendo eles infinitivo, forma flexionada e forma não finita. Nesse cenário, os contextos com verbo no infinitivo, mais recorrentes na amostra, geraram, assim como esperado, a metade dos objetos sendo realizados através do clítico. Já com formas flexionadas, houve mais ocorrências com SN anafórico (47%) e a categoria vazia (30%), assim como nos poucos dados encontrados com formas não finitas, com 53% e 23%, respectivamente.

Visando compreender, com que força cada uma das variáveis de fato atuava na realização do OD anafórico de terceira pessoa, Lima (2022) realizou, então, duas análises multivariadas com conjuntos distintos. A primeira, uma análise multinomial, foi executada com os dados de clítico, SN e objeto nulo, excluindo-se os pronomes lexicais, posto que, por aparecerem apenas nas tirinhas, poderiam enviesar os dados. Nela, foram consideradas as variáveis relacionadas ao *continuum*, ao traço semântico e à forma verbal.

Através da análise multinomial, observou-se que havia uma diferença significativa entre as porcentagens de distribuição das diferentes variantes ao longo do *continuum* proposto. Como previamente hipotetizado, o clítico foi a variante favorecida no extremo de [+ letramento], dada a clara relação com as orientações normativas, enquanto as demais formas foram desfavorecidas. Já nas outras duas porções, Lima (2022) atestou a maior probabilidade de se encontrarem SNs anafóricos e objetos nulos, como indicado pelos valores de 0,923 e 1,510 *logodds* para essas variantes, respectivamente. Tais resultados são explicados pela autora ao relacioná-los às noções de norma, sendo que os textos do polo de maior letramento seriam mais propensos a apresentar formas validadas pela gramática tradicional, em contraste com aqueles do caracterizados como de maior oralidade, que poderiam apresentar uma maior liberdade para a realização das escolhas linguísticas.

Quanto ao traço de animacidade, a autora atestou o favorecimento do clítico frente a antecedentes com traço [+ animado], enquanto aqueles de traço [- animado] favoreceram o SN anafórico e o objeto nulo. Frente a esse último resultado, vale mencionar ainda que, em comparação com o clítico, o objeto nulo foi mais favorecido do que o SN, uma vez que o primeiro apresentou uma estimativa de 2,199 *logodds* e o segundo de 0,589 *logodds*. Isso ratifica a hipótese de que o uso da categoria vazia estaria relacionado ao traço [- animado],

levantada com base na literatura acerca do fenômeno, posto que, dentre as três formas, é essa a mais favorecida nesse contexto. Sobre isso, a estudiosa afirma que esse favorecimento do nulo “[...] pode ser explicado pelo fato de que a mudança linguística que implementou essa forma na fala brasileira começou justamente na retomada de antecedentes com traço [-animado] possível, que era o caso do clítico neutro [...]” (Lima, 2022, p. 211). Além disso, ela também destaca que o fato de o objeto nulo não ter sido tão frequente em seu *corpus* de PB escrito, ao contrário do que se reporta acontecer na língua oral, já aponta para uma diferença entre essas duas modalidades.

Fechando a análise multinomial, Lima (2022) observou que as tendências pressentidas ao investigar as proporções e frequências de cada variante quando relacionadas à forma verbal foram confirmadas. Dessa maneira, o modelo estatístico demonstrou que as formas de infinitivo favoreceram o emprego do clítico, mas, com as demais formas verbais – flexionadas e não finitas, amalgamadas em um só nível –, as estratégias favorecidas foram o SN anafórico e a categoria vazia.

Já a segunda análise multivariada consistiu em um modelo binomial que considerou as mesmas variáveis acima elencadas, excluindo, porém, as ocorrências de SN anafórico, uma vez que, de acordo com Lima (2022, p. 212), essa variante “[...] não parece ser sensível à variação estilística observada no *continuum*”. Com isso, Lima (2022) teve como objetivo verificar a influência das variáveis linguísticas quanto ao preenchimento ou não da posição de objeto com a variante padrão determinada pelas gramáticas, dado que, segundo a autora, seria plausível pensar que, nos textos de sua amostra, “[...] “deixar de apagar a posição do objeto direto anafórico” signifique maior consciência dessa posição, na contramão do que ocorre na modalidade oral [...]” (Lima, 2022, p. 213).

Em resumo, os resultados para o novo modelo indicaram que a primeira variável a atuar significativamente na utilização do clítico acusativo é o *continuum*, seguida pelo traço semântico do antecedente e a forma verbal. Com base nos valores gerados pelo programa estatístico, essa variante, assim como esperado, foi favorecida no extremo de maior letramento do *continuum*, com antecedentes de traço [+ animado] e infinitivo, corroborando os condicionamentos levantados por pesquisas anteriores. Isso indica que “[...] o clítico acusativo, apesar de praticamente ausente na fala brasileira, é uma das variantes preferidas da modalidade escrita e é favorecido nos gêneros aos quais se atribui maior monitoração estilística [...]” (Lima, 2022, p. 215).

Percebe-se, portanto, que o trabalho de Lima (2022), assim como o Marques de Sousa (2021), apresenta um *corpus* com materiais atuais, de maneira que se torna possível averiguar

o estado atual da norma culta escrita no PB, confirmando que os mesmos contextos apontados por estudos mais antigos continuam sendo relevantes para o fenômeno. Ademais, a preocupação da autora em se aprofundar na caracterização dos gêneros por ela considerados apresenta uma nova dimensão à sua pesquisa, reafirmando a importância de um olhar mais atento para os gêneros textuais-discursivos ao analisar processos de variação e mudança linguísticas.

Em suma, Omena (1978) e Duarte (1986) produzem as primeiras pesquisas acerca do tema no PB, destacando variáveis que até hoje se mostram atuantes na variação que cerca o fenômeno. Cyrino (1994), em sua perspectiva diacrônica, destaca o desenvolvimento das variantes ao longo do tempo, proporcionando uma visão mais abrangente do percurso percorrido pela língua, tendo sido, também, o primeiro estudo a trabalhar extensivamente com textos escritos para a investigação da realização do OD anafórico. Freire (2005), por sua vez, apresenta indícios de uma atenção maior voltada ao gênero em relação à variação, provando a validade de se utilizar *continua* para distribuir os gêneros. E, por fim, Marques de Sousa (2021) e Lima (2022) apresentam um panorama do estado atual de ambas as modalidades da língua, unindo aspectos levantados pelas inúmeras pesquisas que os precederam.

Assim, os trabalhos apresentados nesta subseção trazem, cada qual à sua maneira, contribuições extremamente relevantes para o estudo acerca do OD anafórico de terceira pessoa e, certamente, para o estudo que aqui se deseja empreender. Foi com base neles que foram tomadas as decisões metodológicas descritas a seguir, especialmente com relação às variáveis independentes a serem analisadas nesta pesquisa, e será também a partir deles que os comentários acerca dos resultados serão traçados.

4 A TRADUÇÃO E A LEGENDA AUDIOVISUAL

Em um estudo como este, que se propõe a trabalhar com a variação linguística do OD anafórico de terceira pessoa em legendas audiovisuais, é imprescindível que se tracem reflexões acerca da tradução como atividade, além de considerar as características do processo envolvido na tradução desse tipo de material. Dessa maneira, a presente seção visa apresentar alguns conceitos básicos sobre tradução, contextualizar a área da TAV, e detalhar aspectos específicos da produção de legendas.

4.1 Conceitos básicos de tradução

Quando se fala em tradução, uma das primeiras tarefas do pesquisador é determinar a que atividade ele, de fato, se refere. Isso porque, classicamente, distinguem-se três tipos de tradução: a intralingual, a interlingual e a intersemiótica, assim como definidas por Jakobson (1966)²⁴. Nesse contexto, o primeiro tipo de tradução apontado pelo autor, também chamado de “*rewording*” em inglês, consiste na utilização de sinônimos e outros artifícios para reformular um enunciado, isto é, para realizar a interpretação de signos através de outros na mesma língua. Seria esse o processo empregado ao transformar pensamentos em fala verbalizada ou ao parafrasear um trecho de um texto.

O segundo tipo, ou seja, a tradução no sentido comumente atribuído à palavra, se refere à interpretação e à passagem de um texto em uma língua de partida (LP) – também chamada de língua-fonte (“*source language*”) – para uma língua de chegada (LC) ou língua-alvo (“*target language*”). Dois exemplos dessa atividade seriam a tradução de um romance em outra língua para o português e as próprias legendas do inglês para o português que serão analisadas neste trabalho.

Por fim, a tradução intersemiótica ocorre quando um enunciado realizado em um sistema semiótico é transformado e passa a ser expresso através de um novo sistema. Também chamado de transmutação, esse tipo de prática pode ser encontrado quando se faz a adaptação de uma obra literária para um produto audiovisual ou quando se realiza a audiodescrição de uma imagem.

²⁴ Reconhece-se, aqui, a proposta de Rónai (2012) para acrescentar a essa lista a “tradução sociolinguística”, que ocorre quando “[...] através das fórmulas usadas por nosso interlocutor em obediência a convenções sociais, tentamos descobrir o seu pensamento verdadeiro [...]”. Entretanto, por ser um conceito intimamente ligado a questões pragmáticas e por, nesta pesquisa, o termo “sociolinguística” ser utilizado para indicar uma outra área da linguística, não serão feitos maiores comentários sobre esse quarto tipo de tradução.

Pensando na tradução interlingual – foco deste capítulo daqui em diante –, pode-se dizer que, historicamente, essa atividade remonta quase que aos primórdios da humanidade. Do latim “*transducere*”, a palavra “traduzir” significa, em sua etimologia, “conduzir”, “levar para o outro lado”, definição diretamente ligada àquele que é o objetivo basilar da tradução: permitir que um povo tenha acesso à cultura e aos saberes de outro com o qual não divide a língua. Com efeito, havendo indivíduos de línguas diferentes que entram em contato, por qualquer motivo que seja, é natural que a tradução se torne necessária para o entendimento pleno dos falantes. Nesse sentido, da Antiguidade Clássica aos tempos contemporâneos, salvas as devidas exceções,

O fato é que não há povo tão isolado e tão auto-suficiente que possa dispensar o acervo de conhecimentos de experiências e conhecimentos de outros povos, e o intercâmbio de tais conhecimentos só é possível pela via da tradução: pois mesmo quando lê um texto estrangeiro, o leitor está, afinal, traduzindo, de certa maneira (Campos, 1988, p. 8).

Sob esse ponto de vista, a tradução ganha uma dimensão mais ampla, deixando de significar apenas um exercício linguístico para se transformar em “[...] um instrumento indispensável para a universalidade da cultura [...]” (Campos, 1988, p. 9). Em um mundo como o de hoje, em que o intercâmbio cultural ocorre por diversas vias – atingindo as áreas do comércio, da literatura, do audiovisual, entre outras –, torna-se, portanto, imprescindível sistematizar e formalizar os princípios da tradução, assim como suas aplicações, técnicas e tendências, para que se possa tratar dessa atividade com a objetividade e a criticidade que ela merece. Desse esforço teórico, resultam as pesquisas que hoje compõem o ainda recente campo das Teorias da Tradução, também conhecido como tradutologia, abordado nesta seção.

De maneira geral, há uma série de possibilidades que se apresentam ao tradutor interlingual ao começar seu trabalho. Inicialmente, ele pode optar por realizar uma tradução integral, passando o texto inteiramente da LP para a LC, ou parcial, deixando certas partes do texto no original. Pode, além disso, empreender uma tradução direta, se elaborada a partir do texto em sua língua original, ou indireta, se redigida a partir de outra tradução. Ademais, uma tradução pode ser considerada livre quando é feita de maneira razoavelmente independente do original, podendo se aproximar de uma adaptação, enquanto uma tradução literal é produzida visando uma fidelidade maior à forma e ao conteúdo (Campos, 1988).

A própria definição do que seria uma tradução literal, porém, varia à medida que se altera a noção de fidelidade ao original. Apesar de, ao longo do tempo, se ter observado um movimento pendular na valorização de textos traduzidos que se aproximavam mais ou menos

de seus originais, houve, por muito tempo, um entendimento básico de que a fidelidade – ou algum conceito a ela equivalente, como será discutido adiante –, é fulcral para se atingir uma boa tradução. Definir esse princípio, contudo, se caracteriza como uma tarefa complicada e trabalhosa, posto que ela “[...] varia segundo a época, o horizonte de expectativa – em suma, em função de critérios culturais complexos que é preciso serem levados em conta [...]” (Oustinoff, 2015, p. 93-94).

Para o senso comum, uma tradução fiel seria apenas aquela caracterizada como “palavra por palavra”, isto é, uma produção na qual cada palavra do texto na LP é substituída por uma igual em sentido na LC. Entretanto, basta considerar brevemente as inúmeras especificidades de cada língua para perceber que tal empreitada é virtualmente impossível. Na passagem do inglês para o português, por exemplo, inúmeras construções compostas – como as perguntas no presente simples – não poderiam ser traduzidas dessa maneira, ou seja, não seria possível atingir uma simetria perfeita. Assim, chega-se à ideia de que “[...] a noção de fidelidade implica talvez menos aderência às palavras da língua-fonte do que obediência aos usos e às estruturas da língua-alvo [...]” (Rónai, 2012, p. 22).

Ademais, há muito já se sabe que é apenas na união de palavras, proferidas ou escritas em contextos determinados, com objetivos comunicativos e público-alvo definidos, que um enunciado ganha sentido. Desse modo, a tarefa do tradutor não é traduzir a letra em si, mas os textos em sua totalidade, considerando todos os fatores linguísticos e extralinguísticos que se mostram relevantes para a sua composição. Para isso, é necessário que, antes de tudo, ele realize uma interpretação do texto a ser traduzido, de modo que a tradução possa, então, ser pensada como “[...] um ato hermenêutico por excelência, de penetração no significado e de explicitação dele em outro meio linguístico [...]” (Paes, 1990, p. 39).

Entretanto, como é natural ocorrer em diversas áreas do conhecimento, as tendências no estudo das teorias da tradução também foram se modificando com o passar dos anos, passando a abarcar novas discussões e desenvolvimentos. Nesse cenário, surgiram as teorias funcionalistas da tradução, que trouxeram consequências profundas para a maneira de enxergar essa atividade e a questão da fidelidade ao original. Em oposição às primeiras correntes estruturalistas da década de 60, que pensavam a língua como um código e tinham a equivalência formal como um conceito primordial da tradução, as teorias funcionalistas saem das discussões voltadas apenas à literatura e a textos técnicos para passarem a englobar todos os textos cotidianos passíveis de tradução. Nessa empreitada, defendem que o tradutor deve considerar tanto a função que a tradução vai exercer na comunidade de chegada como as intenções do texto original em seu contexto de criação.

A primeira teoria a fazer isso foi a Teoria do Escopo (Reiss; Vermeer, 2013[1984]), que define como seu princípio básico a ideia de que é o propósito do texto de chegada – o escopo – que irá determinar qual será o método e a estratégia empregados no processo tradutório. Para essa teoria, o propósito do texto é definido pelas instruções de tradução (“*translation brief*”) recebidas do cliente, e não mais apenas pelas características do texto de partida.

Com isso em mente, de acordo com Nord (2006), uma tradução se torna funcional, portanto, quando atinge o propósito para ela definido por suas instruções, ou seja, quando ela funciona para seus receptores da maneira desejada pelo emissor do texto original em uma situação comunicativa específica. Nesse sentido, “[...] o uso da língua em certo texto-fonte é investigado com o intuito de esclarecer em detalhes, primeiramente, que meios linguísticos são usados para realizar funções comunicativas e, secundariamente, como o texto é construído [...]” (Reiss, 1981, p. 127, tradução nossa)²⁵. A língua, aqui, está a serviço da comunicação, sendo essencial que o tradutor observe os objetivos de ambos os textos, original e tradução, a fim de que sua organização possa gerar uma equivalência funcional quando em contato com o receptor.

Diante disso, é necessário, portanto, que o profissional avalie as características de cada público, assim como sua disposição para cooperar na criação dos sentidos de um dado texto, a fim de antecipar os efeitos que possivelmente seriam suscitados nos leitores por formas de expressão diferentes. Além disso, na produção de qualquer texto, inclusive de uma tradução, são utilizados inúmeros marcadores funcionais que indicam os objetivos comunicativos de um texto. Tais marcadores podem incluir o uso de títulos explicativos, formatações específicas, formas verbais determinadas, registros particulares, formas de tratamento, entre outros, que apontam para a função que eles devem cumprir (Nord, 2006). Dessa maneira, “[...] se os receptores reconhecem os marcadores funcionais, eles podem considerar que o texto serve à função pretendida [...]” (Nord, 2006, p. 32, tradução nossa)²⁶, sendo necessário, então, que o tradutor esteja familiarizado com essas estratégias tanto na LP como na LC.

Nesse panorama, à medida que o uso da tradução na comunidade de chegada se mostra tão importante quanto o texto na LP, “[...] a ‘aura sacra’ do original e a força/status desse texto em relação à tradução, como aparecia em outras teorias, é diluída consideravelmente [...]” (Spolidorio, 2017, p. 20). Desse modo, a questão da fidelidade passa a ser tratada com maior flexibilidade nas teorias funcionalistas, sendo substituída pelo conceito de lealdade em

²⁵ “[...] the use of language in a given SL text is investigated in order to clarify in detail, firstly, what linguistic means are used to realize specific communicative functions, and, secondly, how the text is constructed [...]”.

²⁶ “[...] if the receivers recognise the function markers, they may accept the text as serving the intended function [...]”.

tradução. É esse o nome que Nord (2006) dá à limitação ética que restringe os tradutores diante das infinitas possibilidades de escopos que poderiam ser adotados para cada texto. A autora argumenta que

[...] tradutores, em seu papel como mediadores entre duas culturas, têm uma responsabilidade especial no que diz respeito aos seus parceiros, isto é, o autor do texto de partida, o cliente ou solicitante da tradução, e os receptores do texto de chegada, e para consigo mesmos, precisamente naqueles casos em que há diferentes visões do que uma ‘boa’ tradução é ou deveria ser [...] (Nord, 2006, p. 33, tradução nossa)²⁷.

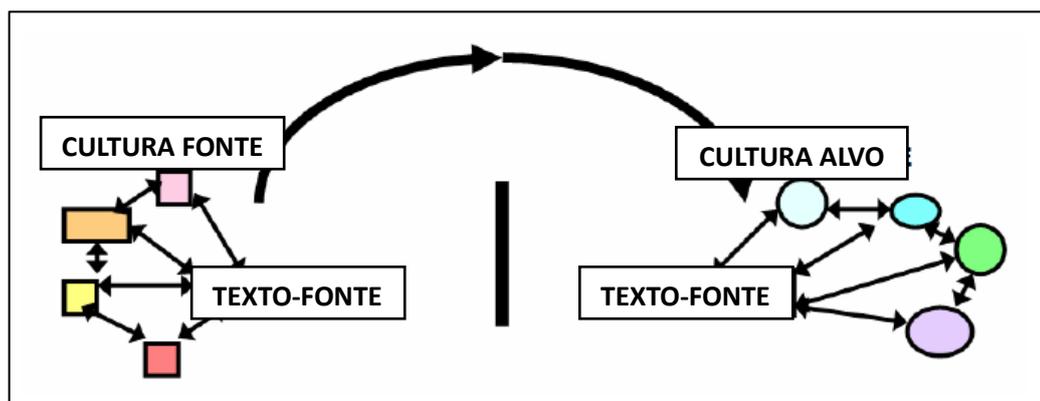
Trata-se, portanto, de uma relação social na qual o tradutor assume a tarefa de realizar a tradução de acordo com as necessidades do cliente e/ou da comunidade de chegada, ao passo que os receptores do texto esperam não ser enganados. Portanto, deixar a noção de fidelidade para trás em favor do princípio da lealdade significa, para além de uma alteração terminológica, uma mudança de postura. Deixa-se uma categoria tradicionalmente intertextual, que se refere às semelhanças linguísticas e/ou estilísticas entre os textos de chegada e de partida, em direção a uma categoria que é essencialmente interpessoal, baseada na confiança mútua e voltada à mediação entre duas culturas (Nord, 2006).

Para que essa mediação aconteça e o processo de tradução seja leal, outro ponto relevante é que o tradutor deve apresentar não só um conhecimento linguístico extensivo das duas línguas em jogo, como já apontado, mas, também, ter consciência de que, assim como o texto original, a tradução, qualquer que seja o seu propósito, “[...] está fadada a entrar no repertório da cultura-alvo e se tornar parte de um sistema intertextual culturalmente específico [...]” (Nord, 2006, p. 38, tradução nossa)²⁸. Isso quer dizer que, ao adentrar a comunidade de chegada, toda tradução estabelecerá relações com os outros textos que nela já circulam, havendo um horizonte de expectativas que deve ser cumprido, independentemente do propósito de cada texto. A Figura 5 abaixo representa graficamente essas conexões, explicitando a multiplicidade de fatores que devem ser levados em conta na elaboração de uma tradução.

²⁷ “[...] translators, in their role as mediators between two cultures, have a special responsibility with regard to their partners, i.e. the source-text author, the client or commissioner of the translation, and the target-text receivers, and towards themselves, precisely in those cases where there are differing views as to what a ‘good’ translation is or should be [...]”.

²⁸ “[...] is bound to enter the target-culture repertoire and become a part of the culture-specific intertextuality system [...]”.

Figura 5 - Representação gráfica da intertextualidade em processos de tradução



Fonte: Traduzido de Nord (2006)

De maneira geral, o que se pode observar, então, é que, nas teorias funcionalistas, a tradução é “[...] um processo interpretativo e comunicativo que consiste na reformulação de um texto pelos meios de outra língua que se desenvolve em um contexto social e com uma finalidade determinada [...]” (Hurtado Albir, 2001, p. 41, tradução nossa)²⁹. Os tradutores, como agentes responsáveis na mediação cultural, têm o papel fundamental de aproximar culturas, respeitando tanto o contexto de produção do original como as expectativas do público-alvo. Por isso, essa corrente torna-se especialmente adequada para observar o material explorado nesta pesquisa, uma vez que muitas das decisões tomadas durante a legendagem de uma produção audiovisual podem ser justificadas ao se observar a função da legenda, conforme será discutido nas subseções a seguir.

4.2 A tradução audiovisual

Como exposto anteriormente, o campo da tradução engloba práticas que envolvem textos de diferentes naturezas, de modo que pensar as especificidades de cada uma delas é imprescindível para um entendimento mais profundo dos aspectos que as afetam. É nesse contexto que se firma o campo da TAV que, acompanhando os desenvolvimentos tecnológicos que permitiram a criação de novas formas de interação, comunicação e arte, preocupa-se em dar acesso a produções audiovisuais em diversas línguas através da tradução.

Pensando de um ponto de vista histórico, pode-se dizer que a TAV se faz presente na sociedade pelo menos desde a criação do cinema mudo em decorrência da necessidade de se traduzir os intertítulos – as sequências de texto filmadas que comumente traziam o diálogo entre

²⁹ “[...] un proceso interpretativo y comunicativo consistente en la reformulación de un texto con los medios de otra lengua que se desarrolla en un contexto social y con una finalidad determinada [...]”.

as personagens (Romero-Fresco, 2013, p. 205). Entretanto, foi, naturalmente, com o advento dos filmes com falas sonorizadas, a partir dos anos vinte do século passado, que esse tipo de tradução começou a se popularizar, tornando-se indispensável em um mundo no qual as telas – e, conseqüentemente, as produções audiovisuais –, são cada vez mais presentes nas vidas dos indivíduos.

Qualquer que seja a maneira como ela toma forma, o fato é que a TAV lida com textos nos quais há várias dimensões em jogo, sendo que “[...] a concorrência de camadas semióticas diferentes, através dos canais visual (imagens, texto escrito, gestos) e auditivo (música, barulho, diálogo), torna a tarefa do tradutor particularmente desafiadora nessa área [...]” (Díaz-Cintas, 2010, p. 344, tradução nossa)³⁰. Além desses dois canais, o tradutor deve levar em conta o fato de que todo texto audiovisual é elaborado, também, através de dois tipos de sinais: os verbais e os não verbais, de modo que “[...] os diferentes sistemas de signos interagem e juntos constituem o texto audiovisual, uma estrutura que é mais complexa do que a soma de suas partes [...]” (Remael, 2010, p. 13, tradução nossa)³¹.

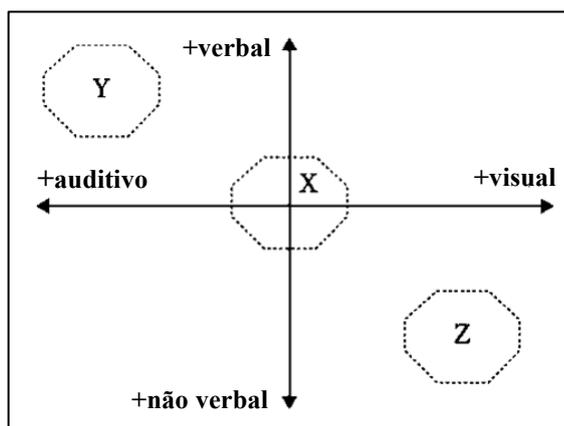
Com efeito, o nível de importância relativa atribuída a cada sistema pode variar a depender do foco do texto, com certas produções dando mais importância à fotografia do que ao diálogo, ou apresentando o roteiro como seu ponto mais forte, por exemplo. Ademais, a integração do componente verbal a um sistema intrincado de signos que devem ser ouvidos, vistos e, por vezes, lidos, comumente leva a uma hibridização desse aspecto, resultando em textos cuja linguagem não é puramente escrita ou falada (Remael, 2010).

Dessa forma, a natureza dos textos audiovisuais não é estanque, com eles podendo se distribuir de maneiras distintas com relação aos aspectos levantados. É a partir dessa constatação que Zabalbeascoa (2008) propõe um diagrama para mapear os objetos de estudo da TAV, composto por dois eixos: um que indica a presença e a importância da comunicação verbal em comparação com outras formas de expressão, e outro que aponta a relevância do som no canal auditivo *versus* a expressividade de signos visuais (Figura 6).

³⁰ “[...] the concurrence of different semiotic layers through the visual (images, written text, gestures) and audio (music, noise, dialogue) channels makes the translator’s task particularly challenging in this field [...]”.

³¹ “[...] the different sign systems interact and together constitute the audiovisual text, a structure that is more complex than the simple summation of its parts [...]”.

Figura 6 - Diagrama das possibilidades de distribuição de textos audiovisuais



Fonte: Traduzido de Zabalbeascoa (2008, p. 29)

Na representação acima, textos audiovisuais localizados mais ao centro, como X, seriam correspondentes àqueles mais prototípicos, em que as intenções e os significados são construídos através dos quatro elementos de maneira relativamente equilibrada. Por outro lado, produções mais periféricas indicariam que mais importância está sendo dada a um aspecto do que a outro. Em Y, poderiam se encontrar, a título de exemplo, filmes focados no roteiro, que tendem a favorecer as palavras ouvidas durante o diálogo; já em Z, estariam textos cujo aspecto mais expressivo são as imagens, em detrimento da comunicação verbal.

Assim, se são inúmeras as possíveis configurações que um texto audiovisual pode assumir, é apenas natural que existam, também, várias maneiras de lidar com essa diversidade. Com efeito, ao longo dos anos, a TAV desenvolveu diferentes modalidades de tradução que, por serem veiculadas em contextos distintos e divergirem quanto à finalidade, apresentam cada qual o seu conjunto de técnicas.

Nesse cenário, Franco e Araújo (2011) apresentam sete modalidades de TAV, que se dividem entre orais e escritas. Às orais, correspondem a dublagem, o *voice-over*, a narração e a audiodescrição. Já as escritas incluem a legendagem para ouvintes, a legendagem para surdos e ensurdecidos e a legendagem eletrônica.³² Apesar de apenas uma dessas formas ser o foco desta pesquisa, todas elas serão apresentadas a seguir, para que se tenha uma melhor compreensão das características básicas de cada uma.

A dublagem consiste na substituição completa do áudio original de uma língua para o de outra, traduzindo-se todos os diálogos e demais informações que possam aparecer no vídeo.

³² Alguns autores, como Gambier (2001), adicionam, ainda, outros itens a essa lista, como a interpretação consecutiva, a interpretação simultânea, a interpretação de sinais e a tradução simultânea. Para Franco e Araújo (2011), porém, essas modalidades estariam mais próximas à área dos Estudos da Interpretação, fora, portanto, do escopo da TAV.

Nela, é necessário que sejam observados os tempos de entrada e saída de fala, assim como a sincronia labial, para que se possa gerar no espectador uma impressão de naturalidade. Trata-se de uma modalidade amplamente conhecida no Brasil, em especial devido à Lei 4.117/62, que torna obrigatório que os produtos audiovisuais transmitidos na televisão aberta sejam dublados.

Já no *voice-over*, ao contrário da dublagem, a substituição completa do áudio não ocorre, havendo apenas a adição do áudio traduzido sobre o original, que pode ser ouvido em volume mais baixo, ao fundo. Não havendo a intenção de manter a ilusão de naturalidade, uma vez que ambas as línguas são veiculadas ao mesmo tempo, a preocupação com a sincronia labial praticamente desaparece, sendo, inclusive, comum que a fala na faixa traduzida termine antes da que ocorre no áudio no idioma original. De maneira geral, o *voice-over* tem um alcance mais restrito, sendo mais utilizado para produtos audiovisuais de não ficção da televisão por assinatura, além de estar presente em entrevistas exibidas em programas da televisão aberta.

Por outro lado, na narração, não se vê o dono da voz em momento algum, de modo que ela é feita por um ‘falante invisível’. Assim sendo, esse tipo de tradução não pressupõe sincronia labial, o que a torna relativamente livre em comparação com as outras. Essa modalidade pode ainda ser utilizada com outras formas de tradução, como a dublagem ou o *voice-over*, em qualquer tipo de produção audiovisual, comumente se fazendo presente quando, no original, também há narração.

Completando o grupo das modalidades orais da TAV, a audiodescrição consiste em uma descrição oral de estímulos visuais. Podendo ser inter ou intralingual, ela acompanha textos audiovisuais veiculados em telas, como programas de televisão, filmes e vídeos de redes sociais, além de constar em espetáculos teatrais, exposições artísticas, entre outros produtos. Diferentemente das outras, trata-se de uma modalidade especificamente voltada à inclusão, buscando dar mais acessibilidade a pessoas com algum tipo de deficiência visual.

Quanto às modalidades escritas, a mais conhecida é a legendagem para ouvintes, corriqueiramente chamada apenas de legendagem, que será apresentada mais a fundo na próxima subseção. Por enquanto, vale dizer que a legendagem diz respeito à tradução na qual os diálogos e as informações em tela de um conteúdo audiovisual são transpostos por escrito, sendo necessário observar uma série de parâmetros técnicos que, de certa forma, restringem o tradutor. É esse o tipo de legendagem mais presente no mercado brasileiro, servindo como base para a definição dos parâmetros básicos dos outros tipos existentes.

Com outros objetivos, a legendagem para surdos e ensurdecidos realiza não só a tradução por escrito dos diálogos do original, obedecendo aos mesmos parâmetros que regem a legendagem para ouvintes, mas também registra em seu texto barulhos, ruídos, músicas,

nomes de personagens ou quaisquer outras informações de natureza sonora que são relevantes para o entendimento da cena.

Por fim, a legendagem eletrônica pode ser utilizada em teatros durante a exibição de peças e óperas em línguas estrangeiras, sendo veiculada em um painel eletrônico que projeta, em sequência, a tradução de todas as falas e músicas. Devido à posição em que aparece, ou seja, acima do palco, é conhecida também com supra-legendagem, apesar de não ser esse o termo mais empregado nos meios que utilizam essa modalidade de TAV.

Pode-se adicionar a esses tipos de tradução, ainda, a legendagem bilíngue que, apesar de não mencionada por Franco e Araújo (2011), é bastante presente em contextos específicos. Em países nos quais convivem diversas línguas, é possível que se apresentem legendas para mais de um idioma ao mesmo tempo, cada uma ocupando uma linha. É o caso da Bélgica, por exemplo, onde o francês e o neerlandês são utilizados, ou da Jordânia, onde são projetadas legendas em hebreu e árabe. Além disso, não é incomum encontrar legendas bilíngues em festivais de cinema, nos quais legendam-se os filmes comumente em inglês e na língua local de onde acontece o evento.

É possível notar, portanto, que uma das maiores preocupações da TAV é tornar um conteúdo audiovisual acessível ao público-alvo que, em geral, não domina a língua original. De fato,

[...] seja o obstáculo uma barreira linguística ou sensorial, o objetivo do processo de tradução é exatamente o mesmo: facilitar o acesso a uma fonte de informação e entretenimento que, do contrário, seria hermética. Desse modo, a acessibilidade se torna o denominador comum que sustenta essas práticas [...] (Díaz-Cintas; Remael, 2014[2007], p. 13, tradução nossa).³³

Diante do exposto, percebe-se que a TAV se configura como uma área particularmente múltipla, que lida com questões complexas que vão desde a própria natureza do material audiovisual com que trabalha até a definição de parâmetros técnicos que devem ser completamente dominados pelo tradutor. Dessa maneira, considerando os objetivos desta pesquisa e pensando na importância desses aspectos, a próxima subseção será dedicada a uma apresentação mais detalhada da legenda, cobrindo tanto suas restrições técnicas como as estratégias empregadas pelos tradutores para lidar com os limites a eles impostos.

³³ “[...] whether the hurdle is a language or a sensorial barrier, the aim of the translation process is exactly the same: to facilitate access to an otherwise hermetic source of information and entertainment. In this way, accessibility becomes a common denominator that underpins these practices”.

4.2.1 A legenda audiovisual

Conforme já mencionado, no campo da TAV, uma das modalidades mais amplamente utilizadas e conhecidas é a legendagem. Expandindo a definição brevemente traçada anteriormente, a

[...] legendagem pode ser definida como uma prática de tradução que consiste em apresentar um texto escrito, geralmente na parte inferior da tela, que busca recontar o diálogo original dos falantes, assim como os elementos discursivos que aparecem na imagem (letras, inserções, grafite, inscrições, placas e afins), e as informações contidas na trilha sonora (músicas, narração) (Díaz-Cintas; Remael, 2014[2007], p. 8, tradução nossa)³⁴.

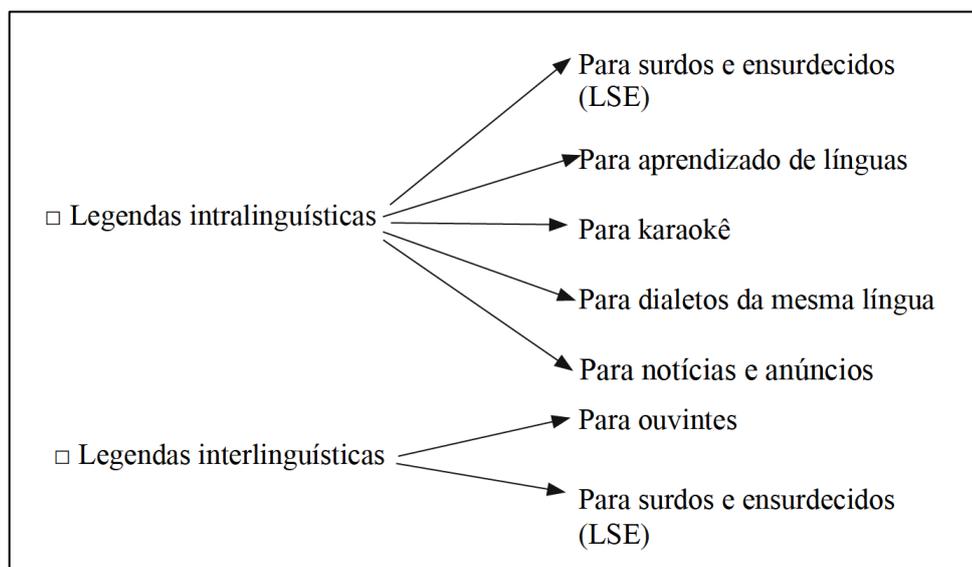
Dessa forma, três são os seus componentes essenciais, de acordo com Díaz-Cintas e Remael (2014[2007]): a palavra falada, a imagem, e a palavra escrita da legenda. Nesse contexto, esses elementos interagem e, junto a outros fatores como a habilidade de leitura do espectador e o tamanho de tela, determinam as características básicas desse meio audiovisual. Por isso, as legendas devem aparecer sempre de maneira sincronizada com a imagem e o diálogo, contendo uma representação da LP que seja semanticamente adequada, isto é, que não desconsidere informações semânticas importantes para o diálogo, e permanecendo em tela por tempo suficiente para que possam ser lidas. Além disso, as legendas não devem chamar a atenção para si mesmas, permanecendo, de certa forma, neutras do ponto de vista técnico e linguístico, a fim de não atrapalhar a experiência do espectador que assiste à produção audiovisual.

Uma vez que diversos fatores fazem parte de sua composição, os mesmos autores propõem cinco maneiras de se classificar as legendas, a partir de dimensões diferentes – a linguística, a do tempo disponível para sua preparação, a técnica, a dos métodos de projeção das legendas, e a do formato de distribuição. Nos próximos parágrafos, serão explorados cada um desses parâmetros, buscando compreender como eles influenciam a elaboração das legendas.

A primeira classificação tradicionalmente feita com relação às legendas é aquela que leva em conta o parâmetro linguístico. Desse ponto de vista, encontram-se dois tipos gerais de legendas, que se subdividem de acordo com o esquema apresentado na Figura 7 a seguir.

³⁴ “[...] subtitling may be defined as a translation practice that consists of presenting a written text, generally on the lower part of the screen, that endeavours to recount the original dialogue of the speakers, as well as the discursive elements that appear in the image (letters, inserts, graffiti, inscriptions, placards and the like), and the information that is contained on the soundtrack (songs, voices off). [...]”.

Figura 7 - Divisão entre os tipos de legenda a partir do parâmetro linguístico



Fonte: Traduzido de Díaz-Cintas e Remael (2014[2007])

Como foi explorado brevemente nas subseções anteriores, uma tradução intralinguística pressupõe uma mudança do meio oral para o escrito, permanecendo, porém, na mesma língua. O primeiro tipo elencado pelos autores, isto é, a LSE, é largamente focada em tornar conteúdos audiovisuais acessíveis para pessoas surdas ou ensurdecidas. Conhecidas também como *closed captions*, essas legendas podem ser ativadas e desativadas pelo próprio usuário e convertem o diálogo proferido pelos atores em texto escrito, que pode ser apresentado em até quatro linhas. Em alguns países, são utilizadas cores diferentes para cada personagem nas legendas, além de serem incorporadas, também, todas as informações paralinguísticas que contribuem para o enredo ou para a atmosfera da produção, tais como a ironia de uma frase, um telefone tocando, risadas, aplausos, entre outros. Outra característica específica da LSE é que esse tipo de legenda, apesar de normalmente figurar na parte inferior da tela, pode ser deslocada para os lados a fim de identificar o falante.

Outro tipo de legenda intralinguística é aquela voltada ao ensino e à aprendizagem de línguas, que já se encontra bem consolidada devido ao seu comprovado valor didático. Nesse caso, as convenções aplicadas às legendas são consideravelmente diferentes das interlinguais, não sendo incomum encontrar legendas de três ou mais linhas, que trazem repetições de elementos lexicais e frases incompletas, e que fazem a representação das falas exatamente como elas foram verbalizadas, forçando o público a demonstrar uma velocidade de leitura mais rápida. Com elas, porém, os espectadores conseguem “[...] ler na tela o diálogo escrito dos

atores e reconhecer ou confirmar o que eles não entenderam ao ouvir [...]” (Díaz-Cintas; Remael, 2014[2007], p. 16, tradução nossa)³⁵.

O terceiro exemplo de legendagem intralingual mencionado pelos autores é o correspondente às legendas para karaokê, que são geralmente utilizadas com músicas ou filmes musicais, para que o público possa cantar ao mesmo tempo que os cantores e/ou as pessoas retratadas nas telas. Atualmente, essas legendas podem ser amplamente encontradas em vídeos de karaokê no *YouTube*, que trazem músicas em uma infinidade de línguas diferentes.

Ainda no campo da legendagem intralingual, encontram-se as legendas que buscam auxiliar na compreensão de pessoas cujos sotaques são difíceis de entender, apesar de, em teoria, falarem a mesma língua. Essas legendas podem ser utilizadas na transmissão de um programa por completo, mas é mais comum que apareçam em momentos específicos em que se fala a língua com um sotaque forte e variações lexicais constantes. No Brasil, esse tipo de legenda costuma acompanhar falas de portugueses em entrevistas veiculadas na televisão aberta, por exemplo, devido às muitas diferenças encontradas entre o PB e o PE.

Finalmente, o último tipo de legenda intralingual destacado por Díaz-Cintas e Remael (2014[2007]) consiste na legendagem de notícias e anúncios. Esse tipo de legenda está presente em estações de metrô e outros espaços públicos nos quais é possível encontrar propagandas, além de poder ser empregada, também, em telejornais, para veicular informações extras acerca do acontecimento reportado pelos âncoras ou apresentar outras notícias na parte inferior da tela. O uso das da forma escrita permite, nesse caso, que a informação seja transmitida sem som, de modo que não incomode o público.

O segundo grande grupo de legendas – e, sem dúvidas, o mais conhecido – é o da legendagem interlingual, voltada à tradução do conteúdo audiovisual de uma LP para uma LC. Chamada por Gottlieb (1994) de legendagem diagonal, uma vez que a mudança ocorre não só entre as línguas, mas, também, entre os modos oral e escrito, esse tipo de tradução é dividido, no esquema de Díaz-Cintas e Remael (2014[2007]), entre dois grupos. O primeiro deles consiste na legendagem para ouvintes, que verte, de acordo com as restrições impostas pelos parâmetros técnicos, todas as falas oralizadas no original. Como anunciado anteriormente, esse tipo de legendagem será explorado em maiores detalhes mais a frente, uma vez que, nesta pesquisa, é precisamente esse o material utilizado. O segundo segue as mesmas convenções da LSE intralingual, buscando apresentar tanto o diálogo traduzido como as informações sonoras necessárias para o entendimento da obra.

³⁵ “[...] read on the screen the written dialogue of the actors and recognize or confirm what they had not understood aurally [...]”.

Quanto à classificação da legendagem a partir do parâmetro do tempo disponível para preparação, também são duas as possibilidades apontadas por Díaz-Cintas e Remael (2014[2007]): as legendas pré-preparadas (legendagem *offline*), elaboradas após a gravação do programa, e as que são feitas em tempo real (legendagem *online*), que são criadas ao mesmo tempo em que o programa é transmitido. Nesse contexto, as legendas pré-preparadas podem ser realizadas de maneira completa, traduzindo-se as frases por inteiro, ainda que se respeitando as restrições técnicas impostas pela modalidade, ou de modo reduzido, quando apenas as informações mais relevantes são efetivamente traduzidas. Já as legendas em tempo real, mais comuns em casos de legendagem intralingual, são empregadas apenas quando não há tempo útil para elaborá-las, normalmente no caso de entrevistas ao vivo, discursos políticos, programas de esporte, entre outros. Atualmente, esse tipo de legenda pode ser criado tanto manualmente, por um tradutor especializado, como por tradutores automáticos conectados a *softwares* de reconhecimento de voz.

Passando a um ponto de vista técnico, pode-se falar, ainda, em legendas abertas e fechadas. As primeiras consistem naquelas que já são embutidas no vídeo e não podem ser removidas ou desligadas, de maneira que o espectador não tem poder de escolha sobre elas. É o caso, por exemplo, das legendas de filmes exibidos nos cinemas, alguns canais de televisão por assinatura e certos vídeos na internet. Já as legendas fechadas podem ser ativadas conforme a vontade de quem assiste ao conteúdo, como sempre foi observado no caso da LSE. Sobre essa diferença, Díaz-Cintas e Remael (2014[2007], p. 22) afirmam que, durante muito tempo, as legendas abertas eram as mais populares em todos os contextos, pelo menos até o advento do DVD. Com o surgimento dessa tecnologia, passou-se a encontrar legendas fechadas cada vez mais, tanto em canais televisivos, mas, sobretudo, em vídeos na internet e, mais recentemente, plataformas de *streaming*.

No que diz respeito aos métodos de projeção das legendas, das cinco possibilidades apresentadas por Díaz-Cintas e Remael (2014[2007]), apenas duas – a legendagem a laser e a eletrônica – são de fato exploradas pelos autores, uma vez que são as únicas ainda utilizadas no mercado³⁶. A legendagem a laser, aplicada desde a década de 80, consiste na utilização de um raio laser para queimar a legenda na cópia positiva do filme, obedecendo às precisas marcações temporais. O resultado são legendas precisas, com boa definição dos caracteres e contornos que

³⁶ A título de curiosidade, os outros três métodos previamente utilizados para a adição de legendas a produtos audiovisuais eram a legendagem mecânica e térmica, a legendagem fotoquímica e a legendagem óptica. Para maiores informações sobre esses processos, pode-se conferir os trabalhos de Ivarsson e Caroll (1998) ou Díaz-Cintas (2001).

facilitam a leitura do texto. A outra alternativa, a legendagem eletrônica³⁷, corresponde ao método que apenas projeta, sobre o filme, a faixa de legendas, ao invés de alterar o original. Como, nesse caso, as legendas não são queimadas na cópia, esse método permite a diminuição dos custos de produção e distribuição de produtos audiovisuais, sendo, hoje, a estratégia preferida das grandes produtoras.

Finalmente, a última categoria apontada por Díaz-Cintas e Remael (2014[2007]) é aquela voltada ao formato de distribuição das legendas, sendo ela uma das que mais influencia a maneira como as legendas são produzidas. De acordo com os autores, encontram-se, então, cinco formatos de distribuição: o cinema, a televisão, o vídeo (VHS), o DVD e a internet. Obviamente, trata-se de uma lista já ultrapassada, dado que, há muito tempo, as fitas VHS e os DVDs perderam espaço no mercado, sendo substituídos, primeiramente, pelo *BluRay* e, em anos mais recentes, pelas plataformas de *streaming* e vídeos distribuídos via internet. Apesar disso, as considerações traçadas pelos estudiosos são relevantes na medida em que deixam clara a relação intrínseca entre o meio de exibição do produto audiovisual e a produção das legendas. Em legendas para o cinema, por exemplo, um número maior de caracteres é permitido do que para produções pensadas para a televisão, devido ao tamanho da tela e das condições de exibição – no cinema, os espectadores apresentam índices mais altos de concentração, posto que não há outros estímulos em jogo além do próprio filme.

Mesmo com todas essas divisões, um ponto que nunca se altera quando se discute legendagem é a relação direta que existe entre essa atividade e as outras semioses mobilizadas pelo produto audiovisual. Como levantado na subseção 4.2, toda TAV lida com produções de alta complexidade semiótica, nas quais diferentes sistemas de signos operam de modo integrado para criar um texto coerente, não sendo diferente com as legendas. Com efeito, no caso das legendas,

[...] há interação dialógica entre os personagens (ficcionais ou não) na tela, mas suas palavras também interagem com os signos visuais do filme, e essa interação determina o que o filme comunica ao espectador [...] (Díaz-Cintas; Remael, 2014[2007], p. 48, tradução nossa)³⁸.

³⁷ Aqui, o termo “legendagem eletrônica” não faz referência à modalidade apontada por Franco e Araújo (2011), que corresponde às supra-legendas, mas especificamente ao processo de projeção das legendas.

³⁸ “[...] there is dialogic interaction between the (fictional or non-fictional) characters on screen, but their words also interact with the film’s visual signs, and this interaction determines what the film communicates to the spectator [...]”.

Assim, em uma legenda audiovisual, não é suficiente que exista apenas uma coesão textual entre os elementos traduzidos, sendo impreterível alcançar uma coesão semiótica. Isso significa que deve haver uma integração entre a linguagem, o áudio e as imagens na tela, lançando mão dessas relações para criar um todo linguístico-visual coerente (Díaz-Cintas; Remael, 2014[2007]). Nesse cenário, lacunas no diálogo legendado podem e devem ser preenchidos pelos espectadores com informações retiradas das imagens na tela, uma vez que, nas palavras de Díaz-Cintas e Remael (2014[2007], p. 51, tradução nossa), “[...] na escrita, um substantivo pode ser substituído por um pronome, enquanto em textos audiovisuais, pronomes no diálogo podem se referir a pessoas ou objetos na tela [...]”³⁹.

Para que isso ocorra, porém, é naturalmente necessário que exista sincronia entre as imagens e falas do texto (audiovisual) original e o texto traduzido na LC, o que não diz respeito apenas aos tempos de entrada e saída de cada legenda, apesar de esses dois aspectos serem, sim, relevantes. A fim de evitar estranhamentos e críticas, é igualmente importante que as legendas demonstrem, na medida do possível e respeitando as restrições do meio, algum nível de aproximação quanto à linguagem empregada no texto de partida. Comumente, o que os tradutores tendem a fazer é, então, realizar a tradução de itens que possuem semelhanças fonéticas ou morfológicas com a LC e respeitar ao máximo as estruturas sintáticas e referências semânticas dos diálogos no conteúdo audiovisual.

Nessa passagem do oral para o escrito, ainda outro aspecto que deve ser observado com atenção é o estilo – aqui colocado de maneira geral, para indicar o conjunto das escolhas linguísticas em um produto audiovisual. De acordo com Díaz-Cintas e Remael (2014[2007], p. 61), existem dois tipos de fala no audiovisual: a espontânea e a roteirizada. Como o próprio nome já adianta, falas espontâneas acontecem sem planejamento prévio, sendo mais comuns em entrevistas e em conteúdos voltados ao improviso. Por outro lado, as falas roteirizadas são planejadas, e podem ser elaboradas também de três maneiras, a saber: (i) para imitar uma conversa natural, sendo chamada de diálogo mimetizado; (ii) para reproduzir a maneira como a língua era falada em outro período, caso em que recebe o nome de diálogo estilizado; e (iii) para ler falas a partir de textos manuscritos ou de teleprompters, como discursos políticos e notícias televisivas.

Dentre todos, o mais comum no meio audiovisual é o diálogo mimético, de modo que deve existir uma preocupação com a naturalidade da fala representada pela legenda, assim como houve no momento da concepção da produção audiovisual. Isso quer dizer que, ao traduzir, o

³⁹ “[...] in writing, a noun may be replaced by a pronoun, whereas in audiovisual texts pronouns in the dialogue can refer to people or objects on the screen [...]”.

legendista deve fazer com que, em algum nível, o estilo de fala apresentado nas legendas seja condizente com aquele utilizado no original. Dessa maneira, reproduzido uma citação já utilizada neste trabalho, “[...] é essencial que a diversidade linguística seja respeitada pelos profissionais ao longo do processo de tradução, utilizando variantes mais ou menos formais ou prestigiadas quando essas se fazem necessárias [...]” (Lopes; Afonso, 2021, p. 146), isto é, de acordo com a natureza dos conteúdos.

O tom da produção também irá definir, de certo modo, o quão flexíveis as legendas poderão ser em comparação com o original. Por exemplo, ao legendar um produto audiovisual de teor mais factual, como documentários ou notícias, é necessário que o tradutor seja bastante criterioso na escolha dos termos empregados, enquanto produções ficcionais permitem que se legende com maior liberdade. Nesse contexto, pode-se afirmar, portanto, que

[...] legendar fatos é algo como adicionar legendas a uma imagem, como em um jornal, enquanto legendar ficção é recriar falas, como no teatro. Então, ao lidar com programas factuais, o teste decisivo seria “É assim que nós normalmente chamamos os objetos ou eventos que vemos aqui?”. Na ficção, interpretando *pessoas*, nós teríamos que dizer “É isso o que esse tipo de pessoa diria nessa situação?” [...] (Gottlieb, 1994, p. 111, tradução nossa)⁴⁰.

De todo modo, mesmo que o tradutor esteja ciente dos pontos acima levantados, o fato é que a transição entre o meio oral e o escrito implica, assim como em toda tradução, algumas perdas que devem ser compensadas. Nas legendas, independentemente do tipo de produto audiovisual com que se trabalha, certos traços da língua falada, tais como repetições, pausas, hesitações, exclamações, entre outros, são inevitavelmente alterados, em especial por causa das restrições inerentes a essa modalidade de tradução. Surge, então, um estilo característico das legendas, que “[...] tem um impacto na gramática e no registro, assim como no âmbito interacional e em outros traços orais do diálogo [...]” (Díaz-Cintas; Remael, 2014[2007], p. 63, tradução nossa)⁴¹.

Com vistas a entender esse estilo, é necessário, então, conhecer cada um dos parâmetros técnicos que limitam as legendas, buscando compreender como, exatamente, eles têm o poder de influenciar as escolhas linguísticas dos tradutores. Por isso, cada um deles será apresentado nos parágrafos que se seguem, assim como algumas orientações gerais que derivam da

⁴⁰ “[...] subtitling facts is somewhat like adding captions to a picture, as in a newspaper, whereas subtitling fiction is recreating lines, as in drama. So, when dealing with factual programs, the litmus test would be “Is this what we normally call the objects or events we see here?”. In fiction, interpreting *people*, we would have to say “Is this what his kind of person would say in this situation?” [...]”.

⁴¹ “[...] an impact on grammar and register, as well as on the interacional and other oral features of dialogue [...]”.

necessidade de obedecer a esses limites. Antes, porém, é interessante tecer um breve comentário acerca do processo de padronização das legendas a nível internacional.

Como já foi possível perceber, o trabalho com a legendagem é essencialmente multifacetado, de modo que, ao longo dos anos, inúmeras foram as convenções criadas para auxiliar os legendistas em sua função. Contudo, essas convenções foram, por muito tempo, específicas de cada produtora ou país, não havendo um padrão que pudesse ser aplicado em larga escala. Diante dessa situação, um dos primeiros esforços para criar uma padronização quanto à legendagem foi a sistematização elaborada por Carroll e Ivarsson (1998), que propuseram seu *Código de Boas Práticas em Legendagem (Code of Good Subtitling Practice*, no original).

Apesar de antigas, várias das recomendações feitas pelos estudiosos acabaram se cristalizando no mercado das legendas, sendo retomadas por diversos autores (cf. Karamitroglou, 1998; Díaz-Cintas; Remael, 2014[2007]; Nunes, 2012; Naves *et al.*, 2016), e consistentemente empregadas até os dias atuais com poucas ou nenhuma alteração. No presente texto, os principais pontos indicados nesses trabalhos serão apresentados em dois blocos, seguindo a divisão proposta por Díaz-Cintas e Remael (2014[2007]): um para os parâmetros espaciais e outro para os parâmetros temporais. Aqui, não serão abordadas, porém, as convenções relacionadas à pontuação e/ou ao uso de outros caracteres e estratégias – tais como o uso de itálico ou de cores –, uma vez que elas não são relevantes para as escolhas linguísticas dos tradutores⁴².

Dando início à discussão dos parâmetros espaciais, observam-se, como um dos mais importantes, o número e a posição das linhas que podem ser usados em cada legenda. No geral, devido à necessidade de a legenda não chamar a atenção para si mesma, o ideal é que ela ocupe o mínimo de espaço na tela, motivo pelo qual se convencionou a utilização de, no máximo, duas linhas. Ademais, por padrão, as legendas são posicionadas horizontalmente na parte inferior da tela e alinhadas ao centro, dado que a ação em produtos audiovisuais não costuma ocorrer nessa área. Quando se tem uma legenda com apenas uma linha, cabe à produtora definir qual será a melhor posição para ela: se na linha superior ou na inferior.

Quanto à escolha por uma ou duas linhas, os manuais comumente afirmam que se deve dar preferência a legendas de uma linha sempre que possível, já que a inserção de quebras desnecessárias pode forçar o espectador a cobrir uma área maior da tela em menos tempo, prejudicando a velocidade de leitura. Assim, quebras de linha devem ser empregadas somente

⁴² Para maiores informações sobre essas convenções, conferir Díaz-Cintas e Remael (2014[2007]).

quando uma legenda for demasiadamente longa, sendo que, havendo duas linhas, recomenda-se que a de cima seja preferencialmente menor do que a de baixo, a fim de não poluir a imagem. De qualquer forma, “[...] o fator primordial para a apresentação de legendas em duas linhas na tela é garantir quebras de linha que sejam sintaticamente apropriadas, respeitando e refletindo a lógica da sentença [...]” (Díaz-Cintas; Remael, 2014[2007], p. 87, tradução nossa)⁴³.

Outro aspecto espacial especialmente relevante diz respeito à configuração dos caracteres nas legendas. Costumeiramente, são utilizadas fontes sem serifa, nas cores branca ou amarela, com sombreamento e contorno preto para aumentar a legibilidade. Além disso, existe um limite de CPL permitido, que varia de acordo com o formato de distribuição e tem o objetivo de preservar o conforto durante a leitura das legendas. Para a televisão, por exemplo, tal limite gira em torno de 37 CPL, incluindo espaços e sinais tipográficos, enquanto, para filmes, o limite costuma aumentar para 43. Não há, todavia, um número mínimo de CPL convencionalmente aceito, apesar de, no geral, legendas com menos de 4 ou 5 caracteres serem raros. Vale mencionar, ainda, que, apesar de existirem orientações quanto a esse parâmetro técnico, a palavra final quanto ao número de CPL permitidos em uma legenda será ditada pelo guia de legendagem adotado ou pelo cliente.

Passando agora aos parâmetros relacionados à dimensão temporal que envolve a legendagem, deve-se pensar, primeiramente, na questão da marcação (*spotting*), ou seja, da determinação do tempo de entrada e de saída de cada legenda. Naturalmente, a marcação de um diálogo deve refletir o ritmo da produção audiovisual, levando em consideração as pausas, as entonações e outros traços da fala original. Nesse sentido, o ideal é que uma legenda comece no momento preciso em que alguém inicia seu turno de fala e saia de tela quando se para de falar, respeitando, também, movimentos de câmera, mudanças nas tomadas, entre outros.

Portanto, um dos aspectos mais importantes para a legendagem é, sem dúvida, a sincronização. De fato, legendas que aparecem em tela muito antes ou muito depois do início das falas, ou que saem de cena sem seguir o original, podem deixar os espectadores confusos e impedir que o produto audiovisual seja apreciado, posto que isso gera um descompasso entre a palavra escrita e os outros elementos veiculados. Naturalmente, nem sempre é possível atingir uma sincronização perfeita, principalmente ao trabalhar com diálogos semanticamente densos e com muita informação, sendo certo grau de flexibilidade permitido nesses casos. De todo modo, é importante lembrar que,

⁴³ “[...] the overriding factor when presenting two-liners on screen is to secure line-breaks which are appropriate syntactically, respecting and reflecting the logic of the sentences [...]”.

[...] uma marcação precisa é crucial para uma legendagem ideal, já que ela reforça a coesão interna do programa traduzido e tem o papel essencial de ajudar o espectador a identificar quem diz o que no programa [...] (Díaz-Cintas; Remael, 2014[2007], p 90, tradução nossa)⁴⁴.

Ao lidar com a marcação e a sincronização, o legendista não pode se esquecer, entretanto, de voltar um olhar cuidadoso também para a duração da legenda, que está intimamente ligada a questões como a velocidade de leitura e o número de CPS que pode ser utilizado. Como padrão, tem-se que cada legenda deve permanecer em tela por, no mínimo, um e, no máximo, seis segundos, tanto para garantir que o espectador tenha tempo suficiente para ler a legenda como para se assegurar de que ele não comece a reler seu conteúdo, conseqüentemente perdendo o foco nos outros canais do texto.

A partir desses valores, pode-se, então, calcular qual seria o limite de CPS que seria permitido para cada programa, podendo haver uma variação no que diz respeito à velocidade de leitura do público-alvo. Ao legendar um programa ou filme infantil, cujos espectadores ainda não desenvolveram plenamente sua capacidade leitora, deve-se adotar uma velocidade de leitura e um número de CPS mais baixos, ao passo que em programas voltados para adultos, por exemplo, espera-se uma maior habilidade de leitura, o que permitiria limites de CPS mais altos.

Todas essas convenções trazem, como era de se esperar, conseqüências diretas para o plano linguístico das legendas. Em face das restrições espaciais e temporais impostas ao tradutor, assim como das dificuldades suscitadas pela mudança do meio oral para o escrito, os legendistas têm de encontrar soluções para se manterem leais ao original, ao mesmo tempo que produzem traduções idiomáticas que obedeçam aos parâmetros definidos. Por isso, os tradutores frequentemente lançam mão de condensações, reformulações e omissões, que podem acontecer no nível da palavra ou da frase, e são largamente empregadas em todos os tipos de legendas. Ao omitir sujeitos, simplificar tempos verbais, fazer uso de sinônimos e recorrer a pronomes para realizar as retomadas, para citar apenas algumas das estratégias utilizadas pelos legendistas, chega-se a unidades sintáticas simples e livres de repetições, que permitem o entendimento rápido do espectador.

Para garantir a otimização da leitura, ainda outro ponto que deve ser considerado pelos tradutores é a segmentação das frases. Como já anunciado anteriormente, ao inserir uma quebra de linha, é necessário que o legendista a faça de maneira a garantir que a interrupção ocorra de

⁴⁴ “[...] accurate timing is crucial for optimal subtitling since it reinforces the internal cohesion of the translated programme and plays the essential role of helping the viewer identify who is saying what in the programme [...]”.

maneira lógica e sintaticamente apropriada. Isso significa que cada legenda deve apresentar uma carga semântica completa, ou seja, deve fazer sentido por si só, mesmo que não corresponda a uma sentença completa e seja necessário conectá-la com a legenda subsequente. Nesse contexto, caso haja a necessidade de realizar a segmentação, seguindo a orientação de Karamitroglou (1998), a divisão entre as linhas deve corresponder a divisões sintáticas já encontradas nas frases, evitando ao máximo a separação de um mesmo sintagma em duas partes.

Desse modo, o tradutor de legendas precisa dispor de uma série de habilidades linguísticas e técnicas que vão muito além do pleno domínio das línguas de partida e de chegada, a fim de realizar seu trabalho da melhor maneira possível e evitar as críticas que comumente atingem as legendas. Em outras palavras,

Para criar legendas de qualidade, é necessário que a pessoa responsável saiba fazer não só as marcações de tempo e inserção das legendas por meio do *software*, mas também usar de forma satisfatória as estratégias de tradução necessárias para a legendagem, refletindo criticamente sobre suas escolhas e as repercussões dessas e também sobre o papel social de quem traduz [...] (Spolidorio, 2017, p. 41).

Pensando nisso, um último aspecto que ainda deve ser observado é o contexto de produção das legendas, uma vez que, com o passar dos anos, a legendagem deixou de ser uma atividade restrita aos tradutores profissionais, chegando, inclusive, a ser realizada pelos próprios fãs das produções audiovisuais. Nas próximas subseções, serão apresentados esses dois tipos de prática de legendagem, uma vez que as diferenças entre eles são relevantes para a pesquisa aqui desenvolvida.

4.2.1.1 As legendas profissionais

Nesta pesquisa, seguindo o que propôs Spolidorio (2017), considera-se profissional uma legenda “[...] cuja produção foi formalmente encomendada diretamente por cliente ou por empresa de legendagem, havendo a compensação financeira pelos serviços prestados [...]” (p. 46). Trata-se, portanto, da tradução elaborada por um indivíduo com formação e/ou experiência na área suficientes para que se possa considerá-lo um profissional da linguagem, sendo compartilhada através de meios oficiais de distribuição, tais quais o cinema, a televisão, e as plataformas de *streaming*.

Nesse cenário, o processo de tradução passa por diversas etapas, nas quais vários profissionais podem estar envolvidos, conforme descrito por Díaz-Cintas e Remael (2014[2007]). A primeira delas consiste no recebimento da solicitação de tradução,

normalmente feita por uma produtora, uma distribuidora, um canal de televisão ou, em dias atuais, uma plataforma de *streaming*. É nesse instante que a empresa de tradução ou o tradutor recebem uma cópia do material a ser legendado e as instruções de tradução, que devem ser seguidas e auxiliarão na determinação dos métodos e estratégias empregados.

Em seguida, o marcador entra em cena para determinar os momentos exatos em que uma legenda deve aparecer na tela e sair de visualização, de acordo com as limitações espaciais e temporais exploradas na subseção anterior. Preferencialmente, a marcação deve ser feita por um tradutor que já possui experiência na área da TAV, de modo que o trabalho possa ser executado da maneira mais precisa possível. Por vezes, o solicitante da tradução disponibiliza uma lista de marcação que já apresenta o diálogo segmentado de acordo com a língua original, a fim de que os mesmos cortes sejam seguidos na tradução, mas nem sempre é esse o caso.

O próximo passo consiste na tradução do original em si, realizada pelo legendista. Dependendo do contexto, o tradutor pode receber tanto a cópia do produto audiovisual como uma lista de diálogos, sendo mais comum, porém, que receba apenas um dos dois materiais. Caso tenha acesso ao vídeo e o prazo permita, o profissional pode assistir ao vídeo completo antes de prosseguir com seu trabalho, buscando tomar conhecimento do todo da obra. O legendista realiza, então, a tradução do texto audiovisual recebido, também em observância aos parâmetros definidos pelo cliente ou pela empresa de tradução, incluindo os elementos acústicos e visuais que precisam ser traduzidos.

Ao final da tradução, o material é enviado diretamente para a empresa ou o cliente, para a aprovação do trabalho. Caso o tradutor tenha realizado apenas a tradução dos diálogos em si, é necessário que ela passe pelas mãos de um técnico ou adaptador que ajuste o texto de acordo com as limitações estabelecidas para aquelas legendas.

Após essa etapa, as legendas passam por uma revisão fina realizada por um profissional específico. Nela, o revisor busca corrigir quaisquer erros – sejam eles de natureza linguística ou de tradução –, de modo a garantir um produto final de alta qualidade e evitar que as legendas recebam críticas desnecessárias. Idealmente, o tradutor deve receber uma cópia final da tradução, a fim de observar e aprovar as mudanças feitas pelo revisor.

Por fim, a legenda passa por uma simulação na qual o cliente assiste à produção audiovisual já legendada, aprovando-a ou não. Pode-se, então, incorporar os últimos ajustes à

tradução, antes que ela seja finalizada e gravada definitivamente no filme pelo legendador⁴⁵ ou enviada para exibição eletrônica.

Vale mencionar, porém, que o caminho acima descrito é ilustrativo apenas de uma das maneiras de produzir legendas no âmbito profissional, podendo haver diferenças entre o que foi reportado e a maneira como o processo efetivamente toma forma. Com as mudanças tecnológicas que atingem a todo tempo a área do audiovisual – e, conseqüentemente, a da TAV –, por exemplo, muito mudou na área da legendagem, inclusive no que diz respeito aos próprios tradutores que, devido à facilidade de uso dos novos *softwares* de tradução criados nas últimas décadas, se capacitaram para executar a marcação, a tradução e a adaptação das legendas.

Apesar dessa transformação em direção a uma área de atuação polivalente, os legendistas ainda são pouco reconhecidos pelo trabalho que fazem, tanto financeira como socialmente. Mesmo tendo o direito de ver seus nomes exibidos nas obras que traduzem, assim como é feito no caso de traduções de literatura, os tradutores dificilmente constam nos créditos dos produtos audiovisuais, o que, de certa forma, acaba por invisibilizá-los. Além disso, após a remuneração inicial pelo trabalho, os legendistas não recebem quaisquer remunerações correspondentes a direitos autorais, mesmo que suas traduções continuem sendo utilizadas e redistribuídas pelas empresas que as encomendaram.

Quanto aos parâmetros técnicos adotados na elaboração de legendas profissionais, pode-se dizer que as convenções seguidas muito se aproximam daquelas anteriormente mencionadas nesta subseção, uma vez que elas tendem a seguir um padrão internacional mais institucionalizado. Contudo, cada empresa possui seu próprio guia de legendagem, sendo esse documento aquele que, em última análise, define quais serão as regras a serem obedecidas pelo legendista⁴⁶.

4.2.1.2 As *fansubs*

Nas últimas décadas, a popularização das legendas na internet (Díaz-Cintas, 2005), a criação de *softwares* de legendagem gratuitos e a disponibilização ilegal de produções audiovisuais para *download* tiveram como consequência o surgimento de um tipo de legenda até então inédito na TAV: as legendas feitas por fãs.

⁴⁵ Reforça-se, aqui, a diferença entre o legendista e o legendador. Enquanto o primeiro se encarrega da tradução em si, o segundo é responsável pela gravação da legenda no filme – função essa que vem desaparecendo desde o surgimento das legendas digitais.

⁴⁶ Como exemplo, pode-se observar o guia de legendagem do *Amazon Prime Video*, disponível em: https://videocentral.amazon.com/home/help?topicId=GBKB422Q9GYC7DWE&ref_=avd_sup_GBKB422Q9GYC7DWE. Acesso em: 9 fev. 2024.

Convencionalmente chamadas de *fansubs* – termo criado a partir da união das palavras “fã” (“*fan*”) e “legenda” (“*subtitle*”) em inglês –, essas legendas se diferenciam das outras formas de legendagem amadora principalmente por existir uma “[...] interação entre quem produz e quem consome as legendas dentro das comunidades virtuais [...]” (Spolidorio, 2017, p. 46). Nesse sentido, desde que surgiu na década de 80, com a legendagem informal de *animes* para o inglês, o *fansubbing* teve sempre como objetivo a produção de legendas “de fãs para fãs”, caracterizando-se não só como uma atividade de tradução, mas como um “[...] fenômeno social em massa na internet [...]” (Díaz-Cintas; Sánchez, 2006, p. 37, tradução nossa)⁴⁷.

Embora criada nesse contexto específico, desde seu aparecimento, a prática do *fansubbing* ganhou espaço e passou a abranger outros tipos de produção audiovisual além dos *animes*, como filmes, *reality shows* e seriados televisivos, em um movimento motivado, principalmente, pela demora na elaboração e distribuição de legendas comerciais (O’Hagan, 2009). Com essa prática, criou-se uma pressão sobre o mercado televisivo, forçando-o, inclusive, a exibir novos episódios de diversos seriados com mais antecedência, para que pudesse competir com a rápida produção das *fansubs* (Calazans, 2010; Bold, 2011), de modo que ficam claras a força e a proporção que esse tipo de legenda tomou com o passar do tempo.

No cenário brasileiro, de acordo com Spolidorio (2017), a popularização das *fansubs* teve início a partir de 2004, ano de estreia do seriado *Lost*. Considerada, à época, a melhor série em exibição na TV, essa produção angariou inúmeros fãs que se reuniam em comunidades *online* para discutir os mistérios do enredo enquanto novos episódios eram ansiosamente esperados. Essa inquietação, somada à facilidade de acesso ao seriado na internet de maneira ilegal, acabou levando alguns fãs a se unirem para legendar a série mais rapidamente, de modo que se formou a primeira equipe brasileira de legendagem, a Psicopatas.

A partir dessa primeira experiência, porém, diversas outras equipes de fãs se consolidaram no país, também como uma resposta à insatisfação com as legendas oficiais que, por vezes, apresentavam problemas relacionados à sincronia, ao tempo de exibição e, sobretudo, a legendas que pecavam na tradução de referências relevantes para a série. Sobre isso, Spolidorio (2017, p. 84) afirma que

[...] algumas séries representam universos muito particulares e têm um público bastante dedicado, então, quando o conteúdo das legendas é simplificado e generalizado, ou não apresentam a terminologia ou nomenclatura consideradas corretas pelos fãs, as reclamações também aparecem em maior número [...].

⁴⁷ “[...] a mass social phenomenon on Internet [...]”.

Para lidar, portanto, com as especificidades de cada seriado, os grupos de legendagem no Brasil apresentam, em geral, uma composição bastante diversificada, sendo frequente que contem com membros especialistas nos assuntos abordados pelas séries. Por exemplo, não é incomum a presença de estudantes ou profissionais da área médica legendando séries como *Grey's Anatomy*, ou da área do direito trabalhando em seriados como *Law and Order*, além de fãs do universo *geek* traduzindo conteúdos de ficção científica ou super-heróis. Essa familiarização com o contexto da série, tanto em relação ao enredo como aos seus tópicos específicos, parece elevar a qualidade das legendas do ponto de vista dos fãs (Bandeira, 2009), compensando, assim, uma suposta falta de treinamento nas técnicas da legendagem em si (O'Hagan, 2009).

Por muito tempo, essas equipes se organizavam em torno do *site* Legendas.TV, retirado do ar no final de 2022, que servia como um repositório *online* das legendas produzidas pelos vários *legenders*⁴⁸ – autodenominação cunhada pelos tradutores-fãs. Além disso, o *site*, que contava com mais de 300 equipes, com algumas delas sendo responsáveis por mais de 20 séries (Spolidorio, 2017), funcionava como uma comunidade na qual os espectadores poderiam interagir com os tradutores. Assim, era naturalmente necessário haver alguma regulamentação que controlasse o monopólio de cada seriado e definisse as diretrizes que deveriam ser obedecidas para compartilhar as legendas, incluindo o prazo máximo de uma semana para a postagem. Tal tarefa era executada pela própria administração do Legendas.TV, cujas determinações continuam a ser seguida mesmo após o fim do *site*. Desse modo,

[...] as redes de *fansubbing* no Brasil parecem funcionar de maneira bastante organizada para lidar com os prazos extremamente curtos impostos por elas mesmas e atingir as expectativas das comunidades de fãs [...] (Bold, 2011, p. 8, tradução nossa)⁴⁹.

Parte dessa sistematização, inclusive, se reflete na estrutura interna dos grupos de legendagem que, por vezes, contam com processos seletivos, treinamentos e períodos probatórios para seus novos membros, com vistas a garantir a qualidade de seu trabalho. Ainda, apesar de terem seus próprios procedimentos, todas as equipes de legendagem seguem, no geral,

⁴⁸ Como mencionado, o termo “*legender*” foi criado pelos próprios fãs brasileiros que realizavam as traduções. Essa denominação se opõe ao nome “*fansubber*” que, em geral, faz referência específica aos fãs que legendam *animes*.

⁴⁹ “[...] the fansubbing networks in Brazil seem to function in a very organized manner to be able to deal with the extremely short self-imposed deadlines and fulfill the expectations of the fan communities. This is especially true for the translation and distribution of American TV series, a genre that has been attracting more and more the attention of Brazilian audiences ever since the popularization of cable TV in that country [...]”.

as mesmas etapas de produção que uma legenda profissional (Díaz-Cintas; Sanchez, 2006; Bugocki, 2009; Bold, 2011), que geralmente são realizadas por membros diferentes.

Esse processo começa, usualmente, dois ou três dias antes da exibição do episódio no país de origem, quando os membros responsáveis pela série entram em contato com os tradutores para solicitar a participação deles na legendagem. Com a equipe fechada, aguarda-se a disponibilização do episódio na internet após a exibição e, assim que o arquivo é obtido, é elaborada uma escala de legenda com a divisão dos trechos pelos quais cada tradutor ficará responsável. Essa escala é enviada aos tradutores juntamente com uma indicação de onde fazer o *download* do episódio e, se possível, com as legendas em inglês do original. Começa, então, a etapa da tradução, que leva em torno de duas horas, na qual diversos membros trabalham de maneira separada traduzindo os trechos que lhes foram indicados. Segue-se com a etapa da sincronia, usualmente realizada pelo mesmo tradutor, que, ao terminar, envia sua parte para o revisor. Com todas as partes em mãos, o revisor faz as correções necessárias na legenda, adequa a sincronia, resolve problemas de tempo de leitura, entre outros, até que se atinja o produto final, normalmente em dois ou três dias. As legendas passam, ainda, por um processo de resincronização, para que possam ser disponibilizados arquivos diferentes para cada versão do episódio que pode ser baixada *online*, e, por fim, são postadas no repositório para que possam ser utilizadas por outros fãs (InSUBs, 2010, 2013)⁵⁰.

Com tantas pessoas envolvidas na criação de legendas, a prática mais comum entre as equipes de fãs é a elaboração de guias de legendagem, a fim de que as produções do grupo sejam padronizadas, a qualidade do trabalho desenvolvido seja mantida e haja auxílio no treinamento de novos membros. Nesses guias, são encontradas orientações quanto aos parâmetros técnicos a serem obedecidos durante a legendagem, à utilização de programas de legendagem, ao emprego de estratégias para realizar a tradução das falas e à própria linguagem a ser usada na legenda⁵¹. Em alguns casos, apresentam-se até mesmo definições do que seria a tradução, como é o caso do guia da equipe InSUBs⁵², explorado em mais detalhes na seção 6.

Nesse sentido, o *fansubbing* se modificou consideravelmente desde seus primórdios, quando as práticas tradutórias dos fãs se distanciavam das recomendações do âmbito comercial (Díaz-Cintas; Sánchez, 2006), atraindo comentários que desconsideravam sua qualidade. Pelo

⁵⁰ A fonte dessas informações foram postagens realizadas no blog da equipe InSUBs, uma das mais reconhecidas do país. Disponível em: <https://insubs-blog-blog.tumblr.com/>. Acesso em: 30 dez. 2023.

⁵¹ Para traçar essas considerações, foram observados os guias de legendagem do Legendas.TV e de quatro equipes: InSUBs, NERDS, LeS e United.

⁵² O guia da equipe InSUBs pode ser baixado gratuitamente a partir do endereço <https://www.mediafire.com/?vc3w6864e5xtllw>. Acesso em: 9 fev. 2024.

contrário, o que se observou ao longo dos últimos anos foi uma aproximação cada vez maior entre as regras estabelecidas pelas equipes de legendagem e as convenções utilizadas na legendagem profissional. O resultado disso é um produto com cada vez mais qualidade, que une, ao menos em teoria, o melhor dos dois mundos: a rigorosidade dos parâmetros das legendas comerciais, que promovem maior legibilidade ao assistir a esse material, e o envolvimento profundo com a produção audiovisual que só um fã poderia ter.

Tendo em vista o que foi aqui levantado, pode-se dizer, portanto, que as *fansubs* se diferenciam notavelmente das legendas profissionais principalmente no que toca aos seus contextos de produção e aos tradutores envolvidos na tarefa. Por um lado, tem-se a legendagem profissional, cujos interesses são puramente comerciais, havendo um distanciamento entre o tradutor e o público; por outro, as *fansubs*, que se materializam precisamente através da interação, movidas por um desejo de compartilhamento. Dessa forma, no contexto das legendas feitas por fãs,

[...] há uma grande dedicação não só na produção das legendas, mas também em se capacitar para estar apto a produzi-las. Mais do que simplesmente ‘fãs que decidiram traduzir’, *legenders* são uma categoria emocionalmente investida na tarefa de compartilhar o conteúdo midiático do qual são fãs com o resto da comunidade e do *fandom* [...] (Spolidorio, 2017, p. 91).

5 PROCEDIMENTOS METODOLÓGICOS

A pesquisa aqui descrita, como explicitado na seção 1, possui dois grandes objetivos: analisar como ocorre a realização do OD anafórico em legendas audiovisuais, levando em conta a influência de características específicas desse gênero e averiguando, também, se há diferenças na expressão do fenômeno em legendas profissionais e *fansubs*; e descrever a legenda como um gênero textual-discursivo a partir de seus aspectos estruturais e situacionais. Dessa maneira, para o desenvolvimento da pesquisa, diversos procedimentos metodológicos foram empregados.

Começou-se, naturalmente, com a composição de um *corpus* de legendas audiovisuais que incluiu tanto produções profissionais como amadoras. A partir desse material, foi realizada tanto a descrição estrutural da legenda, considerando as características descritas na seção 4, como a abordagem situacional desse gênero, a partir dos parâmetros propostos por Biber e Conrad (2009).

Em seguida, com o mesmo material, foi elaborada uma amostra-piloto para auxiliar na definição do objeto de estudo, uma vez que, inicialmente, eram três os fenômenos variáveis que seriam potencialmente analisados. Por se tratar de uma investigação essencialmente variacionista, os passos prototípicos de uma pesquisa sociolinguística foram então cumpridos, tanto na observação da amostra-piloto como no trabalho com a amostra final completa. Assim, o presente estudo passou pelos seguintes estágios, tendo sempre em mente o gênero textual-discursivo abordado: (i) definição da variável dependente; (ii) seleção e coleta de todos os dados dos fenômenos variáveis em jogo; (iii) definição das variáveis independentes que possivelmente motivariam a utilização de uma ou outra variante; (iv) codificação dos dados coletados; (v) realização de análises estatísticas uni e multivariadas por meio da plataforma R (R Core Team, 2023); e (vi) descrição e análise dos resultados.

Nas subseções seguintes, relata-se, portanto, cada uma das etapas empreendidas nesta pesquisa. Primeiramente, descreve-se a composição do *corpus* de análise, passando pela escolha da série, os recortes realizados, a obtenção e o tratamento do material. Em seguida, apresenta-se o trabalho realizado com a amostra-piloto, detalhando sua elaboração e o caminho percorrido até a definição do objeto de estudo. São traçadas, então, considerações acerca da coleta, da codificação e das análises estatísticas empreendidas, seguidas pela identificação das variáveis do estudo, incluindo as extralinguísticas, relacionadas aos aspectos estruturais e situacionais do gênero textual-discursivo em questão. Por fim, explicita-se o procedimento adotado para analisar os aspectos situacionais do gênero legenda.

5.1 A elaboração do *corpus*

5.1.1 Definição da série e recortes metodológicos

Na perspectiva da Linguística de *Corpus*, que apresenta pressupostos metodológicos para a composição de amostras linguísticas, um *corpus* é

[...] um conjunto de dados linguísticos (pertencentes ao uso oral ou escrito da língua, ou a ambos), sistematizados segundo determinados critérios, suficientemente extensos em amplitude e profundidade, de maneira que sejam representativos da totalidade de uso linguístico ou de algum de seus âmbitos, dispostos de tal modo que possam ser processados por computador, com a finalidade de propiciar resultados vários e úteis para a descrição e análise [...] (Sanchez, 1995, p. 8-9, *apud* Sardinha, 2000, p. 338)

Com isso em mente, e considerando, também, os objetivos do estudo ora descrito, foi constituído um *corpus* de legendas audiovisuais com duas amostras: uma de legendas profissionais e outra de *fansubs*, já que, por hipótese, esses dois expoentes do mesmo gênero poderiam gerar resultados diferentes. Ambas as amostras correspondem a legendas para os mesmos episódios da série televisiva *Grey's Anatomy*, a fim de que se possa traçar uma comparação entre elas.

Tal produção foi escolhida, assim como mencionado anteriormente (cf. Introdução), por alguns motivos. O primeiro deles foi a consolidada popularidade da produção norte-americana no Brasil, onde a série é atualmente exibida no canal de televisão por assinatura *Sony* e na plataforma de *streaming* *Star+* (o seriado também já fez parte do catálogo de outros três serviços de *streaming* no país, sendo eles *Netflix*, *Amazon Prime Video* e *Globoplay*). Desde que chegou ao Brasil, o drama médico alcançou um público dilatado e diverso por meio das dublagens e legendagens disponibilizadas, dando origem a inúmeros grupos e páginas de fãs nas redes sociais, que são acompanhados por milhares de usuários. A título de ilustração, podem ser mencionados, por exemplo, os grupos *Grey's Anatomy* Brasil e Loucos por *Grey's Anatomy*, no *Facebook*, que contam, no momento da escrita desta dissertação, com 399 mil e 494 mil membros, respectivamente; a página *Grey's Anatomy* Brasil, na mesma rede social, com 362 mil curtidas; e o perfil *Grey's* Depressão no *Instagram*, seguido por 690 mil pessoas (cf. nota de rodapé 7).

Outra razão para selecionar as legendas de *Grey's Anatomy* como foco desta investigação foi a longevidade da série, que estreou em 2005 na emissora americana *ABC* e continua a ser transmitida atualmente, alcançando 19 temporadas. Ao longo de seus 18 anos de existência, o seriado levou ao ar 420 episódios focados na vida pessoal e profissional de

médicos cirurgiões de um hospital em Seattle, retratando uma gama de interações e situações comunicativas, o que gerou um material rico e extenso para pesquisas em diversas áreas da linguística.

Devido ao desejo de se trabalhar com legendas profissionais e amadoras, o passo seguinte foi determinar quais *fansubs* seriam consideradas e de qual plataforma de *streaming* seriam extraídas as legendas oficiais. Quanto às legendas feitas por fãs, seguiu-se um levantamento das equipes que, ao longo dos anos, legendaram a série *Grey's Anatomy* de maneira voluntária. Diante das informações encontradas nos *sites* *Legendas.TV* e *Addic7ed*, ambos repositórios *online* de legendas, constatou-se que a equipe InSUBs foi aquela que legendou a série por mais tempo, entre a 4ª e a 16ª temporadas, motivo pelo qual se decidiu trabalhar com as legendas por ela produzidas. Trata-se, também, de um grupo que gozou de um *status* renomado no país, tendo legendado mais de 350 produções durante seus 12 anos de existência e chegando a contar com mais de mil colaboradores, entre tradutores, revisores e administradores⁵³. Ademais, o guia de legendagem utilizado pela equipe para padronização pode ser atualmente encontrado na internet, o que facilitou a identificação dos parâmetros técnicos adotados pelo grupo na elaboração das legendas e permitiu uma visão mais ampla da maneira como a InSUBs enxergava as questões linguísticas envolvidas em seu trabalho, fator que poderia, por hipótese, influenciar a variação encontrada no material.

No que diz respeito às legendas profissionais, observou-se, à época da elaboração do *corpus* desta pesquisa, quais plataformas de *streaming* disponibilizavam a série em seu catálogo no Brasil. Dentre as três, *Amazon Prime Video*, *Globoplay* e *Star+*, optou-se por analisar as legendas veiculadas pela primeira, principalmente por duas razões. Primeiramente, considerou-se o fato de que era ela a única cujo guia de legendagem elaborado pela empresa para o PB também poderia ser acessado *online*, de modo a favorecer a comparabilidade dos materiais obtidos para ambas as amostras de legendas. Em segundo lugar, por ser um *streaming* mais antigo, já existiam maneiras automáticas de extrair as legendas do *Amazon Prime Video* (cf. subseção 5.1.2), o que promoveria uma maior rapidez para a compilação das legendas da amostra.

Feita essa seleção, foi necessário, então, realizar um recorte para determinar de que temporadas seriam as legendas que viriam a compor o *corpus*, posto que, dadas a extensão da série e as ferramentas escolhidas para manusear essa amostra, seria inviável, no prazo disponível para a execução desta pesquisa, considerar todo o material disponível. A primeira

⁵³ Disponível em: <https://entretenimento.uol.com.br/noticias/redacao/2020/01/24/insubs-site-de-legendas-anuncia-o-fim-apos-12-anos-e-gera-comocao.htm>. Acesso em: 29 nov. 2023.

decisão foi desconsiderar as três primeiras e as três últimas temporadas do seriado, já que não havia legendas criadas pela equipe InSUBs para esses episódios. Diante das temporadas que restaram – 4ª à 16ª –, tinha-se a hipótese de que o guia encontrado para as *fansubs* muito provavelmente não se aplicaria às primeiras legendas elaboradas pelo grupo, uma vez que o próprio documento é resultado de uma reelaboração dos padrões da equipe. Nesse sentido, não seria possível garantir que as legendas daquela época foram criadas tendo em mente os mesmos parâmetros técnicos posteriormente definidos, o que dificultaria o trabalho com esses aspectos. Por isso, optou-se por descartar, também, as duas primeiras temporadas legendadas pela equipe. Com vistas a manter uma simetria no *corpus*, foram excluídas, ainda, as duas últimas temporadas que apresentavam legendas feitas pela InSUBs. Pela mesma razão, estabeleceu-se, por fim, que seriam consideradas as legendas da 6ª, da 10ª e da 14ª temporadas da série, buscando manter o mesmo intervalo entre as escolhidas.

5.1.2 Obtenção e tratamento das legendas

Com os recortes acima realizados, a etapa seguinte foi a obtenção das legendas, que ocorreu de maneira diferente para cada amostra: no caso das legendas feitas por fãs, fazendo o *download* dos arquivos e, no das profissionais, realizando a extração das legendas a partir do *site* do *Amazon Prime Video*. Ambos os processos serão descritos em mais detalhes a seguir.

Quanto às *fansubs*, decidiu-se, em primeiro lugar, baixar as legendas do repositório *online* Legendas.TV⁵⁴, que, nas palavras da equipe InSUBs,

[...] É uma comunidade que reúne todas as pessoas que gostam de legendar, e é o lugar onde se concentra a organização da legendagem. É a partir desta comunidade que são divididas as séries (cada equipe fica com uma, para não ter problema de termos várias pessoas trabalhando na mesma série) [...] (INSUBS, 2013).⁵⁵

A escolha por esse *site*, que compartilhava legendas de filmes e séries elaboradas por equipes informais de maneira gratuita e ficou no ar de 2009 a 2022, se deu não só pelo fato de a página existir já há bastante tempo, apesar de esse ter sido um fator, mas também em razão da proximidade estabelecida entre ele e as comunidades de fãs através de suas redes sociais, do contato direto entre os administradores do *site* e as equipes de legendagem, e da rapidez com

⁵⁴ À época da composição do *corpus* desta pesquisa, o Legendas.TV podia ser acessado em: <<http://legendas.tv/>>. Desde o fim dos trabalhos do *site*, porém, ao abrir a página, encontra-se uma mensagem de despedida da equipe, justificando o porquê de encerrarem as atividades do repositório.

⁵⁵ Disponível em: <https://insubs-blog-blog.tumblr.com/>. Acesso em: 02 dez. 2023.

que as legendas eram disponibilizadas no domínio. Com efeito, para que as equipes de legendagem pudessem ter seu trabalho em destaque no *site* – havia, em seu *layout*, um espaço reservado para legendas recentes e em andamento, dando proeminência às equipes mais prolíficas –, era necessário que as legendas fossem enviadas até uma semana após a exibição do episódio no exterior, garantindo agilidade no processo de produção.

Para fazer o *download* das legendas, que ocorreu de maneira gratuita, foi necessário apenas criar uma conta de usuário no Legendas.TV. Isso feito, buscou-se por cada uma das temporadas selecionadas e verificou-se que era possível baixar as legendas para cada episódio separadamente ou para a temporada completa de maneira conjunta. Em virtude da praticidade, os arquivos foram baixados em três *packs*, como são chamados os arquivos que compilam mais de uma legenda para um seriado, sendo um para cada temporada e tomando-se o cuidado de verificar se todas as legendas, de fato, haviam sido produzidas pela equipe InSUBs. Em seguida, os arquivos foram extraídos através do programa 7-Zip (Pavlov, 2023), ficando, então, prontos para a próxima etapa.

Já no que concerne às legendas profissionais, o processo ocorreu de maneira completamente diversa. A fim de evitar ter que fazer a transcrição manual das legendas dos episódios veiculados nas plataformas de *streaming*, investigou-se a possibilidade de extrair os arquivos diretamente dos *sites*. A princípio, considerou-se seguir o passo a passo sugerido por Silva (2019) para a compor *corpora* de legendas, que consistia na obtenção dos arquivos via acesso às ferramentas do desenvolvedor do navegador. Entretanto, dessa maneira seria necessário reproduzir cada um dos episódios e buscar pelos arquivos de legenda, fazer o *download* e convertê-los em *.srt* (o formato mais utilizado em legendas, que pode ser aberto através do Bloco de Notas, aplicativo nativo do *Windows*), uma vez que, utilizando esse método, as legendas eram baixadas na extensão *.html*, o que demandaria mais tempo para a composição do *corpus*. Por isso, decidiu-se explorar alternativas que facilitassem esse processo.

Ao buscar outras maneiras de fazer a extração de legendas de plataformas de *streaming*, encontrou-se, então, a possibilidade de utilizar um *script* de usuário elaborado especificamente para esse fim, método que foi empregado nesta pesquisa. *Scripts* de usuário são programas comumente escritos na linguagem *JavaScript*, cuja função essencial é modificar aspectos de uma página na internet e aumentar sua navegabilidade, sendo o único requisito para sua aplicação a instalação de um gerenciador de *scripts*. Disponibilizados gratuitamente em repositórios *online*, esses códigos prontos têm a capacidade, dentre outras coisas, de adicionar botões, criar atalhos, controlar a velocidade de reprodução de vídeos, melhorar o histórico de

navegação, entre outros, já que o autor de cada *script* tem a liberdade de criar diferentes funções de acordo com seus objetivos.

Neste estudo, a extração das legendas profissionais foi realizada por intermédio do *script Amazon Video – subtitle downloader*⁵⁶, versão 1.9.3, disponibilizado para *download* pelo usuário Tithen-Firion no *site Greasy Fork*⁵⁷, um repositório de *scripts* gratuitos mantido pelo programador canadense Jason Barnabe. Para adicioná-lo ao *Google Chrome*, foi instalada ainda a extensão *Tampermonkey*⁵⁸, desenvolvida por Jan Biniok, que pode ser adicionada gratuitamente a diversos navegadores.

Quando ativado, o *Amazon Video – subtitle downloader* cria, na página do *Amazon Prime Video*, botões que permitem o *download* de todas as legendas disponíveis para um conteúdo, seja ele um filme ou um seriado (Figura 8). No caso das séries, o *script* adiciona, também, um botão que possibilita que o usuário baixe todas as legendas para uma temporada de uma só vez (Figura 9), agilizando consideravelmente o processo de extração das legendas. Os arquivos na extensão *.srt* são, então, baixados em uma pasta compactada, que pode ser extraída para se obter acesso às legendas em PB⁵⁹.

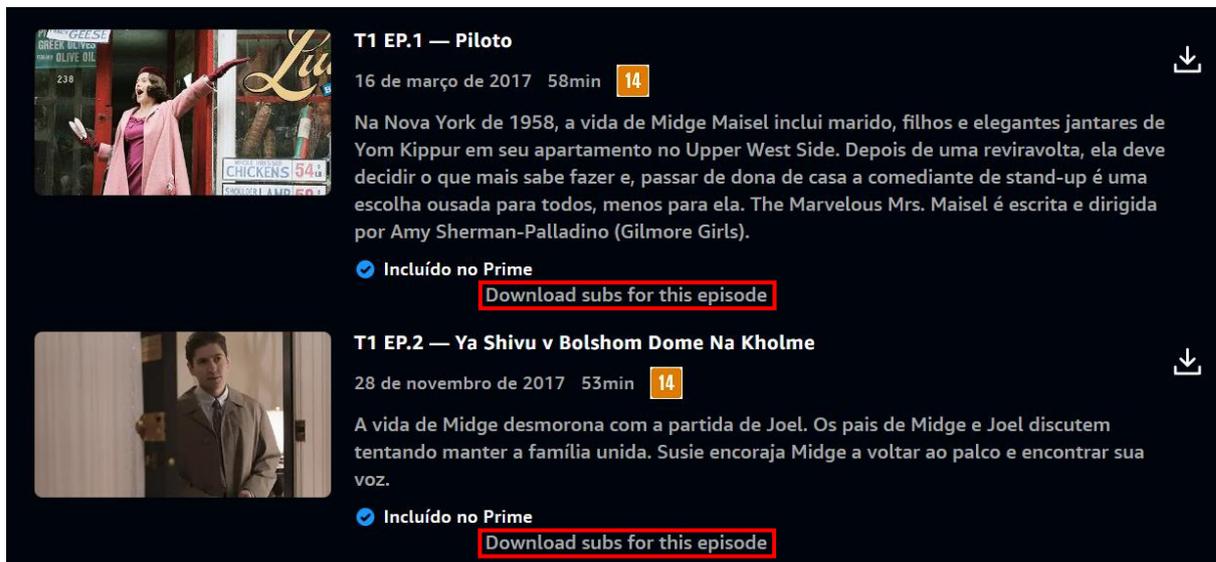
⁵⁶ Disponível em: <https://greasyfork.org/pt-BR/scripts/34885-amazon-video-subtitle-downloader>. Acesso em: 02 dez. 2023.

⁵⁷ Disponível em: <https://greasyfork.org/>. Acesso em: 02 dez. 2023.

⁵⁸ Disponível em: <https://www.tampermonkey.net/>. Acesso em: 02 dez. 2023. A extensão está disponível para os navegadores *Google Chrome*, *Microsoft Edge*, *Firefox*, *Opera Next* e *Safari*, e pode ser baixada tanto através das lojas de extensões de cada um deles como de seu *site* oficial.

⁵⁹ A título de informação, vale mencionar que existem alguns *scripts* de usuários com as mesmas funções para a *Netflix* e o *Hulu*, além de códigos para a extração de legendas do *YouTube*. Para outras plataformas de *streaming*, porém, não há conhecimento sobre *scripts* com essas características. Nesse caso, sugere-se seguir o passo a passo proposto por Silva (2019).

Figura 8 - Captura de tela do *site do Amazon Prime Video* após a ativação do *script* de usuário, mostrando o botão para *download* das legendas de um episódio



Fonte: *Amazon Prime Video*

Figura 9 - Captura de tela do *site do Amazon Prime Video* após a ativação do *script* de usuário, mostrando o botão para *download* das legendas de uma temporada



Fonte: *Amazon Prime Video*

Uma vez obtidas todas as legendas, prosseguiu-se com o próximo passo: o tratamento dos arquivos para a coleta de dados. Tal estágio foi necessário, pois todos os arquivos de legenda, de ambas as amostras, apresentavam informações características desse gênero textual-

discursivo que, porém, não são relevantes para esta pesquisa, tais como a numeração das legendas e as marcações de tempo (*timestamps*), que prejudicariam, inclusive, a contagem de palavras para cada legenda. A Figura 10 ilustra a formatação original das legendas, para que se possa melhor visualizar as características descritas acima:

Figura 10 - Captura de tela do arquivo original da *fansub* para o episódio 6x10 de *Grey's Anatomy*

```

FT06E01 - Bloco de Notas
Arquivo Editar Formatar Exibir Ajuda
1
00:00:02,912 --> 00:00:06,880
De acordo com Elizabeth Kubler
Ross, quando estamos morrendo,

2
00:00:06,915 --> 00:00:09,667
ou sofremos
uma perda catastrófica,

3
00:00:09,702 --> 00:00:13,229
passamos por 5 estágios de luto.

4
00:00:13,264 --> 00:00:15,451
Passamos pela negação.

5
00:00:15,600 --> 00:00:17,000
Mais uma ampola de cálcio...

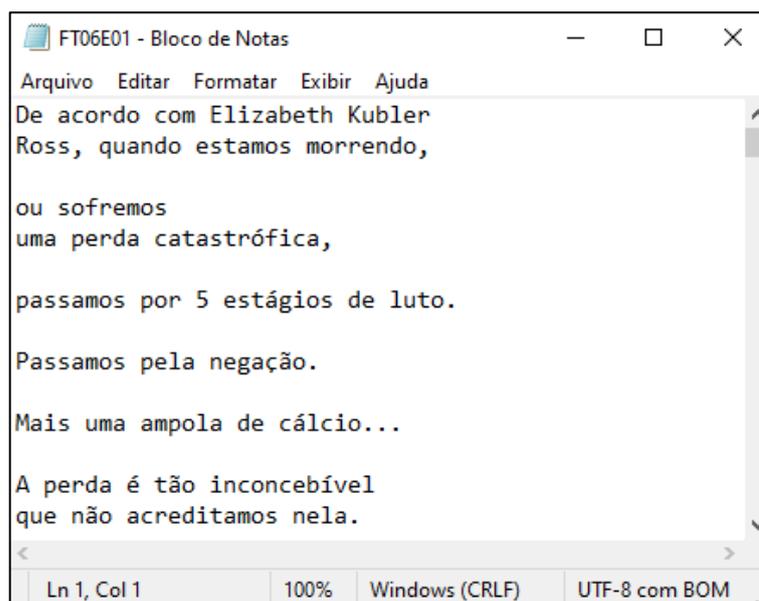
6
00:00:17,001 --> 00:00:21,106
A perda é tão inconcebível
que não acreditamos nela.
Ln 1, Col 1 100% Windows (CRLF) ANSI

```

Fonte: Elaborada pela autora

Com o intuito de realizar essa “limpeza” dos arquivos, deixando-os somente com o texto das falas traduzidas, foi utilizado o *site Subtitle Tools*, que oferta ferramentas gratuitas para converter, sincronizar, limpar, corrigir, combinar e formatar legendas, tudo de modo *online*. Nesse contexto, aplicou-se o conversor de arquivos *.srt* para texto simples, que funciona da seguinte maneira: (i) faz-se o *upload* das legendas que se deseja converter, havendo um limite de 20 arquivos por conversão na versão gratuita do serviço; (ii) escolhem-se as características a serem mantidas, se for esse o caso (pode-se optar por manter uma linha vazia entre as falas, as marcações de tempo e o nome do arquivo); (iii) inicia-se a conversão, sendo que, ao final, todos os arquivos ficam disponíveis para *download* no formato *.txt*. Para esta pesquisa, foram mantidas apenas as linhas entre as falas, para que a visualização dos turnos fosse mais clara. Ao final da conversão, a legenda da figura acima ficou como pode ser observado a seguir:

Figura 11 - Captura de tela do arquivo convertido da *fansub* para o episódio 6x10 de *Grey's Anatomy*



Fonte: Elaborada pela autora

Por fim, a última parte da preparação das legendas para a coleta de dados foi a identificação das personagens em cada uma das falas, a fim de que, depois, essa informação pudesse ser retomada e controlada. Desse modo, todos os episódios que viriam a ser analisados (cf. Tabela 4) foram assistidos atentamente e, a cada duas ou três falas, pausados para que as anotações pudessem ser feitas tanto na *fansub* como na legenda profissional. Um código composto por uma combinação de letras e números foi dado a cada personagem que aparecia, diferenciando-se as falas realizadas em conversas e em narração (todo episódio de *Grey's Anatomy* começa e termina com o monólogo de uma personagem, geralmente a protagonista, realizado fora de cena, assim como as linhas que traduziam texto em tela, músicas e que correspondiam a trechos em que não era possível identificar qual personagem proferia a fala. A título de exemplo, o código M1 foi colocado no início das falas da protagonista Meredith Grey, enquanto M1N correspondia às falas da mesma personagem em contextos de narração, e TT1 foi a combinação atribuída às linhas de texto em tela. A lista com os códigos das personagens que aparecem nos exemplos trazidos nesta dissertação pode ser conferida no Apêndice A.

5.2 Das legendas ao objeto de estudo: o trabalho com a amostra-piloto

Nesta pesquisa a proposta era definir o objeto de estudo a ser investigado a partir do material reunido para o *corpus*, sendo que, de início, eram três os fenômenos variáveis que potencialmente seriam analisados: a expressão de futuro, a colocação pronominal e a realização

do OD anafórico. Assim, após a reunião das duas amostras que compõem o *corpus*, decidiu-se elaborar uma amostra-piloto, a partir da qual todos os dados correspondentes às três variáveis dependentes foram coletados e codificados de maneira simplificada. Foram feitas, então, análises preliminares com o objetivo de observar como esses fenômenos pareciam ocorrer nas legendas, buscando compreender qual deles era o mais relevante no *corpus*. Todo o procedimento relacionado a esse passo da pesquisa será detalhado a seguir.

Diante da necessidade de decidir o fenômeno variável que seria analisado neste estudo, foi composta uma amostra-piloto inicial com o primeiro episódio de cada temporada selecionada, tanto para legendas amadoras como para as profissionais. Trabalhou-se, portanto, com uma subamostra de *fansubs* com 12759 palavras, e uma de legendas profissionais com 12875, somando, no total, 25634 palavras. A distribuição do número de palavras por episódio pode ser conferida na Tabela 2:

Tabela 2 - Número de palavras na amostra-piloto inicial, de acordo com episódio e tipo de legenda

Episódio	<i>Fansub</i>	Profissional
6x01	3876	4097
10x01	4628	4519
14x01	4255	4259
Total	12759	12875

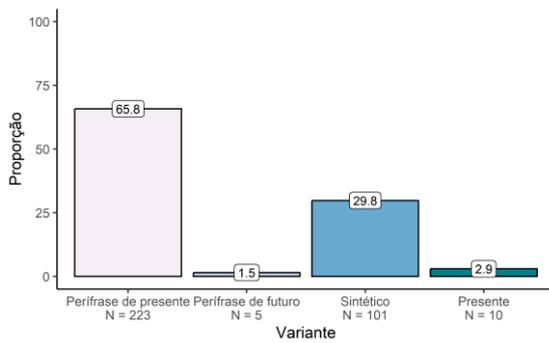
Fonte: Elaborada pela autora

Com a amostra-piloto pronta, o próximo passo foi um breve levantamento da literatura acerca dos três fenômenos em questão, para que se pudesse definir as variantes de cada variável dependente com segurança. Sobre a expressão de futuro, observaram-se os trabalhos de Gibbon (2000), Oliveira (2006), Bragança (2008) e Tesch (2011), que levaram à definição das variantes perífrase de ir no futuro simples + infinitivo e de ir no presente + infinitivo, futuro sintético e presente do indicativo; sobre a colocação pronominal, os de Vieira (2002), Peterson (2010), Biazolli (2010, 2016) e Lima (2022), resultando na seleção de próclise e ênclise como variantes; e sobre a realização do OD anafórico, os de Duarte (1986), Freire (2005), Marques de Sousa (2021) e Lima (2022), com base nos quais se elegeu o clítico pronominal, o objeto nulo, o pronome lexical e o SN anafórico como variantes.

Todos os dados de cada uma das variantes apontadas nos estudos foram, então, coletados manualmente e codificados em uma planilha do *Excel* de acordo com o tipo de legenda em que ocorreram. Em seguida, o material foi submetido à plataforma *R* (R Core Team, 2023), a fim

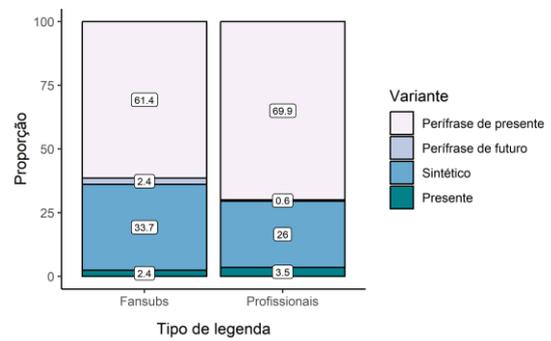
de calcular frequências e proporções e realizar testes exatos de Fisher (explicado em mais detalhes na subseção 5.4.3), buscando identificar se havia diferenças significativas entre a realização de cada fenômeno nos dois tipos de legenda. Os resultados dessas análises preliminares, assim como as decisões tomadas com base neles, são apresentados de maneira resumida nos gráficos abaixo.

Gráfico 1 - Frequências e proporções das variantes da expressão de futuro na amostra-piloto inicial



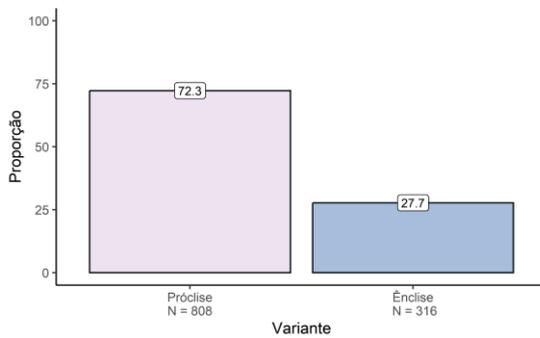
Fonte: Elaborado pela autora

Gráfico 2 – Expressão de futuro na amostra-piloto inicial por tipo de legenda



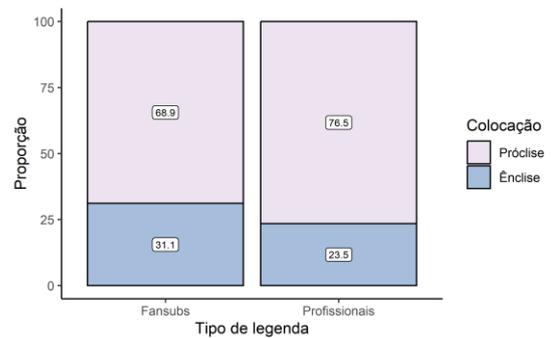
Fonte: Elaborado pela autora

Gráfico 3 - Frequências e proporções de próclise e ênclise na amostra-piloto inicial



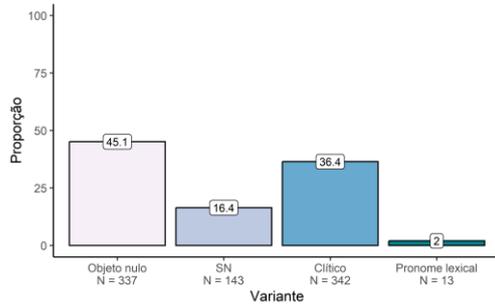
Fonte: Elaborado pela autora

Gráfico 4 – Colocação pronominal na amostra-piloto inicial por tipo de legenda



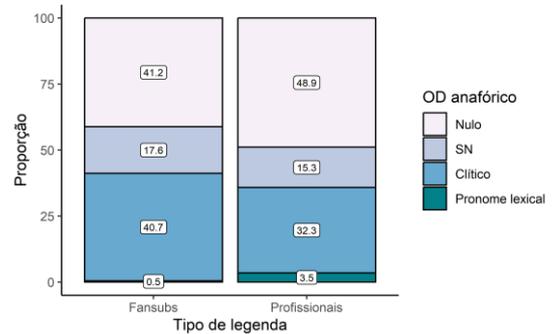
Fonte: Elaborado pela autora

Gráfico 5 - Frequências e proporções das variantes da realização do OD anafórico na amostra-piloto inicial



Fonte: Elaborado pela autora

Gráfico 6 - Realização do OD anafórico na amostra-piloto inicial por tipo de legenda



Fonte: Elaborada pela autora

Diante dos resultados obtidos para os três fenômenos apontados, apresentados resumidamente acima, decidiu-se manter a colocação pronominal e a realização do OD anafórico como objetos que, possivelmente, seriam analisados com mais afinco nesta pesquisa, pelo fato de seus resultados terem, aparentemente, contrariado algumas hipóteses iniciais da pesquisa. A fim de, finalmente, tomar essa decisão, propôs-se a expansão da amostra-piloto para uma nova rodada de testes com os dois fenômenos restantes. Passou-se, então, a considerar não só os primeiros episódios de cada temporada, mas, também, os últimos. Assim, a configuração da amostra-piloto após sua expansão foi a seguinte:

Tabela 3 - Número de palavras na amostra-piloto expandida, de acordo com episódio e tipo de legenda

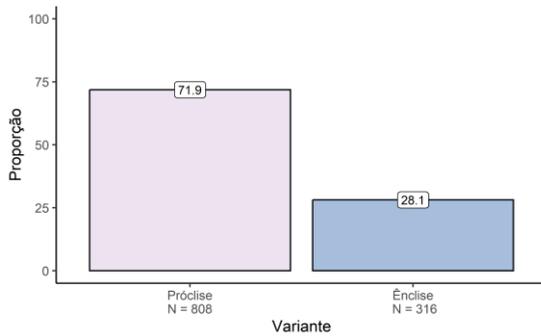
Episódio	Fansub	Profissional
6x01	3876	4097
6x24	3285	3445
10x01	4628	4519
10x24	4976	5367
14x01	4255	4259
14x24	4555	4262
Total	25575	25949

Fonte: Elaborada pela autora

Com os novos episódios incluídos na amostra-piloto expandida, foram coletadas e codificadas as novas ocorrências, com base nos mesmos critérios, para que as frequências e as proporções pudessem, novamente, ser calculadas. O mesmo tipo de teste realizado com a

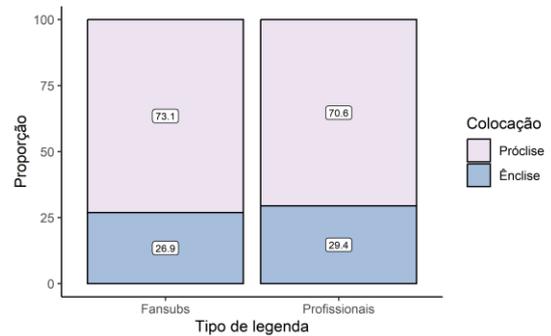
amostra-piloto inicial foi repetido na expandida, tanto para a colocação pronominal como para a realização do OD anafórico. Os resultados para essas etapas, assim como a decisão final quanto ao objeto de estudo, são apresentados a seguir.

Gráfico 7 - Frequências e proporções de próclise e ênclise na amostra-piloto expandida



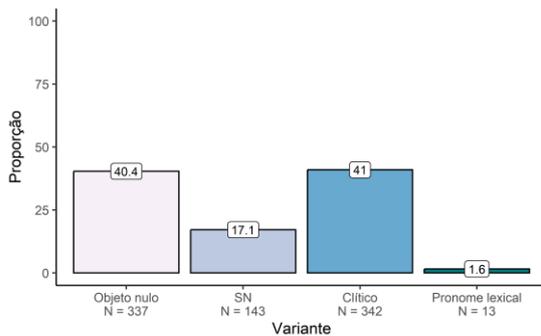
Fonte: Elaborado pela autora

Gráfico 8 - Colocação pronominal na amostra-piloto expandida por tipo de legenda



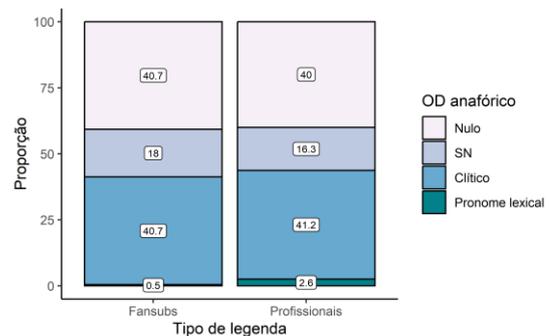
Fonte: Elaborado pela autora

Gráfico 9 - Frequências e proporções das variantes da realização do OD anafórico na amostra-piloto expandida



Fonte: Elaborado pela autora

Gráfico 10 - Realização do OD anafórico na amostra-piloto expandida por tipo de legenda



Fonte: Elaborado pela autora

Tendo em vista, todos os resultados obtidos ao longo do trabalho com as amostras-piloto, optou-se, finalmente, por eleger a realização do OD anafórico como objeto de estudo desta pesquisa. Tal decisão se deu por ser esse um fenômeno com quatro variantes, com muita variação entre fala e escrita, o que o tornaria de especial interesse nas legendas, um material híbrido que lida, o tempo todo, com o limiar entre a fala e a escrita em sua concretização. Além disso, surpreendeu a grande expressividade de clíticos pronominais nas amostras, dado que essa

é uma estratégia quase improdutiva na fala, modalidade que, ao menos em teoria, a legenda deveria tentar emular.

5.3 A composição final do *corpus*

A partir das decisões metodológicas expostas na subseção 5.1, assim como das necessidades identificadas a partir do trabalho com a amostra-piloto, descrito na subseção 5.2, decidiu-se adicionar mais um episódio de cada temporada ao *corpus*, visando investigar não só se as distribuições encontradas na amostra-piloto se manteriam com a inclusão de mais episódios no estudo, mas também quais outros fatores, linguísticos, extralinguísticos e relacionados ao gênero textual-discursivo, condicionariam a escolha entre uma ou outra variante.

Desse modo, optou-se por adicionar ao *corpus* o décimo terceiro episódio de cada uma das temporadas, que normalmente corresponde à *winter finale*, ou seja, o último episódio transmitido nos Estados Unidos antes da pausa nas produções para o inverno. Tal escolha se deu, por, geralmente, serem abordados pontos importantes do enredo, de maneira que eles, assim como o primeiro e o último episódios das temporadas, costumam atrair um grande número de telespectadores.

Uma vez executados os mesmos passos descritos anteriormente para o tratamento das legendas (cf. subseção 5.1.2), chegou-se a um conjunto de 18 legendas, correspondentes às legendas profissionais e às *fansubs* para 9 episódios de 3 temporadas diferentes da série *Grey's Anatomy*. Assim, a elaboração do *corpus* foi finalizada com um total de 81886 palavras, 41397 delas provenientes de legendas profissionais e 40489 de *fansubs*, ultrapassando o número de palavras utilizado em outros trabalhos como os de Biazolli (2016) e de Lima (2022), conforme se apresenta na tabela abaixo.

Tabela 4 - Número de palavras no *corpus*, de acordo com episódio e tipo de legenda

Episódio	Fansub	Profissional
6x01	3876	4097
6x13	5170	4941
6x24	3285	3445
10x01	4628	4519
10x13	4888	4921
10x24	4976	5367
14x01	4255	4259
14x13	4856	5586
14x24	4555	4262
Total	40489	41397

Fonte: Elaborada pela autora

5.4 Trabalhando com os dados: procedimentos de coleta, codificação e análise

Objetiva-se, aqui, descrever os procedimentos metodológicos adotados durante o trabalho com os dados. A subseção 5.4.1 expõe como foi feita a coleta de todos os dados; a 5.4.1 aborda a codificação realizada; e a 5.4.3 se ocupa das análises estatísticas empregadas.

5.4.1 Coleta dos dados

Após a composição do *corpus* de legendas e das duas rodadas da amostra-piloto, o passo seguinte da pesquisa foi a coleta das ocorrências do OD anafórico de terceira pessoa nos episódios que ainda não haviam sido explorados, considerando todos os contextos em que esse fenômeno poderia ser observado no gênero.

Da mesma maneira que o descrito para a amostra-piloto, toda a coleta foi feita manualmente, mesmo havendo ferramentas – tais como o programa *Antconc* ou pacotes como o *dmsocio* (Oushiro, 2018) para a plataforma *R* – que, através da busca por expressões regulares, executam a localização e a extração dos dados. Tal decisão se deu, pois, tendo em vista as características do fenômeno em questão, a extração do objeto nulo e do SN anafórico dependeria, ainda, da utilização de um etiquetador sintático ou outra ferramenta que pudesse auxiliar nesse passo, o que não era viável à época da coleta de dados para esta pesquisa. Além disso, caso a extração dos clíticos e dos pronomes lexicais fosse feita automaticamente, as buscas por essas estruturas retornariam também resultados que não interessariam à pesquisa (como os artigos definidos e os pronomes *ele* e *ela* em função de sujeito), de modo que seria

necessário conferir e filtrar as ocorrências extraídas pela ferramenta utilizada. Assim, como já seria necessário realizar a leitura de todos os textos do *corpus* para identificar as ocorrências de objetos nulos e SNs anafóricos, decidiu-se realizar a coleta de todas as variantes de maneira manual.

Além disso, a coleta manual dos dados do OD anafórico levou à observação não só das ocorrências do fenômeno, mas, também, da própria legenda e suas características estruturais e situacionais, promovendo o contato tão necessário para que esse gênero textual-discursivo pudesse ser analisado e, posteriormente, descrito. Foi nesse momento que se percebeu certa regularidade nos tópicos das conversas entre as personagens e nas relações entre elas, o que possibilitaria a quantificação e o controle estatístico desses dois aspectos situacionais das legendas como variáveis a serem consideradas neste estudo, assim como descrito na subseção 5.5.1.2 à frente.

Todas as legendas foram, então, cuidadosamente lidas em três momentos diferentes, para que se pudesse ter certeza de que todos os dados haviam sido coletados. Ademais, ocorrências que geraram dúvidas foram coletadas e marcadas para, posteriormente, serem revisadas e discutidas, de modo a garantir que todos os dados incluídos, de fato, seriam representativos do fenômeno.

5.4.2 Codificação dos dados

Com todas as ocorrências do *corpus* coletadas, passou-se à próxima etapa de uma pesquisa sociolinguística: a codificação dos dados de acordo com todas as variáveis definidas para o estudo, definidas também a partir do gênero textual-discursivo e que serão mais amplamente exploradas na subseção 5.5.

Para que isso fosse feito, todos os dados foram organizados em uma planilha do *Excel*, salva no formato .csv, de acordo com as convenções do modelo “caso por variável”, sugerido por Gries (2019). Foram seguidas, também, as recomendações do autor para a formatação do texto na planilha, tais como utilizar caixa alta para o nome das variáveis e não fazer uso de caracteres especiais na codificação das variáveis.

Ao codificar as variáveis, lançou-se mão de diversas ferramentas do *Excel*, especialmente de filtros, do preenchimento automático e da fórmula =SE (esta última adaptada do que foi proposto por Lima (2022)), evitando que erros de codificação fossem cometidos e automatizando as partes do processo que fossem possíveis. Ainda, para codificar as variáveis número de CPL e número de CPS – apresentadas na subseção 5.5.1.3 adiante –, os arquivos

originais das legendas foram inseridos no programa de legendagem *Aegisub* (Aegisub, 2014), que mostra automaticamente essas duas informações, facilitando o processo.

5.4.3 Análises estatísticas

Conforme explorado na seção 2, uma pesquisa que se encaixa no âmbito da Sociolinguística Variacionista deve sempre realizar um tratamento estatístico dos dados, buscando descobrir quais variáveis exercem influência sobre a escolha de uma ou outra variante, tanto separadamente como quando consideradas em conjunto. Dessa maneira, espera-se que, em algum momento, todos os dados sejam submetidos a algum programa ou instrumento que permita a realização de testes estatísticos com eles.

Nesse contexto, a ferramenta que, atualmente, apresenta mais possibilidades para a realização de diferentes análises é a plataforma *R* (R Core Team, 2023), que possui diversas vantagens em relação a programas comumente utilizados por sociolinguistas, tais como o *Goldvarb-X* (Sankoff *et al.*, 2005). Em primeiro lugar, o *R*, por ser uma linguagem de programação, permite que os usuários realizem inúmeras tarefas, desde a elaboração de *sites* e a conversão de arquivos a análises estatísticas, sem se restringirem a funções já definidas de um programa, sendo, inclusive, utilizado por indivíduos das mais diversas áreas do conhecimento. Por esse mesmo motivo, é possível que cada usuário crie suas próprias funções e/ou *scripts* – os arquivos com os códigos de programação do *R* – e os disponibilize de maneira gratuita na internet, podendo ser utilizados por todos, de modo que as possibilidades de uso da plataforma são, virtualmente, infindáveis.

Pensando em aplicações (sócio)linguísticas, o *R* pode ser empregado não só para a realização de análises multivariadas fixas de regressão logística, como faz o *Goldvarb-X*, mas, também, para análises univariadas, como o teste de qui-quadrado e o teste exato de Fisher, por exemplo, e análises multivariadas de regressão logística ou linear, de maneira binomial ou não. Desse modo, a plataforma permite a investigação de fenômenos cujas variáveis não são binárias com maior precisão, possibilita o trabalho com variáveis numéricas contínuas – como o número de CPS neste estudo –, possui maneiras de aferir o efeito de variáveis aleatórias no estudo, entre tantas outras funções.

Em virtude disso, escolheu-se trabalhar com o *R* nesta pesquisa, utilizando-o através da interface *RStudio*, que conta com janelas que facilitam tanto a visualização dos dados e dos gráficos gerados como a própria elaboração dos códigos a serem empregados, por exemplo. Através dessa linguagem, os dados foram, então, carregados a partir da planilha no formato *.csv*, visualizados e inicialmente analisados através de frequências e proporções. Recorreu-se,

em seguida, a diferentes testes estatísticos para executar as análises uni e multivariadas. Tais testes serão brevemente apresentados a seguir.

O primeiro deles, aplicado nas análises univariadas, é o teste exato de Fisher. Este teste é utilizado para verificar a existência de correlação entre duas variáveis nominais, sendo ele adequado quando o objetivo é comparar proporções entre grupos, especialmente quando um dos níveis da variável apresenta menos do que cinco ocorrências. Nele, assim como em outros testes estatísticos, comparam-se valores observados com valores esperados em uma distribuição, sendo seu resultado dado em um valor: o valor- p , também chamado de significância. Convencionalmente, estabelece-se que um valor- p abaixo de 0.05 é significativo, o que indica, no contexto das análises aqui empreendidas, que existe uma diferença estatística entre os grupos. Por exemplo, na subseção 5.2.1.3, apresentou-se o resultado de um teste exato de Fisher que buscava investigar se a realização do OD anafórico acontecia de maneira diferente em legendas profissionais e *fansubs*; nele, o valor- p encontrado foi menor do que 0.05 ($p = 0.02493$), de modo que se chegou à conclusão de que, estatisticamente, havia diferença entre os dois conjuntos.

Já para as análises multivariadas, cujo objetivo geral é verificar o efeito concomitante de diversas variáveis, o teste aplicado foi uma regressão logística multinomial⁶⁰, devido ao número de variantes sendo consideradas. Neste ponto, é importante traçar um comentário acerca do nome dado aos diferentes tipos de variáveis que fazem parte da análise. Contrariamente ao que ocorre em testes univariados, os termos variável dependente (VD) e variável independente (VI) não são adequados para fazer referência à variável relacionada ao fenômeno linguístico e à variável a ela correlacionada em uma análise multivariada, utilizando-se, em seu lugar, variável resposta (VR) e variável previsora (VP). Sobre essa diferença, Oushiro (2022) explica que, quando se realiza uma análise desse tipo, é possível que exista interação entre mais de uma variável, de modo que seu efeito não seria, de fato, independente – por isso a necessidade de utilizar termos diferentes.

Voltando aos testes empregados nas análises multivariadas dos dados, é necessário, ainda, traçar alguns comentários. Uma regressão logística multinomial, assim como descrito por Levshina (2015), é um modelo de análise adequado quando se tem uma VR com três ou mais níveis, como é o caso do fenômeno deste estudo, já que a realização do OD anafórico de terceira pessoa conta com quatro variantes. Nela, o resultado é dado em *logodds*, uma medida

⁶⁰ Regressões logísticas multinomiais também foram utilizadas para observar o efeito de duas variáveis relacionadas ao gênero – os números de CPL e CPS –, uma vez que essa é uma maneira de testar a correlação entre uma variável resposta enéria e uma variável previsora numérica.

centrada em zero, de modo que valores negativos indicam o desfavorecimento de um dos fatores da VR em relação aos níveis de referência definidos pelo pesquisador e valores positivos, favorecimento, sendo apresentada, também, a significância de cada VP. Diferentemente do que ocorre na regressão logística binomial, esse modelo de análise contrasta, ao mesmo tempo, o nível de referência da VR com todas as outras variantes, apresentando-se estimativas ajustadas para cada uma delas, inclusive em relação a todas as VPs incluídas no modelo. Assim, por meio de regressões logísticas multinomiais, é possível determinar não só quais são as VPs que mais influenciam a ocorrência de uma das variantes da VR, mas também medir o tamanho dessa força.

Ao longo da seção 6, que apresenta a discussão e a análise dos resultados obtidos durante a investigação proposta neste estudo, serão identificados, também, todos os testes realizados. Serão retomadas e aprofundadas, também, algumas explicações aqui traçadas, utilizando exemplos práticos para facilitar a compreensão dos resultados.

5.5 Variáveis e hipóteses

Conforme já foi mencionado anteriormente (cf. seção 2), todo estudo que se encaixa na Sociolinguística Variacionista trabalha com variáveis, com vistas a investigar quais são os fatores que condicionam a escolha por uma ou outra variante de um fenômeno linguístico. Nesse contexto dois são os tipos de variáveis consideradas: as variáveis dependentes ou resposta e as independentes ou previsoras, adotando-se cada termo a depender do tipo de teste estatístico empreendido (cf. subseção 5.4.3).

Com base na literatura acerca do fenômeno aqui investigado, definiram-se, então, quatro variantes, sendo elas o clítico acusativo, o objeto nulo, o SN anafórico e o pronome lexical⁶¹ (Omena, 1978; Duarte, 1986; Cyrino, 1994; Freire, 2000, 2005; Arruda, 2006, 2012; Costa, S., 2011; Costa, T., 2012; Lauer, 2015; Pivetta, 2015; Rodrigues, 2018; Spinelli, 2018; Marques de Sousa, 2021; Lima, 2022). Os exemplos (36) a (39) abaixo, retirados dos episódios que foram adicionados mais tardiamente ao *corpus*, representam cada uma dessas estratégias.

(36) C2: Pergunte logo.

M1: Cadê o Shane?

⁶¹ Neste estudo, optou-se por trabalhar com essas quatro estratégias, pois, além de serem elas as variantes consideradas em pesquisas clássicas acerca do fenômeno sendo analisado, a incorporação do SN anafórico nesta investigação (incluindo os casos com pronome demonstrativo “isso”) permitiria que fosse averiguado um possível desfavorecimento dessa forma nas legendas em decorrência das restrições de espaço e tempo características desse gênero, que, por hipótese, levariam a um uso mais expressivo de estratégias curtas.

Depois que nos forçou
a aceitá-*lo* de volta... (FT10E13)

(37) AZ1: Ainda assim, gostaria de me certificar.

Dr. Karev, como quer prosseguir?

A1: -Com uma TC com triplo contraste.

AZ1: -Então, faça [Ø] (PT06E13)

(38) O1: Bailey, 5 minutos se passaram,

o intestino não está rosa

e não escuto nenhum fluxo.

Precisamos pensar

sobre ressecar *o intestino*. (FT14E13)

(39) PC12: - Eles estão aqui?

A1: - Não.

PC12: Tinha medo

que se eu contasse a você

quanto era importante

você fazer *eles* virem,

você poderia ligar para eles

e dizer para esquecerem isso. (FT10E13)

Já no que diz respeito às variáveis independentes/previsoras, foram definidas tanto variáveis linguísticas – em especial, com base em Lima (2022), por ser o estudo mais recente acerca do objeto de estudo desta pesquisa –, como variáveis extralinguísticas, relacionadas ao contexto de comunicação retratado nos episódios de *Grey's Anatomy* e ao gênero textual-discursivo em questão, levantadas a partir da literatura acerca da TAV e do processo de legendagem e da descrição da legenda empreendida.

Assim, e considerando que “[...] cada variável previsora é uma hipótese a respeito da variável resposta, e as hipóteses [...] devem ser bem motivadas [...]” (Oushiro, 2022, p. 290),

todas as variáveis independentes/previsoras serão elencadas na subseção a seguir, assim como as hipóteses a elas relacionadas, e exemplos retirados do *corpus*.

5.5.1 Variáveis independentes e/ou previsoras

5.5.1.1 Variáveis linguísticas

Ao longo dos anos, os estudos voltados à descrição do OD anafórico no PB analisaram diversas variáveis linguísticas, com algumas delas se mostrando consistentemente relevantes tanto para a identificação dos contextos nos quais, de fato, há variação, como para a realização do fenômeno em si. É o caso da natureza do antecedente, do traço semântico do antecedente, da função sintática do antecedente, da forma verbal do predicador do OD anafórico e da estrutura sintática da oração na qual se encontra o OD, que serão apresentados nas subseções a seguir.

5.5.1.1.1 Natureza do antecedente

Conforme apontado na seção 3, que explora a caracterização do objeto de estudo desta pesquisa, um OD anafórico de terceira pessoa pode não só retomar um SN como uma sentença completa ou um predicado. Com efeito, o *corpus* reunido para a realização da investigação aqui proposta conta com ocorrências desses tipos, como exemplificado pelos trechos (40) a (42).

(40) A1: - Vá pegar cefazolina.

J2: - Não temos [Ø].

A1: - Ache. [Ø]. (PT10E01)

(41) A3: Está apaixonada pelo Alex.

Tudo bem. Não estou julgando.

Eu suspeitava há anos,

mas é um dia ruim

para revelar [Ø]. (FT14E24)

(42) L2: São nossos chefes

e somos dispensáveis para eles.

Mas não deveríamos ser [Ø]. (PT10E13)

Pesquisas anteriores já demonstraram que, no PB, antecedentes oracionais não são retomados por todas as variantes do OD anafórico de terceira pessoa, sendo o objeto nulo a estratégia de retomada mais utilizada, seguida pelo emprego do SN anafórico com diferentes determinantes (cf. Duarte, 1986; Freire, 2005). Assim, nesse contexto, espera-se encontrar apenas objetos nulos e SNs anafóricos retomando referentes sentenciais, além de se supor, também, que grande parte dos SNs corresponderão ao uso do demonstrativo “isso”, muitas vezes relacionados à presença de um “*that*” na fala original em inglês. Caso seja esse o cenário encontrado, a proposta será utilizar essa variável como um critério de separação de dados, assim tal como feito por Lima (2022), com vistas a analisar apenas as ocorrências que apresentarem SNs como antecedentes, uma vez que é nesse contexto que as quatro formas da VD/VR podem ser encontradas.

5.5.1.1.2 Traço semântico do antecedente

De acordo com trabalhos previamente realizados acerca da expressão do OD anafórico de terceira pessoa, umas das variáveis que parece exercer maior influência sobre a realização desse fenômeno é o traço semântico do antecedente (cf. Omena, 1978; Duarte, 1986; Cyrino, 1994; Freire, 2005; Marques de Sousa, 2021; Lima, 2022). De fato, como demonstram Cyrino, Duarte e Kato (2000), o avanço da categoria vazia afetou, inicialmente, antecedentes de traços menos referenciais, que, de maneira geral, deixaram de ser representados pelo clítico neutro “o”, ao passo que os pronomes são mais comumente realizados quando retomam antecedentes mais referenciais.

Com isso em mente, as pesquisas que se debruçaram sobre a questão do acusativo anafórico sugeriram diversas maneiras de se codificar essa variável, podendo considerar apenas questões relacionadas à animacidade, à referencialidade, ou a ambas. Nesse cenário, Cyrino, Duarte e Kato (2000) propuseram uma escala de referencialidade que leva em conta tanto os aspectos [+/- animado] e [+/- referencial] já mencionados como o traço [+/- específico], o que a levou a ser retomada por diversos estudos posteriores. Um deles foi o de Freire (2005), cuja proposta, que conta com três níveis – [- referencial]; [+ referencial; - animado] e [+ referencial; + animado] –, será utilizada na pesquisa aqui descrita⁶².

⁶² Vale destacar que, em consonância com o mesmo autor, serão considerados [+ referenciais] os antecedentes que fazem referência a uma entidade no mundo biossocial, incluindo, por exemplo, substantivos coletivos.

Assim, para essa variável, espera-se que antecedentes de traço [+ referencial; - animado] favoreçam a utilização do objeto nulo, enquanto aqueles de característica [+ referencial; + animado] se correlacionem com a realização plena dos pronomes. Além disso, supõe-se que os dados com antecedente [- referencial], por representarem objetos sentenciais, se apresentarão apenas com objetos nulos ou SNs anafóricos. Os três níveis da variável em questão são ilustrados pelos exemplos (43) a (45) abaixo, na mesma ordem em que foram apresentados no parágrafo anterior:

(43) O1: Temos pouquíssimos remédios.

O pessoal não veio
por causa dos alagamentos.

C2: Então precisamos
fechar o PS.

AZ1:- Mas podemos fazer *isso*? (FT10E01)

(44) O1: Está bem. Apenas relaxe.

Vamos tomar conta de vocês dois, está bem?

Q1: -Devo tampar esses ferimentos?

B2: -Isso pode esperar.

Apenas cubra [Ø] com gaze em solução salina. (PT14E13)

(45) CP1: Pode contar a ela
mesmo assim?

Pode encontrá-*la*
e dizer que *a* amei?

Eu *a* amava tanto. (FT06E24)

Vale mencionar, ainda, que, caso sejam excluídos os dados cujo antecedente é de natureza sentencial/predicativa, como sugerido na subseção anterior, serão mantidos na análise apenas os níveis [+ referencial; - animado] e [+ referencial; - animado], sendo eles simplificados como [+ animado] e [- animado].

5.5.1.1.3 Função sintática do antecedente

Em estudos anteriores, outra variável que se mostrou relevante para a realização do OD anafórico foi a função sintática do antecedente. Inicialmente, Omena (1978) demonstrou que antecedentes que exerciam função de complemento, ou seja, que apresentavam a mesma função sintática que os ODs, condicionam a utilização da categoria vazia, achado que foi posteriormente confirmado, por exemplo, por Marafoni (2004) e Santana (2016), em estudos sobre a língua oral. Por outro lado, Lima (2022), estudando textos jornalísticos, atestou que, quando a função sintática do antecedente também é de OD, as porcentagens para o objeto nulo eram maiores, apesar de não ter sido essa a variante mais utilizada nesse contexto.

A partir desses resultados, hipotetiza-se que, no *corpus* elaborado para esta pesquisa, os objetos nulos sejam favorecidos quando o antecedente cumprir a mesma função sintática que o OD anafórico, mesmo que ela se dê através do meio escrito, principalmente por a legenda ter como um objetivo se aproximar, em certa medida, da fala original. Para trabalhar com essa variável, com base em Marafoni (2004) e Santana (2016), foram definidos, então, dois níveis: “igual” e “diferente”, representados pelos exemplos (46) e (47), respectivamente.

(46) A2: Marilyn,

vou te dar um sedativo.

Respire, certo?

Como o Leo se parece?

APC14: Tem 10 anos, cabelo castanho
e grande, não quer cortar [Ø]. (FT10E24)

(47) O1: Talvez porque ela já passou

por dor demais na vida?

N1: Exato! Não tem porquê fazê-*la*
passar por isso. (PT14E01)

Durante a codificação, porém, percebeu-se que certos dados não se encaixariam em nenhum dos níveis estabelecidos, devido à natureza do gênero textual-discursivo em questão. Trata-se de ocorrências que se encaixam em três casos: (i) quando o antecedente é identificado visualmente, não sendo verbalizado e não cumprindo, portanto, uma função sintática; (ii) quando o antecedente foi mencionado em uma cena anterior e/ou fora de cena, mas pode ser

presumido pelo contexto; (iii) quando o antecedente é verbalizado, mas, devido a elipses, não é possível identificar sua função sintática. Os trechos (48) a (50) abaixo exemplificam cada uma dessas possibilidades na ordem em que foram apresentadas.

(48) S3: Desculpe.

J2: Por que você não tira [Ø]? (FT14E01)

(49) A2: Não consigo achá-*los*.

E o pastor não chegou. (PT14E24)

(50) B1: - Ácido láctico.

A1: - 2.

B1: Isso não pode estar certo.

Ele tinha que estar acidótico.

- A cintura está aumentando.

A1: - Aqui diz...

B1: Eu sei o que diz,

mas está errado.

Refaça [Ø]. (FT10E01)

No exemplo (48), os óculos do personagem Levi Schmitt, identificado como S3, caem enquanto ele troca beijos com Jo Wilson (J2), o que a leva a sugerir que ele tire o acessório antes que eles continuem. O antecedente é, portanto, um objeto apresentado ao espectador de maneira visual, de modo que a anáfora se concretiza de maneira multimodal. Já no dado apresentado em (49), April Kepner (A2) inicia a cena afirmando que não consegue achar Alex Karev e Jo Wilson, que são mencionados somente em cenas anteriores. No enredo do episódio, os dois desaparecem no dia de seu casamento, fazendo com que outros personagens procurem por eles ao longo do capítulo, de modo que, quando April utiliza o clítico acusativo no exemplo, fica claro a quem ela está se referindo. Por fim, no trecho (50), o antecedente consiste no exame de ácido láctico mencionado por Miranda Bailey (M1). Entretanto, a personagem diz apenas essas duas palavras – “ácido láctico” – para pedir o resultado, omitindo o restante da frase. Apesar de ser facilmente entendida, não se pode, porém, ter certeza se a sentença completa seria

algo como “Qual o resultado do ácido lático?” ou como “O ácido lático deu quanto?”, sendo impossível determinar a função sintática do antecedente com precisão.

Assim, por não se encaixarem nas categorias adotadas para a variável em questão, ODs anafóricos como os apresentados acima foram codificados como “NA”, por entender-se que eles não cumpriam uma função sintática de fato. Para esse nível, a expectativa é que sejam empregados mais objetos nulos, uma vez que os antecedentes em questão podem ser facilmente recuperados, podendo ser até mesmo completamente omitidos do diálogo.

5.5.1.1.4 Forma verbal

Ainda outra variável que se mostrou relevante em diversas pesquisas foi a forma verbal do verbo que precede o OD anafórico (cf. Duarte, 1986; Freire, 2005; Santana, 2016; Lima, 2022). Sobre essa variável, Freire (2005) afirma que a ocorrência do clítico só consegue ser mais expressiva do que a soma das demais variantes quando encontrado após formas verbais simples não finitas, tendo destaque, nesse contexto, o infinitivo. A mesma tendência foi encontrada por Lima (2022), sendo o SN anafórico e o objeto nulo as estratégias preferidas em seu conjunto de dados apenas quando empregados com formas flexionadas.

A partir dos trabalhos mencionados, espera-se encontrar, portanto, um número maior de clíticos utilizados com formas não finitas, especialmente o infinitivo. Para investigar essa hipótese, foram inicialmente adotados três níveis, com base em Freire (2005): formas simples não flexionadas, formas simples flexionadas e formas complexas. Entretanto, conforme será relatado na próxima seção, após o início das análises, percebeu-se que o clítico superava as outras variantes de maneira significativa principalmente com formas no infinitivo. A partir disso, optou-se por reorganizar a variável em três outros níveis: formas de infinitivo (simples ou complexas), formas simples flexionadas e outras formas não finitas (simples ou complexas). Os excertos (51) a (53) ilustram cada um desses níveis, na ordem em que foram elencados.

(51) R1: Aquilo foi cruel.

Ele estava se debatendo,
e você estava gostando.

O1: Me desculpe
se meu estilo de liderança

não envolva chamá-*lo*

para almoçar. (FT10E13)

(52) AZ1: Como April devolverá o presente

se nem deixamos os presentes
com cartões?

C1: Certo, acharemos *os cartões*.

Arizona! Pare. (PT10E13)

(53) C2: Owen está sendo muito

homem das cavernas.

E isso é excitante. Meio
perturbador, mas bem excitante.

Acho que ainda está chateado
por eu tê-*lo* oferecido a Teddy. (FT06E13)

5.5.1.1.5 Estrutura sintática

Por fim, a última variável linguística considerada por este estudo será a estrutura sintática projetada pelo verbo que antecede o OD. Ao longo dos anos, tal variável foi caracterizada de diferentes maneiras, considerando-se níveis diferentes a depender da pesquisa em questão. Duarte (1986) e Freire (2005), por exemplo, consideram oito níveis em suas investigações, ao passo que Santana (2016) trabalha com seis, e Lima (2022), com apenas três níveis. De modo geral, porém, são observados os seguintes tipos de estrutura: (i) estruturas simples, que seguem o padrão S V OD; (ii) estruturas com complementos, como S V OD OI e S V OD Oblíquo; e (iii) estruturas que incluem miniorações, que se configuram como S V OD + predicativo, S V OD + infinitivo e S V OD + gerúndio.

Como resultados, Duarte (1986) e Marafoni (2004) encontraram, em suas pesquisas acerca da língua oral, que estruturas com miniorações favorecem mais o uso do pronome lexical do que outras, uma vez que o OD exerce a dupla função de objeto e sujeito ao mesmo tempo. Por outro lado, estudos acerca da escrita (cf. Freire, 2005; Lima, 2022) demonstram haver uma presença considerável do clítico em todos os contextos nos quais o antecedente é um SN, concorrendo, por vezes, com as outras variantes. Assim, a hipótese para essa variável supõe que, nas legendas, o mesmo cenário será atestado, com pronomes lexicais sendo mais comuns

em estruturas que projetam miniorações e o clítico figurando como uma variante expressiva em diversos contextos.

Nesta investigação, foram consideradas, a princípio, cinco das seis estruturas acima mencionadas, já que criar amálgamas posteriormente é mais prático do que ter que recodificar os dados caso seja necessário desfazer a junção de um ou mais níveis inicialmente adotados⁶³. Assim como no caso da forma verbal, os resultados das primeiras análises indicaram uma preferência pelo clítico mais expressiva em certos contextos, o que levou à criação de mais um amálgama. Passou-se a trabalhar, então, com três níveis nessa variável, quais sejam: S V OD, S V O + complementos e S V OD + minioração. Os exemplos (54) a (56), organizados nos tópicos abaixo, ilustram os três níveis dessa variável.

- **S V OD**

(54) S1: Você viu o dr. Webber?

C2: Eu *o* vi antes da cirurgia. (FT10E01)

- **S V OD + complementos**

(55) C1: Sai. Pior não pode ficar.

Já está péssimo.

Preciso levá-*lo* lá para cima

e pôr um fixador externo já. (PT10E01)

- **S V OD + minioração**

(56) C2: - Owen.

O1: - Estou aqui.

- Calma, converse comigo.

GC1: - Pare de operá-lo.

Deixe *ele* morrer. (FT06E24)

⁶³ Excluiu-se o nível S V OD + gerúndio das análises empreendidas nesta dissertação por não ter sido encontrado nenhum dado com essa estrutura.

5.5.1.2 Variáveis extralinguísticas

Em pesquisas sociolinguísticas, um passo comum e esperado é a análise dos condicionamentos relacionados a variáveis extralinguísticas tais como idade, escolaridade, sexo/gênero, entre outras. Entretanto, por este estudo se propor a investigar a realização de um fenômeno em um material como a legenda, que não se desvincula do conteúdo audiovisual original, e por se reconhecer a importância de se considerar os gêneros textuais-discursivos de maneira detalhada ao investigar processos de variação e mudança (cf. subseção 2.3), optou-se por trabalhar com variáveis voltadas às características estruturais e situacionais da legenda que, por hipótese, poderiam apresentar um efeito mais relevante na escolha por uma ou outra variante do acusativo anafórico de terceira pessoa do que se analisadas as características sociais das personagens e/ou dos tradutores⁶⁴.

Desse modo, decidiu-se trabalhar com cinco variáveis extralinguísticas divididas em dois blocos diferentes no presente trabalho: aquelas voltadas a aspectos estruturais e as correspondentes a aspectos situacionais. No primeiro grupo, incluiu-se o tópico, relacionado ao tema geral das conversas retratadas na série, a relação entre os participantes, que volta o olhar sobre os diferentes vínculos existentes entre as personagens de *Grey's Anatomy*, e o tipo de legenda. Já no segundo, foram considerados o número de CPL e o de CPS (cf. seção 4). Nas subseções a seguir, essas variáveis serão mais aprofundadas e exemplificadas com exemplos do *corpus*.

5.5.1.2.1 Tópico

Na seção 2, foi apresentada a proposta de Biber e Conrad (2009) para a categorização dos aspectos situacionais de diversos gêneros textuais-discursivos, sendo um deles o Tópico. Conforme discutido anteriormente (cf. subseção 2.2), uma das dimensões que podem ser exploradas é a do tópico geral, isto é, a natureza dos assuntos mobilizados em uma dada situação comunicativa. Com isso em mente, e levando em consideração que a legenda se faz em conjunto com a produção audiovisual à qual ela corresponde, julgou-se ser importante investigar, nesta pesquisa, se o tópico das conversas representadas no seriado exerce influência sobre a realização do OD anafórico de terceira pessoa.

⁶⁴ Reconhece-se, aqui, a importância de um estudo que possa observar como é feita a tradução de personagens que representam diferentes categorias sociais, buscando averiguar se os tradutores não só percebem essas diferentes categorias, mas buscam refleti-las na linguagem utilizada. Contudo, por uma questão de recorte, esta pesquisa não tratará dessas questões, focando-se de maneira mais aprofundada em questões relacionadas ao gênero textual-discursivo em questão.

Posto que *Grey's Anatomy* se caracteriza como uma série que explora tanto o dia a dia dos médicos cirurgiões no hospital como os dramas pessoais de seus protagonistas, decidiu-se trabalhar com dois níveis na variável “tópico”: profissional e pessoal, que abarcam, de maneira geral, todas as interações entre as personagens. Desse modo, sob o rótulo “profissional” estarão as conversas que versam sobre cirurgias, pacientes, procedimentos, questões burocráticas, entre outros, e, no âmbito do termo “pessoal”, se encaixarão os diálogos acerca de relacionamentos familiares e amorosos, problemas de convivência, emoções das personagens etc.

Por hipótese, tem-se que as conversas profissionais favoreçam ou o uso do clítico, por ser uma forma comumente associada à formalidade, ou do objeto nulo, já que essa é uma estratégia sobre a qual não parece recair estigma. Além disso, espera-se que os pronomes lexicais apareçam em maior número em interações pessoais, já que a legenda deveria, em teoria, retratar uma linguagem mais próxima da que é, de fato, empregada oralmente. Os trechos (57) e (58) abaixo representam os dois níveis dessa variável.

(57) O1: Ele é doador?

R1: Não sei [Ø]. Liguei para a mãe dele.

Ela está vindo.

Vamos suturar [Ø] e transferi-*lo*

para a UTI, caso ela queira vê-*lo*. (PT06E01)

(58) A1: - Ela chamou Koracick.

M4: - Meu deus. Péssimo.

A1: - Um babaca.

Veio para transar com ela.

A3: - Não veio, não.

A1: - Veio para transar.

Você transou com ele

da última vez.

A3: - Não deveria ter contado *isso*. (FT14E13)

5.5.1.2.2 Relação entre participantes

Não é incomum que estudos sociolinguísticos, sobretudo aqueles que se debruçam sobre a variação dos pronomes de segunda pessoa ou das formas de tratamento, incluam questões voltadas à relação entre os falantes, pautando-se, muitas vezes, na Teoria da Polidez (Brown; Lewinson, 1987) e na Teoria do Poder e da Solidariedade (Brown; Gilman, 1960). Nesse contexto, os autores discorrem sobre relações que podem ser mais simétricas, quando os falantes estão mais próximos na hierarquia social, ou mais assimétricas, quando suas posições são distantes, sendo que ambas podem exercer um efeito sobre as escolhas linguísticas dos envolvidos.

De maneira semelhante, apesar de partirem de outra orientação teórica, Biber e Conrad (2009) também elencam como um aspecto essencial para a caracterização de um gênero textual-discursivo a relação entre os participantes, sugerindo que sejam observados pontos como interatividade, conhecimento compartilhado, papéis sociais e relações pessoais. Em sua obra, os autores associam os papéis sociais ao *status* relativo que um indivíduo detém diante da sociedade e/ou ao poder que ele possui, e as relações pessoais ao tipo de vínculo estabelecido entre os participantes: se são amigos, colegas, estranhos, entre outros.

Traçando um paralelo entre as duas abordagens mencionadas, esta pesquisa se propõe a observar se, nas legendas dos diálogos de *Grey's Anatomy*, as relações entre os participantes das cenas – isto é, entre as personagens – influenciam a realização do OD anafórico de terceira pessoa. Para investigar essa variável, definem-se dois níveis: simétrico, representando relações entre pares como médico-médico e paciente-paciente; e assimétrico, abarcando as conversas entre médico e paciente, médico e superior, médico e enfermeiro, entre outros. Caso a relevância desse fator se confirme, supõe-se que serão encontrados mais clíticos quando a relação entre os participantes for assimétrica e variantes mais inovadoras, como a categoria vazia e o pronome lexical, quando ela for simétrica. O trecho (59) a seguir representa um diálogo que envolve uma relação simétrica, entre médicos amigos, e o (60) uma relação assimétrica, entre médico e paciente.

(59) O1: - Você disse sim?

A2: - Como?

O1: Para o Matthew.

Eu ajudei [Ø] com o pedido.

A2: Sim, eu disse sim. (FT10E01)

(60) AZ1: Houve um acidente, e um dos médicos da nossa equipe morreu.

APC1: Andy é a minha equipe.

Ele é o único no meu time

e algo está errado com ele.

Então, estou perguntando para você:

devo levá-*lo* para outro lugar? (PT06E01)

5.5.1.2.3 Tipo de legenda

O primeiro fator controlado relacionado ao gênero foi o tipo de legenda, tendo essa variável feito parte das investigações desde o trabalho inicial com a amostra-piloto. Conforme exposto na seção 4, a legenda audiovisual pode ser produzida tanto por tradutores profissionais, inseridos em empresas de tradução ou estúdios, como por tradutores-fãs, que criam legendas para suas séries favoritas, disponibilizando-as de maneira gratuita na internet.

Dados seus contextos de produção e de recepção distintos, uma das principais hipóteses elaboradas para esta pesquisa é a de que a realização do OD anafórico de terceira pessoa ocorreria de maneira distinta em legendas profissionais e em *fansubs*. Nesse panorama, supunha-se serem dois os cenários possíveis: (i) as legendas profissionais fariam o uso de variantes que estão de acordo com a norma-padrão, por serem produzidas por profissionais da língua, enquanto as *fansubs* apresentariam usos mais próximos da fala natural, por serem o resultado de um trabalho amador; ou (ii) devido à necessidade de validação de um trabalho essencialmente informal, as legendas feitas por fãs se aproximariam mais do padrão linguístico ditado pelas gramáticas, enquanto as legendas institucionalizadas demonstrariam mais liberdade na escolha das variantes utilizadas, por vezes apontando para um conhecimento maior dos tradutores acerca de questões de adequação linguística.

Após as duas rodadas de testes com a amostra-piloto, que tiveram resultados divergentes (cf. subseções 5.2.1.3 e 5.2.2.2), julgou-se necessário manter essa variável no estudo a fim de confirmar se ela se correlaciona ou não com o fenômeno em questão. Havendo correlação, as expectativas são aquelas descritas no parágrafo acima; por outro lado, caso a diferença entre a realização do acusativo anafórico de terceira pessoa nos dois tipos de legenda não seja

significativa, tal resultado pode indicar que o gênero legenda se apresenta de maneira mais homogênea do que inicialmente foi pensado, especialmente se o mesmo for futuramente verificado ao se analisar outras variáveis, ou seja, se outros fenômenos também não forem condicionados pelo tipo de legenda.

Como já foi mencionado, essa variável conta com dois níveis, profissional e *fansub*, que não serão aqui abertamente exemplificados, posto que todos os trechos trazidos nesta dissertação foram reproduzidos já com a indicação da amostra de que vieram, isto é, a codificação com relação ao tipo de legenda.

5.5.1.2.4 Número de CPL

Uma das maiores restrições enfrentadas pelos tradutores audiovisuais é, sem dúvida, o número máximo de CPL que pode ser utilizado. Tal como discutido na subseção 4.2, todo tipo de legenda audiovisual possui um limite para esse parâmetro, o que acaba, por vezes, forçando o tradutor a alterar o texto inicialmente traduzido, adotando estratégias que possam encurtá-lo.

Assim, considerando que as quatro variantes do OD anafórico possuem extensões diferentes, é possível que o número de CPL afete diretamente a escolha final dos tradutores. Supõe-se que formas mais curtas como o objeto nulo e o clítico sejam privilegiadas quando a frase se aproximar do limite estabelecido, e as outras variantes figurem mais comumente em trechos cujo número de caracteres é mais baixo.

Para realizar o controle dessa variável, foram mantidos os valores específicos de CPL de cada legenda, obtidos ao abrir os arquivos originais no programa de legendagem *Aegisubs* (cf. subseção 5.4.2), já que, assim como para outras variáveis, isso permitiria que, posteriormente, amálgamas fossem criados se necessário. Isso significa que essa variável será tratada como numérica a princípio, de modo que será necessário um tipo de teste diferente dos demais para observar o seu efeito. O exemplo (61) a seguir representa uma ocorrência do OD anafórico em uma *fansub*, em frase próxima ao limite de caracteres (26), enquanto o exemplo (62) ilustra um contexto no qual se está distante do número máximo de CPL (11).

(61) HE1: Meu pai é um chefe de bombeiros,
me ensinou [Ø] quando eu tinha sete. (FT14E13)

(62) C1: -Tudo bem.

Pegue. Leve-*os*. (PT06E24)

5.5.1.2.5 Número de CPS

Conforme mencionado na subseção 4.2, tem-se costumeiramente um máximo de 21 CPS em legendas profissionais, enquanto, para as *fansubs*, o limite normal é de um caractere a menos, ou seja, 20. Assim, pelos mesmos motivos apresentados na subseção anterior, o número de (CPS) também acaba diminuindo as possibilidades de escolha do tradutor, de modo que se torna importante lançar um olhar cuidadoso para esse parâmetro técnico.

Ao operacionalizar essa variável, optou-se, como no caso dos CPL, manter os números indicados pelo programa *Aegisub* após carregar as legendas, e tratar dessa variável como contínua, decisão que se justifica pela possibilidade de, no momento das análises, ser possível criar amálgamas de maneira facilitada. Novamente ecoando o que foi dito na subseção anterior, espera-se que as formas curtas do OD anafórico sejam mais empregadas em contextos próximos do limite de CPS e que, em frases nas quais o número de CPS é mais baixo, apareçam, também, as outras variantes do fenômeno em pauta. Seguem-se os exemplos (63) e (64), com 19 e 9 CPS, respectivamente, que ilustram os dois contextos aqui apontados.

(63) L2: Ross e Karev

o levaram para cirurgia. (FT10E24)

(64) B1: “Faremos o seguinte?”

Nós? Não a ouviu? (PT06E13)

5.6 A descrição da legenda como um gênero

Um dos objetivos do presente estudo é, conforme descrito na Introdução, a descrição da legenda como um gênero textual-discursivo, tanto a partir de um viés estrutural como sob um olhar situacional.

Com vistas a caracterizar a legenda a partir de sua forma, serão consideradas, então, as discussões empreendidas na seção 4 acerca da TAV e das especificidades das legendas profissionais e das *fansubs*. A partir dos arquivos reunidos para a realização deste trabalho, será observado, portanto, se os traços apontados como definidores de uma legenda estão, de fato, presentes no material que compõe o *corpus*, além de investigar-se, também, se as recomendações dos guias de legendagem quanto aos parâmetros técnicos são ou não são geralmente respeitadas e as consequências dessa padronização.

Já no que diz respeito à descrição situacional da legenda, serão examinados os sete aspectos para a caracterização e comparação de gêneros textuais-discursivos propostos por

Biber e Conrad (2009), a saber: participantes, relações entre participantes, canal, condições de produção, cenário, propósitos comunicativos e tópico (cf. seção 2). Desse modo, partindo, mais uma vez, do *corpus*, será explorado como cada um desses pontos se materializa nas legendas analisadas, observando se existem grandes diferenças entre as legendas profissionais e as *fansubs* e recorrendo a adaptações, caso seja necessário.

No contexto desta pesquisa, considera-se que a caracterização completa desse gênero textual-discursivo é, também, um resultado, posto que era ela um dos objetivos definidos para o trabalho. Em razão disso, a descrição da legenda elaborada a partir dos procedimentos aqui mencionados será apresentada na próxima seção, juntamente aos resultados referentes à análise da variação linguística do OD anafórico de terceira pessoa.

6 ANÁLISE DE DADOS E DISCUSSÃO DOS RESULTADOS

Nesta seção, apresentam-se tanto os resultados obtidos a partir de análises estatísticas dos dados da realização do OD anafórico de terceira pessoa, provenientes de legendas profissionais e de *fansubs*, como a caracterização da legenda como um gênero textual-discursivo, empreendidas concomitantemente no decorrer da pesquisa. Inicia-se com considerações sobre a distribuição geral dos dados, pela apresentação da análise multivariada e por comentários específicos sobre aquelas variáveis que se mostraram significativamente correlacionadas, e, por fim, traz-se a descrição da legenda a partir do conceito de gênero.

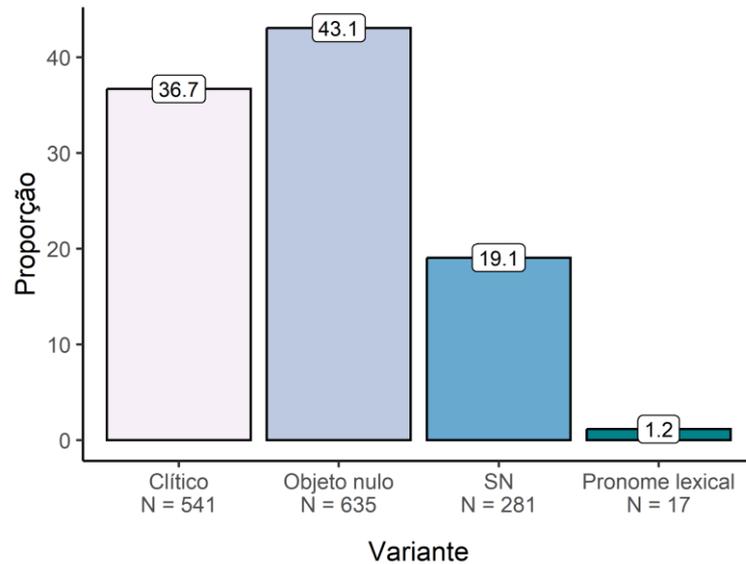
6.1 O OD anafórico de terceira pessoa em legendas audiovisuais

6.1.1 Distribuição geral

Antes de dar início à exposição dos resultados, é importante traçar um comentário acerca da maneira como eles estão organizados. Conforme será discutido na subseção 6.1.2 adiante, os testes estatísticos realizados indicaram não haver diferença significativa entre os dois tipos de legenda no que diz respeito à realização do OD anafórico de terceira pessoa, o que indica que, ao menos para a expressão desse fenômeno, os subconjuntos considerados são estatisticamente iguais. Por isso, os resultados para as duas amostras serão apresentados de maneira conjunta, sem fazer distinção entre legendas profissionais e *fansubs*, com exceção, é claro, das considerações especificamente relacionadas à variável tipo de legenda.

No total, foram encontradas 1474 ocorrências de OD anafórico de terceira pessoa nas legendas que compõem o *corpus*, retomando tanto antecedentes oracionais como sintagmas nominais e predicativos, sendo 754 delas derivadas de legendas profissionais e 720 de legendas feitas por fãs, o que indica equilíbrio entre as amostras. Nesse cenário, a distribuição geral se deu da seguinte maneira:

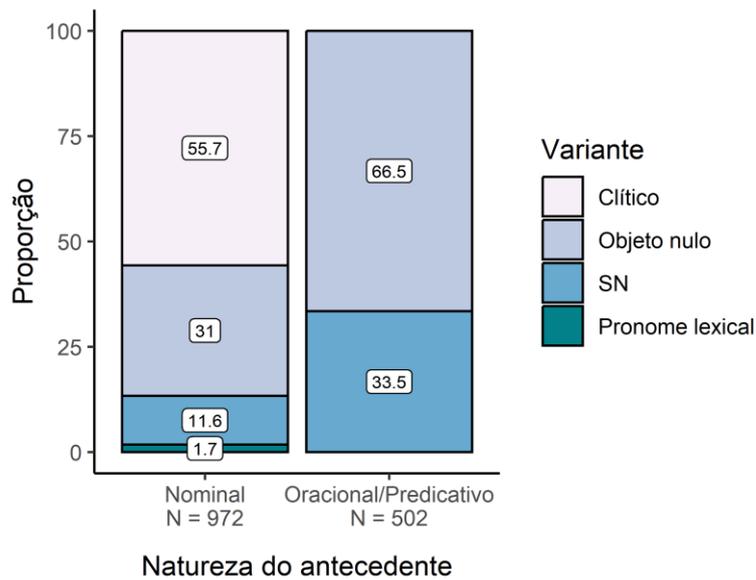
Gráfico 11 - Frequências e proporções das quatro variantes da realização do OD anafórico de terceira pessoa no *corpus* de legenda completo (N = 1474)



Fonte: Elaborado pela autora

Como pode ser observado, considerando ODs que retomam antecedentes de todas as naturezas, a estratégia mais utilizada pelos tradutores foi o objeto nulo, correspondente a 43,1% dos dados, seguido pelo clítico acusativo (36,7%), pelo SN (19,1%) e pelo pronome lexical (1,2%). Entretanto, ao voltar o olhar para a natureza do antecedente, percebe-se uma forte influência dessa variável nos resultados gerais. Com efeito, dos 1474 dados encontrados, 502 são referentes a antecedentes oracionais e predicativos, tendo apenas duas variantes aparecido nesse contexto: o objeto nulo, com 66,5% das ocorrências, e o SN, com 33,5%, como pode ser observado no Gráfico 12.

Gráfico 12 – Proporções do OD anafórico de terceira pessoa de acordo com a natureza do antecedente (N = 1474)



Fonte: Elaborada pela autora.

Outro ponto que chamou a atenção foi o fato de que, das 168 retomadas executadas através de SNs, 166 foram realizadas por meio de pronomes demonstrativos, uso que pode estar relacionado à utilização dos pronomes “*it*” ou “*that*” no original em inglês⁶⁵. Os exemplos (65) e (66) abaixo ilustram as duas variantes empregadas com antecedentes oracionais/predicativos, e os excertos (65’) e (66’)⁶⁶ correspondem às versões originais.

(65) B1: Você é a estrela
de sua turma.

Eu sempre soube [Ø]. (FT10E01)

(65’) B1: *You are the star of your class.*

*I always knew **it**.* (T10E01)

(66) M4: Você transou com ele
na última vez que ele esteve aqui.

⁶⁵ Reconhece-se, aqui, a pertinência de se analisar essa relação, especialmente por ela poder determinar a escolha por uma ou outra variante para representar antecedentes oracionais/predicativos. Contudo, por uma questão de escopo, essa discussão não será empreendida nesta dissertação.

⁶⁶ Uma vez que os trechos em (65’) e (66’) são reproduções diretas do original, não há indicação de legenda profissional ou *fansub*. Além disso, neles os antecedentes são marcados pelo uso do itálico, enquanto o OD é destacado pelo negrito.

Eu nunca deveria ter dito *isso* a vocês. (PT14E13)

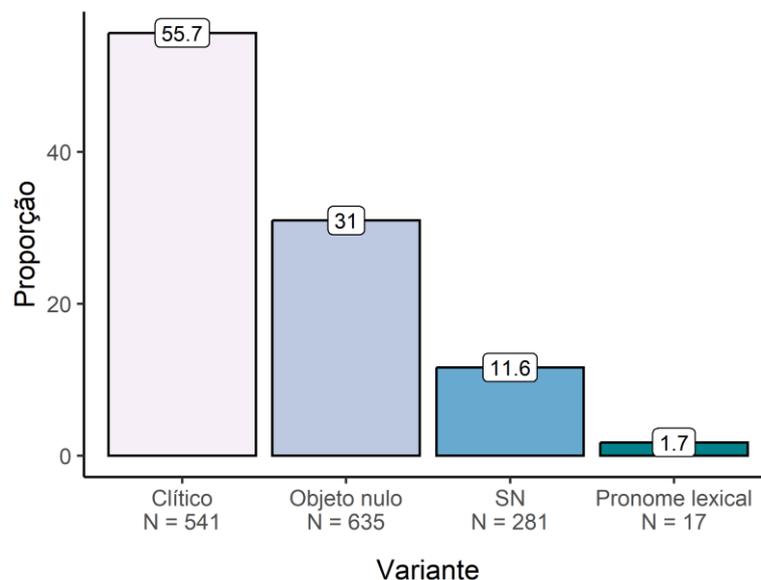
(66) M4: *You did have sex with him*
the last time he was here.

I never should've told you guys that. (T14E13)

Tal resultado está de acordo com o esperado, uma vez que pesquisas anteriores já atestaram que antecedentes oracionais/predicativos favorecem o objeto nulo e o SN, em detrimento das outras duas variantes do fenômeno (cf. Duarte, 1986; Cyrino, 1994; e Freire, 2005, por exemplo). Além disso, não há, nesse caso, opcionalidade para o pronome lexical, uma vez que ele não poderia ser utilizado para retomar um antecedente oracional/predicativo. Com efeito, uma frase como “*Eu nunca deveria ter dito *ele* a vocês” seria considerada agramatical no PB, de modo que, nesse contexto, nunca estiveram disponíveis as quatro variantes do fenômeno, apenas três. Assim, como anunciado na subseção 5.5.1.1.1, foram consideradas para o restante da análise, portanto, apenas as ocorrências de OD anafórico de terceira pessoa que retomassem antecedentes nominais.

Excluindo-se, então, os dados com antecedentes oracionais/predicativos, obteve-se um contingente de 972 ocorrências, das quais 487 foram coletadas de legendas profissionais e 485 de *fansubs*. Nesse conjunto, a distribuição dos dados foi a seguinte:

Gráfico 13 - Frequências e proporções das quatro variantes da realização do OD anafórico de terceira pessoa com antecedentes nominais (N = 972)



Fonte: Elaborado pela autora

Conforme apresentado no Gráfico 13, a estratégia mais expressivamente utilizada para realizar o OD anafórico de terceira pessoa no *corpus* de legenda foi o clítico acusativo, com 55,7% das ocorrências, em contraste com os 31% de objeto nulo, 11,6% de SN, e os inexpressivos 1,7% de pronome lexical. Os excertos (67) a (70) a seguir exemplificam cada uma dessas variantes.

(67) J2:- Ouça-me!

Não estou te largando.

Só não podemos nos casar
porque está abalado pelo seu pai.

A1: Não *o* use como desculpa
por estar hesitante. (PT10E13)

(68) MK1: - Poderemos ligar os braços.

A1PC2: - Eba! Quem achou [*Ø*] fui eu. (FT06E01)

(69) O1: Deve uma ao homem

cujos coração estragou

por não saber
o que estava fazendo,

assim como deve a esta garota
que está na mesa, perdendo sangue,

porque não sabe
como fazer um nó básico.

Saia da sala cirúrgica, Dr. Ross.

Dr. Warren, faça *o nó*. (PT10E13)

(70) Menino, 14 anos, trauma

abdominal e ferimento na coxa.

Sangue na cabeça. Tirei *ele*

e o irmão da casa em chamás. (FT14E13)

No caso dos clíticos, esperava-se encontrar um número menor de ocorrências nesse contexto, uma vez que essa variante vem desaparecendo da modalidade oral do PB (cf. Omena, 1978; e Duarte, 1986; Freire, 2000, por exemplo) e as legendas se propõem a emular a fala original das personagens de maneira natural (Lopes; Afonso; 2021). Uma possível explicação para esse fato pode ser que, apesar de pensada em relação com a oralidade, a legenda não consiga se desprender completamente da modalidade escrita na qual é criada. Desse modo, acaba por privilegiar o uso do clítico, posto que essa é uma estratégia que ainda persiste nesse contexto por força das prescrições gramaticais. Ademais, assim como no caso do objeto nulo, o uso expressivo dos clíticos nesse contexto pode estar ligado às exigências dos aspectos técnicos envolvidos na produção das legendas. Com efeito, os clíticos são, depois da categoria vazia, a forma mais curta de se expressar o OD anafórico de terceira pessoa, o que pode funcionar como alternativa a outras variantes quando se está perto do limite de CPL e CPS.

Contudo, é importante reconhecer que, em muitos dos casos encontrados, não há opcionalidade para todas as variantes. Ao observar a tradução em (69), por exemplo, seria difícil imaginar o emprego de uma das outras formas sem apoio explícito de outros elementos multimodais, principalmente pelo fato de o próprio original realizar a repetição do SN (“*Do the knot*”). Da mesma maneira, no exemplo (70), a única alternativa ao pronome lexical que parece viável seria um SN como “o menino”, dado que um nulo resultaria em uma oração agramatical por não apresentar um OD para o verbo (*Tirei [Ø] e o irmão da casa em chamás), e um uso de um clítico, apesar de gramaticalmente possível, não soaria natural em um OD com dois núcleos (Tirei-o e o irmão da casa em chamás).

Quanto ao uso da categoria vazia nesse contexto, considerando a proposta supracitada de que a legenda se aproxime da fala a ser adaptada na tradução, não é surpreendente que uma estrutura já consolidada no PB falado, e que funciona como uma maneira não estigmatizada de expressar o acusativo de terceira pessoa (Duarte, 1986), seja a escolhida em aproximadamente um terço dos casos. Entretanto, a proporção verificada ajuda a confirmar o que havia sido hipotetizado, posto que, justamente por essa forma ser amplamente difundida na língua oral, esperava-se uma maior produtividade dela em um gênero textual-discursivo que se relaciona diretamente com a fala. Além disso, o emprego do objeto nulo nas legendas também evita que os tradutores utilizem mais caracteres do que o necessário em cada linha, o que contribui para que os parâmetros técnicos sejam respeitados. Assim, é possível que a escolha da categoria vazia como segunda estratégia de realização do OD anafórico preferida em legendas aconteça

não só em decorrência de uma suposta consciência dos tradutores acerca de usos linguísticos comuns, mas, também, devido à necessidade de atender requisitos técnicos.

Por causa dessas restrições técnicas, já era esperado que o número de SNs não fosse tão expressivo na amostra, como ficou comprovado. De fato, a repetição de um SN não mencionado, com ou sem alterações, implica a utilização de mais caracteres, o que pode levar o tradutor a evitar essa alternativa. Ademais, parece haver uma tendência – ainda que não testada nesta pesquisa – de se manter os SNs apenas quando eles são utilizados, também, no original em inglês, o que sugeriria uma influência do decalque linguístico sobre a tradução das legendas do *corpus*.

Da mesma maneira, as expectativas com relação ao pronome lexical também se confirmaram. Apesar de serem amplamente utilizados na fala, os pronomes tônicos em função de OD ainda são comumente estigmatizados pelos próprios falantes, especialmente na modalidade escrita (Duarte, 1986; Freire, 2005; entre outros), de modo que seu uso nas legendas poderia ser fortemente criticado, podendo, na opinião do público, colocar em xeque a competência do tradutor para executar o trabalho que lhe cabe. Entretanto, é relevante mencionar que, apesar disso, existem ocorrências de pronomes lexicais no *corpus*, em ambas as amostras, como demonstrado pelos exemplos (71) e (72), o que, apesar do baixo número de dados, pode indicar que os tradutores têm consciência de que, a depender do contexto, esse é um uso aceitável.

(71) AZ1: O que a Zola achou dele?

M1: Ela quer levar *ele* para mostrar
para todo mundo, na creche. (PT10E01)

(72) A1: Quero *ele* demitido!

Quero *ele* preso! (FT10E13)

6.1.2 A análise multivariada

Evidentemente, as tendências gerais apresentadas na subseção anterior não são, porém, suficientes para entender a variação linguística encontrada na expressão do OD anafórico de terceira pessoa em legendas audiovisuais. Desse modo, a fim de averiguar a influência de cada uma das VPs sobre o fenômeno, foram criados três modelos de regressão logística multinomial através da função *mlogit* no *R*, conferindo sempre se havia colinearidade entre as VPs por meio

da função $vif()$ do pacote para o R , para que se pudesse compará-los e eleger o mais adequado para a realização das análises. O primeiro apresentava os quatro níveis da VR e todas as VPs; o segundo, também os quatro níveis da VR, mas com a exclusão das VPs que não se mostraram significativas no primeiro modelo; e o terceiro, desconsiderando-se o pronome lexical devido à sua ocorrência virtualmente inexpressiva no *corpus*, continha apenas três níveis da VR e todas as VPs.

Dado que todos os modelos apresentaram valores de R^2 de McFadden⁶⁷ entre 0,2 e 0,4 – o que indica um ótimo ajuste (Louviere *et al.*, 2000) –, recorreu-se ao critério de informação de Akaike (AIC)⁶⁸ para auxiliar na seleção do modelo a ser empregado nas análises. Após o cálculo do AIC, também com o suporte da plataforma R , chegou-se a uma estimativa de 1384.726 para o primeiro modelo, 1380.730 para o segundo, e 1253.724 para o terceiro. Posto que, de maneira geral, valores menores de AIC indicam um ajuste melhor, selecionou-se o último modelo para a realização das análises, cujos resultados estão dispostos na Tabela 5:

Tabela 5 - Resultado da análise de regressão logística multinomial para a realização do OD anafórico de terceira pessoa em legendas audiovisuais (N = 955)

	Estimativa	Erro padrão	Valor-z	p		Aplicação/N
Clítico (ref.)						
<i>Intercept</i> 1: ON	-1,273	0,548	-2,321	0,02	*	
<i>Intercept</i> 2: SN	-2,663	0,702	-3,791	<0,001	***	
Traço semântico						
Animado (ref.)						456/530 (86%)
Não animado: ON	3,435	0,226	15,163	<0,001	***	250/425 (59%)
Não animado: SN	3,086	0,289	10,602	<0,001	***	90/425 (21%)
Função do antecedente						
Igual (ref.)						234/510 (46%)
Diferente: ON	-0,477	0,294	-1,618	0,10	+	29/193 (15%)
Diferente: SN	-0,670	0,366	-1,826	0,06	+	13/193 (6%)
NA: ON	0,277	0,242	1,147	0,25	+	81/252 (32%)

⁶⁷ O pseudo R^2 de McFadden é empregado para avaliar o ajuste de diferentes modelos, comparando o log-verossimilhança de um modelo com o modelo nulo (que contém apenas o *intercept*). Valores de R^2 de McFadden entre 0,2 e 0,4 indicam um ótimo ajuste do modelo, correspondendo a um R^2 de 0,7 a 0,9 em modelos lineares.

⁶⁸ O AIC é um método para seleção de modelos baseado na teoria da informação, que preza pelo ajuste e pela simplicidade do modelo. Nele, são considerados o número de variáveis predictoras e a estimativa de máxima verossimilhança do modelo para avaliar quanto da variação pode ser explicada utilizando-se o mínimo de variáveis possível.

NA: SN	-0,656	0,340	-1,928	0,05	+	15/223	(6%)
Forma verbal							
Infinitivo (ref.)						277/419	(66%)
Não finita: ON	0,777	0,215	3,616	<0,001	***	14/29	(48%)
Não finita: SN	0,059	0,265	2,108	<0,001	*	5/29	(17%)
Simples flexionada: ON	1,277	0,659	1,936	0,05	+	186/507	(37%)
Simples flexionada: SN	0,905	0,771	1,173	0,24	+	67/507	(13%)
Estrutura sintática							
S V OD (ref.)						294/591	(50%)
S V OD + complementos: ON	-0,166	0,223	-0,748	0,45	+	86/307	(28%)
S V OD + complementos: SN	-0,924	0,304	-3,030	0,002	**	20/307	(7%)
S V OD + minioração: ON	-0,948	0,526	-1,801	0,07	+	7/57	(12%)
S V OD + minioração: SN	-0,703	0,596	-1,178	0,23	+	4/57	(7%)
Tópico							
Pessoal (ref.)						263/432	(61%)
Profissional: ON	0,731	0,236	3,092	0,001	**	184/523	(35%)
Profissional: SN	0,374	0,291	1,283	0,19	+	61/523	(12%)
Relação entre participantes							
Simétrica (ref.)						219/421	(52%)
Assimétrica: ON	-1,026	0,233	-4,401	<0,001	***	153/514	(30%)
Assimétrica: SN	-0,983	0,288	-3,405	<0,001	***	54/514	(11%)
Tipo de legenda							
Profissional (ref.)						288/476	(61%)
Fansub: ON	0,081	0,211	0,384	0,70	+	166/479	(35%)
Fansub: SN	0,141	0,263	0,536	0,59	+	60/479	(13%)
CPL							
ON	-0,077	0,021	-3,554	<0,001	***		
SN	-0,035	0,026	-1,326	0,18	+		
CPS							
ON	0,028	0,028	0,996	0,31	+		
SN	0,074	0,034	2,124	0,03	*		

Modelo: mlogit(VD1 ~ 1 | TRACO.ANT + FUNC.ANT + FOR.VERB.A + EST.SINT.A + TOPICO + RELACAO + TIPO.LEGENDA + CPL + CPS, data = dfm2, relevel = 1)

Fonte: Elaborada pela autora

Antes de se debruçar, de fato, sobre os resultados, é importante retomar alguns comentários acerca do teste estatístico empregado. Como mencionado na seção 5, uma regressão logística multinomial realizada no modo “*wide*” permite a análise de uma VR fatorial com mais de dois níveis, contrastando, ao mesmo tempo, uma com as demais variantes. No caso desta pesquisa, o nível de referência adotado para a VR foi o clítico – dada sua presença expressiva no *corpus* –, o que significa que, para cada nível de cada VP, foram calculadas duas estimativas: a primeira comparando-o com o objeto nulo e a segunda, com o SN. Por exemplo, nos resultados apresentados na tabela acima, para o preditor “traço semântico do antecedente”, a primeira linha se refere aos coeficientes obtidos ao cotejar a categoria vazia com o clítico, enquanto a segunda traz as estimativas geradas pelo contraste entre o SN e o clítico.

Além disso, toda análise de regressão logística apresenta como *intercept* o nível de referência determinado tanto para a VR como para as diversas VPs categóricas incluídas no modelo⁶⁹. Neste estudo, trata-se, portanto, do clítico acusativo que retoma um antecedente de traço animado e de função sintática igual, empregado com uma forma verbal de infinitivo, em uma estrutura sintática simples (S V OD), ocorrido em uma interação de tópico pessoal e com relação simétrica entre as personagens, em uma legenda profissional. Em regressões logísticas multinomiais, contudo, os valores para o *intercept* são calculados também para cada variante da VR contrastada com a referência. Assim sendo, na análise aqui empreendida, são dois os coeficientes obtidos para o *intercept*: o primeiro associado ao objeto nulo, e o segundo, ao SN. O trecho (73) ilustra um exemplo que representa todos os níveis definidos para referência:

(73) A2: Arizona, ela também sumiu.

Não os encontro em lugar algum.

AZ1: Graças a Deus.

Não é tudo minha culpa.

A2: -Pode me ajudar a procurá-*los*? (PT14E24)

Todas as estimativas geradas pelo modelo adotado são apresentadas em *logodds*, o que permite a observação do tamanho e da direção do efeito de todos os fatores considerados na realização do fenômeno. Coeficientes positivos demonstram um aumento nas chances de certa

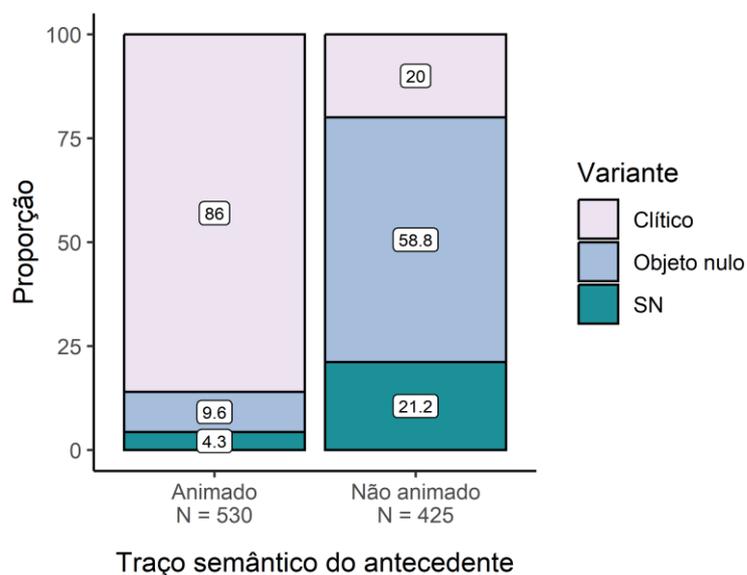
⁶⁹ No que diz respeito às VPs numéricas, como é o caso do número de CPL e de CPS, as estimativas indicam o quanto aumenta ou diminui a chance de certa variante ser empregada conforme aumenta a unidade da VP. Para as variáveis numéricas deste estudo, por exemplo, os valores indicam o quanto a chance de usar uma das formas contrastadas se altera a cada caractere a mais utilizado. Essa especificidade será retomada mais à frente, ao se analisar os resultados para as duas variáveis numéricas da pesquisa.

variante ocorrer em relação à referência, enquanto estimativas negativas apontam para um desfavorecimento. Ademais, cada modelo calcula a significância de cada categoria das VPs na escolha por uma ou outra variante, sendo indicada pelo valor-*p*. Conforme destacado anteriormente, valores-*p* abaixo de 0,05 indicam a probabilidade de se observar uma determinada distribuição caso a hipótese nula seja verdadeira, devendo-se, portanto, acatar a hipótese alternativa. Nesta investigação, cuja hipótese nula é a de que não há correlação entre as variáveis, um valor-*p* abaixo de 0,05 leva à aceitação da hipótese alternativa, isto é, a de que as variáveis consideradas são estatisticamente correlacionadas.

Com isso em mente, é possível perceber, a partir do disposto na Tabela 5, que, no contexto geral das legendas audiovisuais, a categoria vazia é desfavorecida em -1,273 *logodds* em relação ao clítico acusativo, e o SN em -2,663 *logodds*. Isso quer dizer que, comparadas ao nível de referência, as chances de as duas formas serem utilizadas são menores do que as do clítico, sendo o SN a variante menos provável de ocorrer. Ainda, o valor-*p* obtido para essas duas variantes (0,01 para o objeto nulo e <0,001 para o SN) indica uma diferença estatisticamente significativa, confirmando o que foi observado na distribuição geral dos dados.

Voltando o olhar, agora, ao traço semântico – primeiro preditor que se correlacionou com o fenômeno –, pode-se perceber pelas estimativas que o clítico acusativo foi favorecido ao retomar referentes animados, ao passo que a categoria vazia e o SN foram favorecidos em, respectivamente, 3,435 e 3,068 *logodds* quando o antecedente era não animado. Comparando os dois valores, observa-se, também, que o objeto nulo foi a estratégia mais favorecida nesse contexto, o que confirma a hipótese levantada na seção 5 de que antecedentes de traço [-animado] seriam preferencialmente retomados por complementos não lexicalizados. Para uma melhor visualização do resultado da realização do OD anafórico de terceira pessoa em relação ao traço semântico do antecedente, pode-se observar as frequências e proporções apresentadas no Gráfico 14:

Gráfico 14 - Frequências e proporções das variantes do OD anafórico de terceira pessoa de acordo com o traço semântico do antecedente (N = 955)



Fonte: Elaborado pela autora

Como mostra o gráfico, as distribuições das três variantes sob investigação foram bastante distintas a depender do traço semântico do antecedente. Para retomadas de antecedentes animados, a proporção de uso do clítico foi de expressivos 86%, superando em muito a soma do objeto nulo (9,6%) com o SN (4,3%). Já no caso de antecedentes não animados, a estratégia preferida pelos tradutores foi o objeto nulo, correspondente a 58,8% das ocorrências, enquanto as porcentagens de clíticos (20%) e SNs (21,2%) ficaram mais próximas uma da outra. Esses números, em conjunto com as estimativas geradas pelo modelo de regressão logística multinomial utilizado nas análises, permitem afirmar que as diferenças encontradas na realização do OD anafórico de terceira pessoa não são aleatórias, indicando uma correlação entre essa variável e o fenômeno em questão.

Tal constatação corrobora resultados anteriormente reportados na literatura, que afirmam que a dominância do objeto nulo diante de antecedentes de traço [-animado] pode ser explicada através do processo que difundiu essa variante no PB. Com efeito, foi a partir da retomada de antecedentes não referenciais e não animados através da categoria vazia, ao invés do clítico neutro, que a forma inovadora ganhou espaço (cf. Pagotto, 1992; Cyrino, 1994; Cyrino; Duarte; Kato, 2000, entre outros), chegando a ser a variante atualmente preferida no PB oral.

Os exemplos (74) e (75) a seguir trazem casos de antecedentes animados e não animados, referenciados pelo clítico e pelo objeto nulo, respectivamente:

(74) M3: Ele está em Bagdá

com o exército americano.

Não pude adotá-**lo** legalmente,
e ele é um refugiado sírio.

Levará anos para trazê-**lo**
para cá, se eu conseguir. (F14E01)

(75) AZ1: Mas acho que vamos resolver tudo,

que isso vai nos fortalecer.

Sabe? Como um osso.

Às vezes é preciso quebrar [Ø]
para solidificar certo. (PT10E01)

Analisando-se, então, a forma verbal, percebe-se, pelas estimativas dispostas na Tabela 5, que as formas verbais compostas por infinitivo (simples ou complexas) favoreceram o uso do clítico acusativo em comparação com os outros dois níveis da variável. Já diante de outras formas não finitas – como o particípio e o gerúndio –, foram o SN (0,777) e o objeto nulo (0,059) as variantes favorecidas em comparação com o pronome acusativo, sendo a primeira estratégia mais favorecida do que a segunda. Os excertos (76) e (77) que se seguem exemplificam, respectivamente, os contextos favoráveis ao clítico e às outras duas estratégias.

(76) L2: Ela perguntou e você não disse nada.

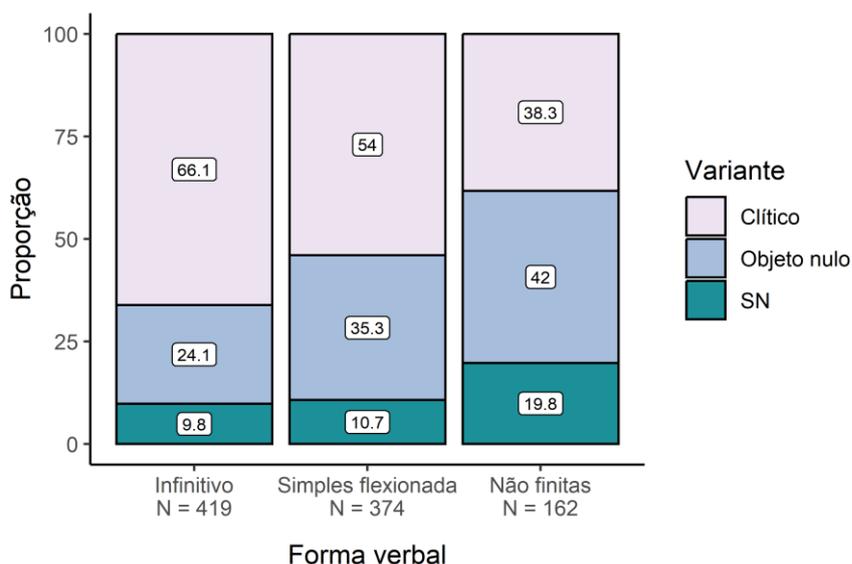
O que, estava esperando o momento
certo para humilhá-**la** publicamente? (PT10E13)

(77) Não vou operar uma paciente

que não me quer operando [Ø]. (FT06E13)

A fim de facilitar a visualização dos resultados, pode-se, ainda, recorrer ao Gráfico 15, que apresenta a distribuição das variantes do OD anafórico de terceira pessoa de acordo com a forma verbal:

Gráfico 15 - Frequências e proporções das variantes do OD anafórico de terceira pessoa de acordo com a forma verbal (N = 955)



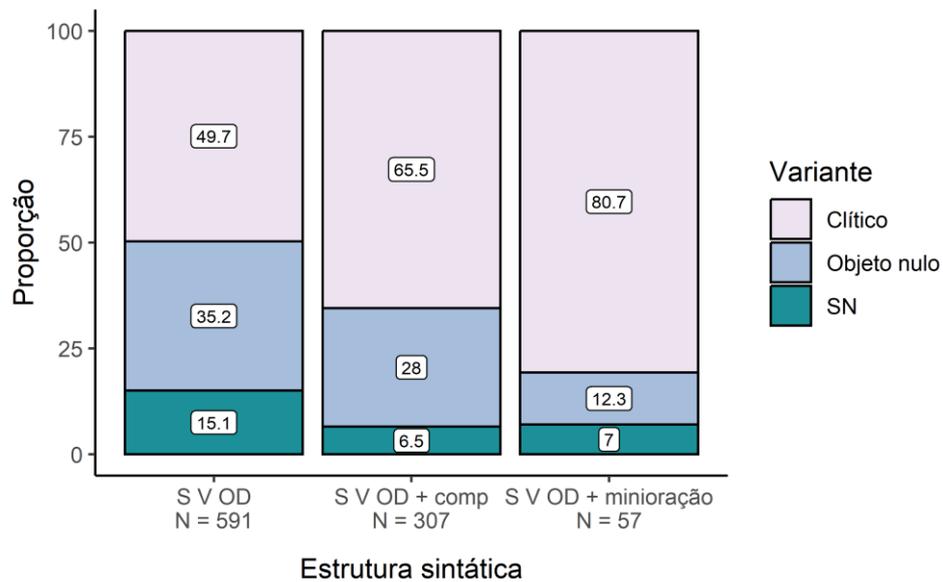
Fonte: Elaborado pela autora

Como pode ser verificado no gráfico, é diante de formas verbais de infinitivo que o OD tende a ser mais realizado através do clítico acusativo, que representou 66,1% dos dados encontrados nesse contexto, em contraste com os 24,1% de objetos nulos e 9,8% de SNs. Por outro lado, quando utilizado em conjunto com outras formas não finitas, a situação é bastante diversa: a variante preferida nos 162 dados analisados passou a ser o objeto nulo, com 42% das ocorrências, que, em conjunto com os SNs (19,8%), superaram a produtividade do pronome acusativo, representante de 38,3% dos ODs coletados. Com efeito, conforme diminui a expressividade do clítico em cada contexto, parecem aumentar as taxas de utilização do objeto nulo e do SN, corroborando, portanto, os resultados da regressão logística multinomial empreendida.

Dessa maneira, as expectativas para essa variável foram, novamente, apenas parcialmente confirmadas. De fato, o clítico foi a variante mais favorecida com formas de infinitivo simples e complexas, superando a soma das outras variantes, conforme descrito, por exemplo, por Marafoni (2004), Freire (2005) e Lima (2022). Entretanto, no que diz respeito às outras formas não finitas, o cenário não foi o mesmo, com a categoria vazia assumindo a posição de variante preferida. Ademais, no contexto das formas simples flexionadas, o clítico seguiu representando metade das ocorrências, apontando para a expressividade dessa variante em diversos contextos.

Já no que toca à estrutura sintática, as estimativas geradas pelo modelo de regressão logística multinomial indicam que essa variável se correlacionou significativamente com a VR, sendo que, quando o SN foi empregado em sentenças de estrutura S V OD + complementos, ou seja, em orações nas quais, além do OD, havia, também, um objeto indireto ou um complemento oblíquo, o SN foi desfavorecido em $-0,924$ *logodds* em relação ao clítico – o nível de referência. A fim de auxiliar na interpretação desses resultados, o Gráfico 16 exhibe as porcentagens calculadas para cada nível da VR.

Gráfico 16 - Frequências e proporções das variantes do OD anafórico de terceira pessoa de acordo com a estrutura sintática (N = 955)



Fonte: Elaborado pela autora

As proporções apresentadas no gráfico demonstram que o clítico foi a forma mais utilizada em todos os tipos de estrutura, seguindo a tendência geral da amostra, sendo maior do que a soma das outras variantes nas estruturas complexas (com complementos ou de minioração). Já o SN, de maneira contrária, foi consistentemente a variante menos empregada, atingindo sua menor taxa de aplicação (6,5%) nas construções de S V OD + complementos. No trecho (78), pode ser observado um exemplo de SN utilizado na estrutura que o desfavoreceu e, no (79), um caso de clítico utilizado em conjunto com uma minioração.

(78) C2: Meredith? O Owen está morto?

Dra. Grey! O Owen está morto?

M1: Ele está vivo. Está vivo.

Inconsciente, mas vivo.

C2: Leve **Owen** para a sala em frente e retire
essa bala. (PT06E24)

(79) M1: Você manda lá fora,
mas eu mando aqui.

Saia do meu caminho.

RP1: Deixem-**na** passar. (FT10E24)

Sobre esse último tipo de estrutura, o que contém miniorações, é válido comentar que, mesmo o pronome lexical não tendo sido incluído na análise multivariada devido ao baixo número de ocorrências dessa estratégia no *corpus*, foi justamente em estruturas com miniorações que ele se mostrou mais produtivo. Dos 17 dados encontrados nas legendas audiovisuais, 7 deles (41%) ocorreram quando o objeto exercia uma dupla função, o que atesta o que já foi reportado na literatura (cf. Duarte; Marafoni, 2004). Para exemplificar, são apresentados os exemplos (80) e (81) a seguir, nos quais o OD é seguido por um verbo e por um predicativo.

(80) D1: Ross, fique aqui.

Bipei a Brooks.

Mande **ela** fazer uma TC de cabeça,
tórax e abdômen da Srta. Ashford. (PT10E01)

(81) LO1: Ronnie e Jerry estão longe,
numa pescaria no Alasca,

e quando voltarem para casa terei
que contar para eles que Georgie...

É demais para mim.

Não posso fazer mais...

Não posso.

Não consigo nem ver
ele desse jeito, não dá. (PT06E01)

Assim, foi possível confirmar as hipóteses formuladas para essa variável, na medida em que o clítico foi, realmente, a variante mais expressiva em todos os contextos, e o pronome lexical, apesar da baixa produtividade, foi mais presente em sentenças de estrutura S V OD + minioração.

Passando, agora, à análise das variáveis extralinguísticas, comenta-se, primeiramente, o efeito do tópico da conversa na expressão do fenômeno aqui em foco. Segundo os coeficientes gerados pelo modelo de regressão logística criado para esta pesquisa, listados na Tabela 5, com relação ao clítico acusativo, o objeto nulo foi favorecido em 0,731 *logodds* quando utilizado em interações cujo tópico era profissional. Isso significa que, se a conversa versa sobre aspectos do trabalho das personagens, aumentam as chances de ser a categoria vazia a variante escolhida para retomar um antecedente em comparação com o pronome acusativo. Os diálogos (82) e (83) a seguir ilustram o contexto que aumentou as chances de realização do objeto nulo e uma interação de cunho pessoal, na qual o clítico foi empregado:

(82) B1: Você mandou uma mensagem
 ao Dr. Webber também?

O1: Não mandei [Ø]. (PT14E13)

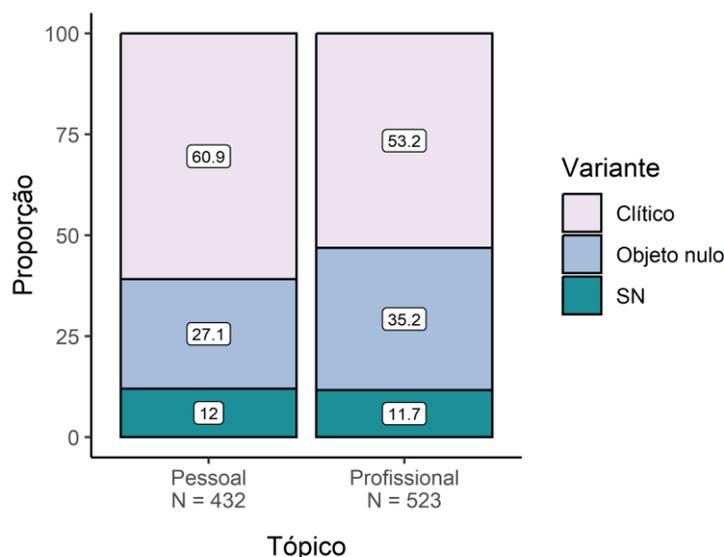
(83) L2: - Você nem gosta dela.

S1: - Caramba.

J2: Ignore-*a*, Steph. (FT10E01)

Tal favorecimento não é suficiente, porém, para fazer com que a categoria vazia seja a variante mais expressiva, mesmo em interações profissionais, conforme pode ser conferido no Gráfico 17:

Gráfico 17 - Frequências e proporções das variantes do OD anafórico de terceira pessoa de acordo com o tópico (N = 955)



Fonte: Elaborado pela autora

De acordo com as informações representadas visualmente no gráfico acima, torna-se fácil perceber que, assim como para a maioria das outras variáveis até aqui apresentadas, a estratégia preferida para a realização do OD anafórico de terceira pessoa nas legendas do *corpus* foi o clítico, independentemente do tópico da conversa entre os personagens. Quando o assunto dos diálogos foi pessoal, o pronome acusativo, variante padrão comumente associada a situações de maior formalidade, correspondeu a 60,9% das ocorrências, enquanto o objeto nulo ficou restrito a 27,1% e o SN a 12% dos dados. Já nas interações profissionais, a categoria vazia parece ganhar mais espaço, atingindo 35,2%, em detrimento do clítico acusativo, cuja proporção cai para 53,2%. Tal resultado, apesar de surpreender em um primeiro momento, acaba por reforçar a preferência pelo clítico nas legendas do *corpus*, apontando para uma tendência generalizada.

Ainda, assim como no caso da estrutura sintática, é importante traçar um comentário acerca do uso de pronomes lexicais nas legendas do *corpus* em relação ao tópico, mesmo que, por uma decisão metodológica, eles tenham sido excluídos da análise multivariada. Dentre as 17 ocorrências dessa variante encontradas nas legendas, como a representada pelo exemplo (84), 13 foram coletadas de diálogos acerca da vida pessoal das personagens, isto é, 76,5% dos dados (cf. Gráfico 18), o que pode indicar, apesar do baixo número de dados, que os tradutores de ambos os tipos de legenda têm consciência de que essa forma é mais adequada em contextos de menor monitoramento.

(84) PC2: - Não acredito que foram

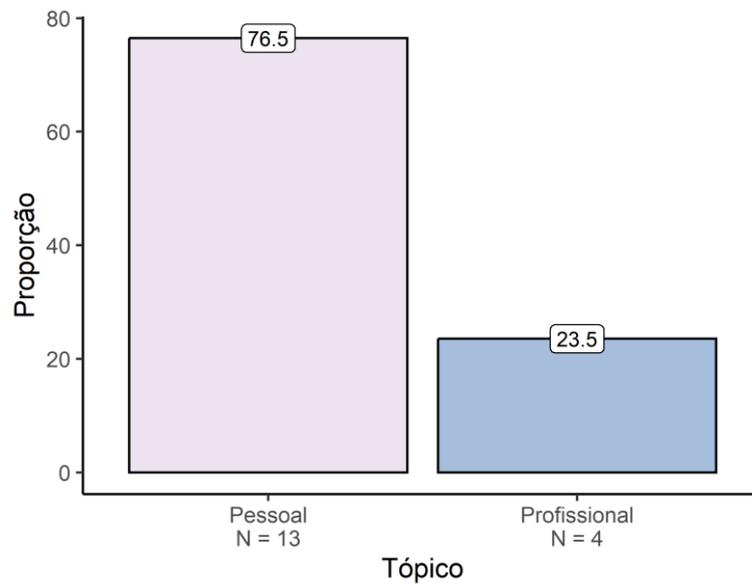
no 'Burning Man', bruxas.

Nunca ouviram falar

do 'Burning Man'.

Odeio *elas*! (FT06E01)

Gráfico 18 – Distribuição dos pronomes lexicais em legendas de acordo com o tópico (N = 17)



Fonte: Elaborada pela autora

Diante do exposto acerca dessa variável, foi possível comprovar as hipóteses levantadas para o tópico, uma vez que o clítico acusativo foi a variante preferida nos dois contextos, além de o objeto nulo ter sido favorecido em diálogos profissionais. Ademais, como se supunha, os pronomes lexicais foram mais encontrados em conversas pessoais do que profissionais, uso que se aproxima da linguagem que, ao menos em teoria, a legenda deveria emular.

Nesse sentido, é interessante observar a conexão da relação entre os participantes com a realização do OD anafórico de terceira pessoa nas legendas. Conforme evidenciado pelas estimativas na Tabela 5, o clítico foi favorecido em diálogos nos quais a relação entre as personagens era assimétrica, enquanto o objeto nulo e o SN foram desfavorecidos de maneira significativa, em -1,026 e -0,983 *logodds*, respectivamente. Desse modo, as chances de se verificar o uso da variante padrão em interações entre personagens que não compartilham o

mesmo nível de poder, como o apresentado no exemplo (85), são maiores do que quando a conversa se dá entre indivíduos do mesmo *status* relativo, caso representado pelo excerto (86).

(85) WP1: Os convidados estão aqui, mas e a noiva?

Ela quer a menor cerimônia
da história?

A2: Vou procurá-*la*. (PT14E24)

(86) C2: - Dê as pás internas.

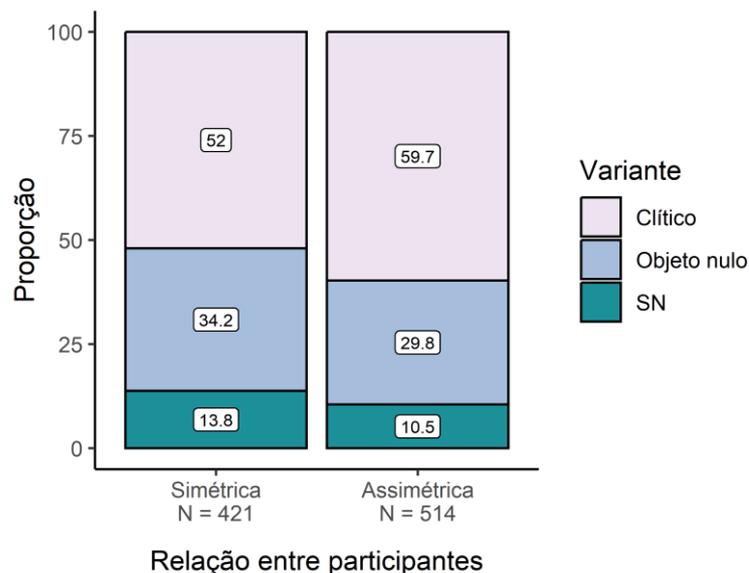
J1: - 1 de epi.

C2: - Abril, pegou [Ø]?

A2: - Sim. (FT06E24)

Para completar o panorama da expressão do acusativo anafórico nas legendas pensado a partir da relação entre os participantes, pode-se, ainda, examinar as frequências e proporções de cada variante, apresentadas no Gráfico 19 a seguir:

Gráfico 19 - Frequências e proporções das variantes do OD anafórico de terceira pessoa de acordo com a relação entre participantes (N = 955)



Fonte: Elaborado pela autora

Assim como vem sendo recorrente para as outras variantes, em termos absolutos, o pronome acusativo foi a variante mais expressiva em ambos os contextos analisados para a variável “relação entre participantes”, correspondendo a 52% dos dados quando a relação era simétrica e a 59,7% das ocorrências quando era assimétrica. Apesar das proporções próximas, verifica-se que as relações simétricas permitem um uso maior das outras formas que concorrem com o clítico, enquanto nas relações assimétricas a variante padrão é responsável por praticamente dois terços do total de ODs coletados.

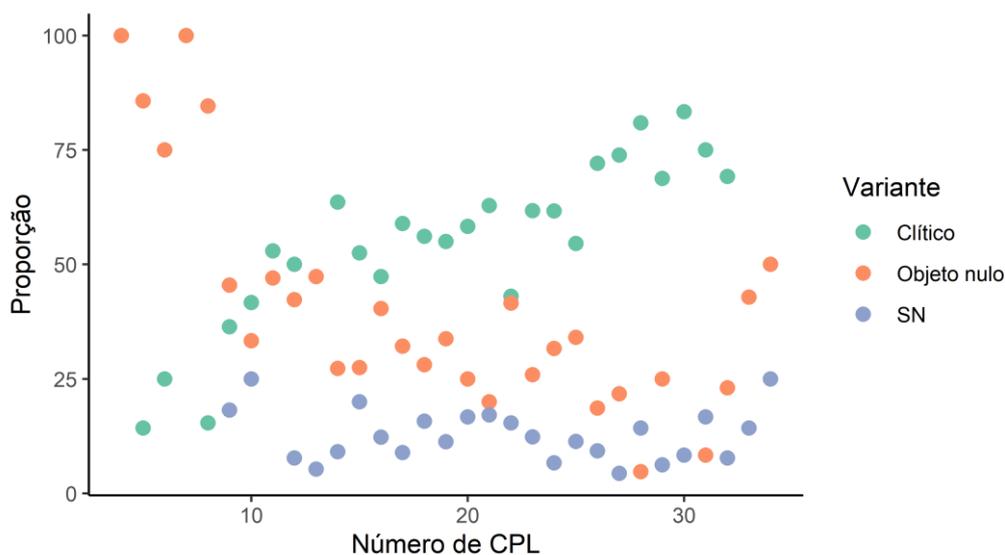
Após a interpretação desses resultados e daqueles do modelo de regressão logística multinomial, pôde-se confirmar as principais hipóteses construídas para essa variável. Assim como previsto, a maior parte dos clíticos foi encontrada quando a relação entre os participantes dos diálogos foi assimétrica, e o objeto nulo teve mais espaço no contexto contrário. Entretanto, cabe mencionar, ainda, que a hipótese para o pronome lexical de que ele seria mais produtivo em relações simétricas foi refutada, uma vez que ele se distribuiu de maneira bastante semelhante entre os dois contextos observados – foram 8 dados em relações simétricas e 9 em relações assimétricas. Ademais, cabe reconhecer que o estudo dessa variável ainda pode ser aprofundado, refinando o olhar sobre as relações assimétricas para pensar se, quando as personagens se dirigem a alguém de maior ou menor poder, elas realizam o OD anafórico de terceira pessoa de maneiras distintas, isto é, se interações assimétricas ascendentes ou descendentes entre as personagens poderiam trazer consequências para a variação.

Chegando, enfim, às variáveis ligadas a aspectos estruturais do gênero textual-discursivo sob investigação, vale fazer um breve comentário acerca da natureza dos resultados gerados na análise multivariada. Diferentemente do que vem sendo apresentado, quando se inclui uma variável numérica como o número de CPL e de CPS em uma regressão logística multinomial, as estimativas calculadas não dizem respeito ao favorecimento ou não de uma forma em relação ao nível de referência, mas, sim, ao quanto aumenta ou diminui a chance de se encontrar uma dada variante, em contraste com a referência, a cada unidade do preditor em questão. No caso das duas variáveis contínuas deste estudo, isso significa que os coeficientes apresentados na Tabela 5 vão indicar se as chances do objeto nulo e do SN nominal serem empregados aumentaram ou diminuíram a cada caractere a mais utilizado na legenda.

No que diz respeito aos resultados para o número de CPL, que se correlacionou com a VR, observa-se que, a cada novo caractere adicionado à legenda, diminuiu em $-0,077 \log odds$ a chance de um objeto nulo ser escolhido para realizar o OD anafórico. Assim, quanto mais CPL tiver uma legenda, menor será a chance de uma categoria vazia ser escolhida ao invés de

um clítico que, mais uma vez, é dominante na amostra. Tal relação pode ser bem observada através do Gráfico 20:

Gráfico 20 - Proporção de utilização das variantes do acusativo anafórico de terceira pessoa de acordo com o número de CPL



Fonte: Elaborado pela autora

Conforme ilustrado no Gráfico 20, observa-se uma correlação negativa entre a categoria vazia e o número de CPL em legendas de até 20 caracteres, com a proporção de objetos nulos chegando a superar a dos clíticos em legendas com menos de 10 CPL. Nesse contexto, foram comuns dados retirados de legendas em que houve repetição do verbo que precede o OD anafórico, em que o objeto nulo apareceu em resposta a uma pergunta do tipo “sim” ou “não” e em enunciados com emprego de imperativo, representados, respectivamente, pelos exemplos (87) a (89). Já no caso dos clíticos, parece haver uma correlação linear entre o número de CPL e a ocorrência dessa variante, na qual quanto mais caracteres são empregados nas legendas, maiores as chances de se encontrar o pronome acusativo. Em contrapartida, a variável número de CPL aparentemente não foi decisiva para o emprego de SNs anafóricos, tendo sua curva se mantido bastante estável independentemente do número de CPL empregados.

(87) M1: Ela merece a dignidade
da história completa.

A1: Merece [Ø]. (PT14E01)

(88) B1: Você mandou uma mensagem

ao Dr. Webber também?

O1: Não mandei [Ø]. (PT14E13)

(89) A1: - O que é isto?

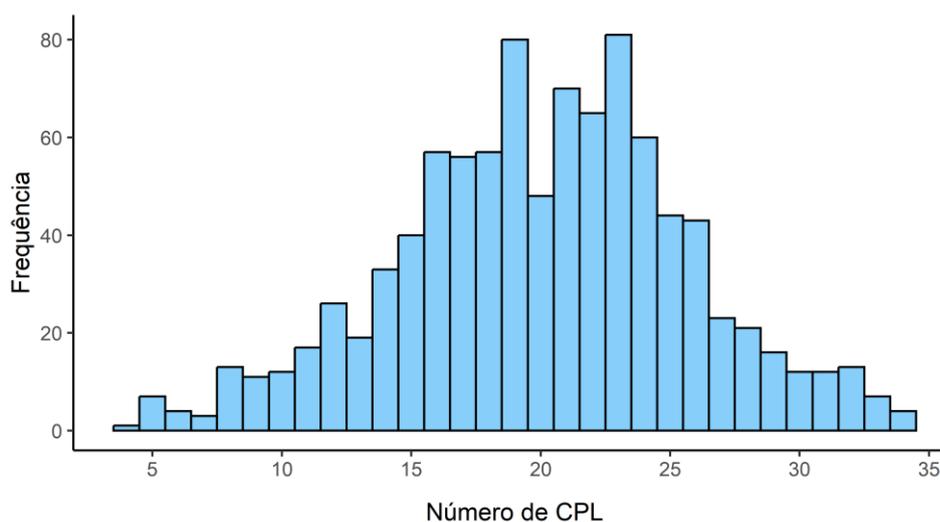
D1: - São as chaves do meu trailer.

A sua nova casa.

Aproveitem [Ø]. (FT06E01)

Esse resultado, que contraria a hipótese inicial de que números de caracteres mais próximos ao limite estabelecido para a legenda favoreceriam o objeto nulo, pode, porém, ser explicado ao se considerar tanto o número de CPL das legendas do *corpus* como o processo de revisão ao qual toda legenda é submetida. Quanto ao primeiro ponto, é necessário reconhecer que a maioria das legendas analisadas, conforme indica o Gráfico 21 abaixo, sequer se aproximou do número máximo de CPL permitido, ficando a média dessa variável em torno de 20 caracteres, ou seja, bem abaixo do limite. Do total de ocorrências consideradas nesta análise, por exemplo, apenas 68 (7%) foram encontradas em legendas que tinham mais do que 30 caracteres, tendo sido 35 o número máximo de CPL verificado, de modo que as legendas parecem respeitar, no geral, as determinações para esse parâmetro técnico.

Gráfico 21 - Distribuição dos dados em relação ao número de CPL das legendas do *corpus*



Fonte: Elaborado pela autora

Já no que toca ao segundo aspecto mencionado, é importante considerar que, ao elaborar uma legenda, o tradutor propõe uma tradução preliminar para as falas, que pode ou não ser posteriormente alterada durante a revisão. Assim, as legendas que chegam para o público podem não ter passado por qualquer processo de adequação que visasse à diminuição de seu número de CPL, caso essa necessidade não tenha sido identificada durante sua criação. É perfeitamente plausível, portanto, pensar que ao menos parte das legendas às quais se teve acesso para a realização desta pesquisa consiste em versões iniciais da tradução, nas quais a utilização do objeto nulo não foi necessariamente afetada pela preocupação com o limite de CPL.

Além disso, o próprio processo de adequação da legenda aos parâmetros técnicos pressupõe uma diminuição no número de caracteres utilizados, o que aumenta a tendência de se ter legendas mais curtas. Por exemplo, caso um SN como “minha mão”, encontrado no exemplo (90) abaixo, fosse substituído por uma categoria vazia, o número de caracteres cairia de 31 para 22, contando o espaço, o que significa que o dado não mais se aproximaria do limite de CPL. Portanto, é também possível que o número maior de objetos nulos em legendas mais sucintas seja, precisamente, resultado desse tipo de operação.

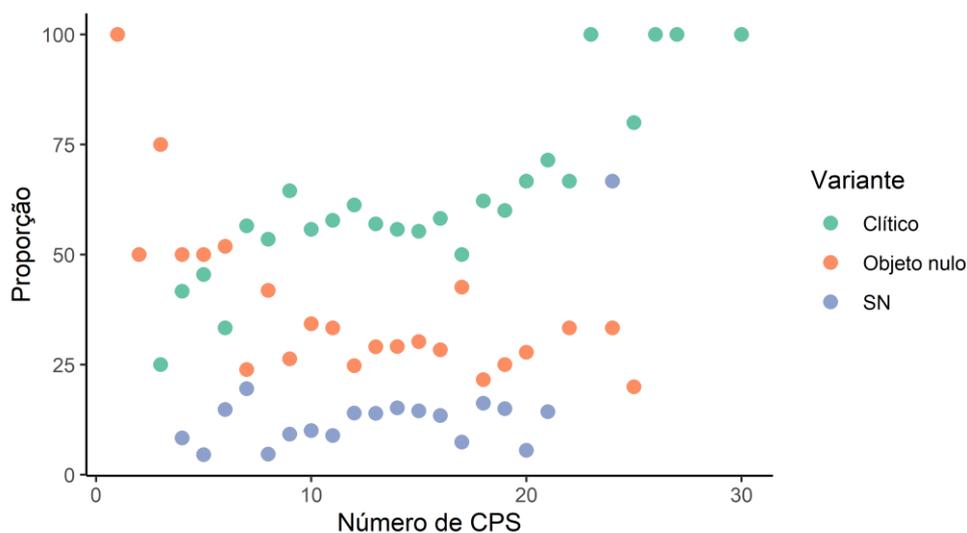
(90) M1: Herrera, tire sua mão daí.

HE1: Você me disse que se eu retirasse

minha mão, ele iria sangrar em minutos. (PT14E13)

Chegando, enfim, ao número de CPS, nota-se de antemão que essa variável, assim como o número de CPL, também se correlacionou com a VR no *corpus* de legendas. De acordo com os resultados dispostos na Tabela 5, a probabilidade de se encontrar um SN anafórico na posição de OD aumentou em 0,074 *logodds* a cada caractere por segundo a mais verificado nas legendas. Em outras palavras, as estimativas indicam que quanto mais CPS uma legenda apresentar, maior será a chance de o SN ser a variante empregada em relação ao clítico, mesmo que esta última se conserve como a preferida dentre as outras. O Gráfico 22 auxilia na visualização dessa tendência:

Gráfico 22 - Proporção de utilização das variantes do acusativo anafórico de terceira pessoa de acordo com o número de CPS



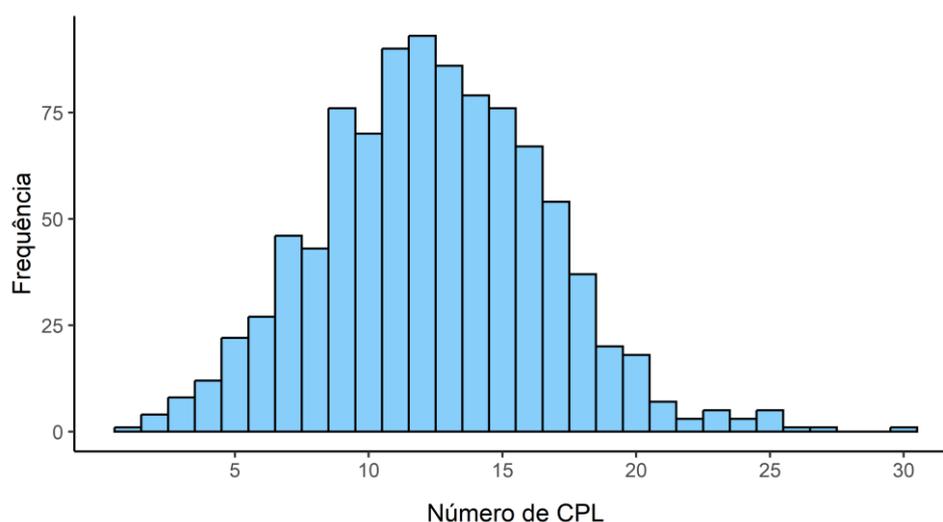
Fonte: Elaborado pela autora

Assim como aconteceu para a variável anteriormente explorada, é possível observar padrões bastante distintos ao considerar as proporções de clíticos e categorias vazias em relação ao número de CPS. Conforme aumenta o número de caracteres, diminui a ocorrência de objetos nulos, que são raramente encontrados em legendas com mais de 20 CPS, e aumenta a produtividade do clítico, que pode ser encontrado em legendas que chegam a apresentar 30 CPS. Da mesma maneira que foi verificado para a variável número de CPL, é apenas em legendas com poucos CPS que o objeto nulo se mostra mais expressivo do que o clítico, sendo muitos desses dados encontrados nos mesmos contextos já destacados acima. O que mais chama a atenção, porém, é que, no caso dos SNs anafóricos, percebe-se um leve aumento nas taxas de aplicação dessa variável conforme aumenta o número de CPS, tendência confirmada pelas estimativas geradas pelo modelo de regressão utilizado.

Dessa maneira, a hipótese levantada para investigar a influência do número de CPS na realização do OD anafórico de terceira pessoa também foi refutada, dado que, quanto mais a legenda se aproximou do limite de CPS, mais cresceram as chances de se encontrar um SN – justamente a forma que se imaginava ser rechaçada nesse contexto, devido à sua extensão. Contudo, conjectura-se que, aqui, os mesmos dois fatores que foram mencionados para explicar os resultados da análise do número de CPL podem estar, também, relacionados ao que foi encontrado para os CPS.

Com efeito, as legendas em geral se distanciaram consideravelmente do limite de CPS definido pelas orientações técnicas dadas aos tradutores (cf. Gráfico 23), sendo a média para essa variável de 12 caracteres, isto é, 9 a menos do que o número máximo permitido. Dentre todos os dados coletados para o *corpus*, somente 65 (6,8%) corresponderam a ocorrências em legendas que ultrapassavam os 19 CPS, com 19 delas extrapolando o teto de 21 caracteres. Mais uma vez, pode ser que esse fato impeça a compreensão do que realmente ocorre na realização do OD anafórico de terceira pessoa em legendas que estejam próximas ao limite de CPS determinado.

Gráfico 23 - Distribuição dos dados em relação ao número de CPS das legendas do *corpus*



Fonte: Elaborado pela autora

Naturalmente, posto que as legendas analisadas não necessariamente precisaram ser adequadas com relação ao número de CPS, pelos mesmos motivos acima discutidos, é possível que elas sejam fruto da primeira versão da tradução, que, novamente, nem sempre é realizada priorizando-se variantes menos extensas de maneira consciente. Ademais, as alterações realizadas a fim de atender aos requisitos técnicos de CPS também resultam em legendas mais curtas. No trecho (91) a seguir, se o tradutor tivesse escolhido utilizar o objeto nulo, por exemplo, no lugar do SN “a ressonância”, o número de CPS diminuiria de 24 – o que estoura o limite – para 17, de modo que não se poderia mais considerar que ela se aproxima do máximo de CPS.

(91) R1: Você pediu uma ressonância 3D
de emergência pra este paciente?

AZ1: Certo.

Está zangado, entendi.

Mas com o plano de
saúde da mãe dele,

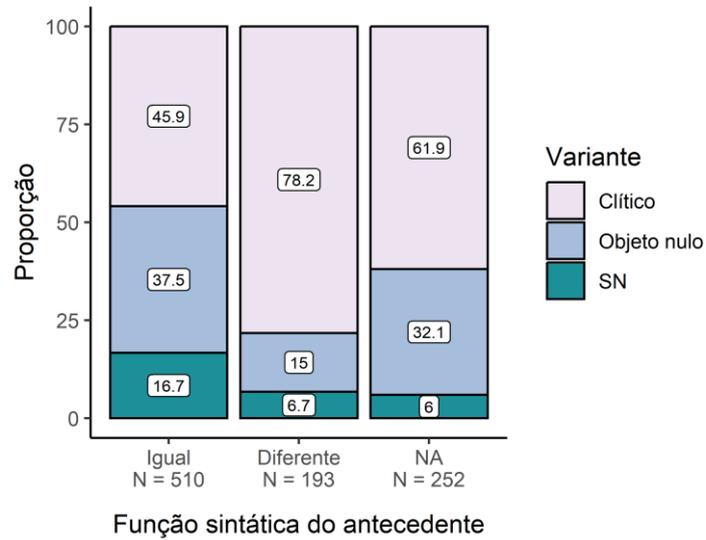
poderia levar semanas para
conseguir *a ressonância*,

principalmente com
as queixas de dor... (FT06E01)

Antes de encerrar as discussões empreendidas nesta seção, faz-se interessante averiguar qual foi a distribuição das variantes de acordo com a função do antecedente e com o tipo de legenda, mesmo que essas duas variáveis não tenham se correlacionado de maneira significativa com o fenômeno aqui em estudo. Com isso, busca-se contrastar as hipóteses para elas elaboradas, além de apontar especificidades que podem ter levado aos resultados encontrados.

No que diz respeito à função do antecedente, as frequências e as proporções das variantes da realização do OD anafórico de terceira pessoa, apresentadas no Gráfico 24, indicam que o clítico foi a variante mais utilizada nos três contextos investigados. Apesar disso, chama a atenção o fato de o objeto nulo ter sido mais expressivo na retomada de antecedentes que também exerciam função de OD, apresentando uma porcentagem de 37,5% e concorrendo com o clítico, consoante estudos anteriores como Cyrino (1994), Marafoni (2004) e Santana (2016).

Gráfico 24 - Frequências e proporções das variantes do OD anafórico de terceira pessoa de acordo com a função sintática do antecedente (N = 955)



Fonte: Elaborado pela autora

Ademais, é relevante observar o que aconteceu para os casos em que o antecedente não pôde ser identificado – por ser apresentado visualmente, ter sido mencionado em outra cena ou não estar encaixado em uma oração recuperável –, já que esse foi um nível definido especialmente em decorrência da natureza do gênero textual-discursivo sendo observado. Conforme apresentado acima, nos casos em que não foi possível determinar a função sintática do antecedente, observou-se uma predominância do clítico acusativo, que representou 61,9% dos dados, seguido pelo objeto nulo, com 32,1% e, enfim, pelo SN, com 6%. Os excertos (92) e (93) a seguir ilustram uma ocorrência com a forma menos expressiva nesse contexto e outro com o objeto nulo fazendo referência a um OD, nessa ordem.

(92) H1: Lidocaína e vários 4x4.

Pode levar *isso* para o Hunt? (PT10E01)

(93) M1: Também sinto muito

por oferecer o emprego.

Se eu fosse você,

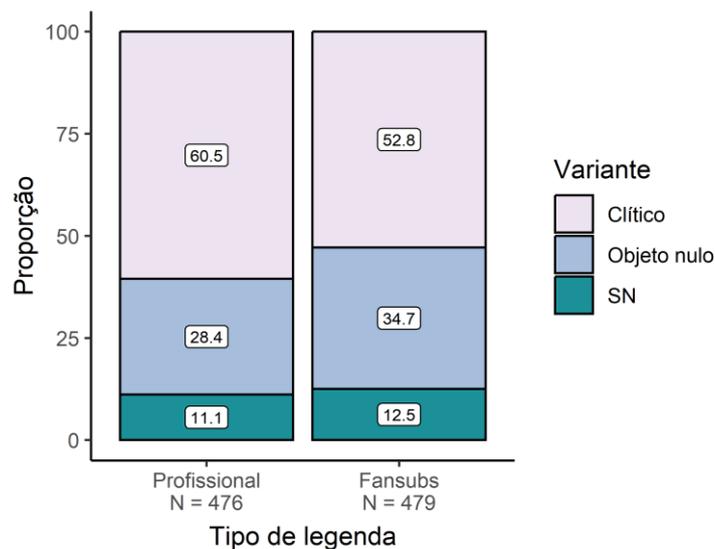
não aceitaria [Ø]. (FT14E24)

Dessa maneira, a partir da interpretação dos dados, não foi possível confirmar completamente as hipóteses levantadas. De fato, a categoria vazia foi mais utilizada diante de

antecedentes de função sintática igual, porém não houve correlação estatisticamente significativa, assim como verificado em Lima (2022). Além disso, indo em direção contrária ao que se supunha, o objeto nulo não foi a variante mais empregada quando a função sintática do antecedente não foi indicada, mas, sim, o clítico, resultado que parece estar em concordância com a tendência geral de emprego do pronome acusativo no *corpus*. Contudo, o fato de o SN ser, aparentemente, menos presente nesse contexto se adequa à lógica que motivou a hipótese: podendo os antecedentes serem facilmente recuperados, não há necessidade de mencioná-los e, muito menos, de empregar processos de referência da mesma maneira que em um texto escrito.

Por fim, resta tecer comentários acerca dos resultados obtidos para a variável tipo de legenda. Desde a concepção desta pesquisa, presumia-se que a linguagem utilizada em legendas profissionais e *fansubs* seria distinta, devido aos seus contextos de produção e de recepção distintos, conforme explorado na seção 4. Entretanto, como indicaram tanto as frequências e proporções calculadas para esse fator, apresentadas no Gráfico 25, como as estimativas geradas pelo modelo de regressão criado (Tabela 5), a língua empregada nesses dois tipos de legenda se aproxima muito mais do que o esperado, sendo que essa variável não se correlacionou com a realização do OD anafórico de terceira pessoa no *corpus*.

Gráfico 25 - Frequências e proporções das variantes do OD anafórico de terceira pessoa de acordo com o tipo de legenda (N = 955)



Fonte: Elaborado pela autora

Como observado acima, a variante preferida pelos tradutores foi o clítico acusativo, que, nos dois conjuntos de dados, representou mais do que a metade das ocorrências – 60,5% para

as profissionais e 52,8% para as legendas amadoras –, corroborando a tendência global observada no *corpus*. Em seguida, também em consonância com o que já foi apresentado, houve uma preferência pelo uso da categoria vazia, seguida pelo SN, a variante menos empregada, apresentando taxa de utilização pouco maior do que 10% nas amostras em questão.

Assim, nenhum dos dois tipos de legenda se aproximou de maneira mais clara dos preceitos da norma-padrão, indo contra ao que foi hipotetizado. Pelo contrário, as duas amostras apresentaram uso expressivo da variante padrão, com pouco espaço para formas estigmatizadas como o pronome lexical (foram 11 ocorrências nas legendas profissionais e apenas 6 nas *fansubs*). Supõe-se que tal cenário se relaciona tanto com recomendações de estilo trazidas nos guias de legendagem, que aconselham o uso da norma culta (cf. subseção 6.2), como ao fato de a legenda ser um gênero particularmente suscetível a críticas, por ser uma tradução que é sempre apresentada em conjunto com o original. Além disso, os resultados levam a crer, também, que a legenda, a partir do conjunto de textos observados aqui, se caracteriza como um gênero muito mais homogêneo do que inicialmente se esperava – aspecto considerado na descrição empreendida na subseção seguinte –, possivelmente em decorrência da padronização que atingiu o universo das *fansubs* nas últimas décadas e as aproximou das produções profissionais (cf. subseção 4.2 e subseção 6.2).

6.2 A legenda audiovisual como um gênero

Conforme delineado na Introdução desta dissertação (cf. seção 1), outro objetivo da pesquisa aqui apresentada consistiu na descrição da legenda audiovisual a partir do conceito de gênero textual-discursivo. É dessa caracterização que se ocupa esta subseção, que traz, inicialmente, uma apreciação geral da legenda, baseada nas considerações acerca do gênero traçadas na subseção 2.2.

Para iniciar essa discussão, é primeiramente importante reconhecer que, assim como no caso de outros gêneros, a legenda se caracteriza como um evento comunicativo. Com efeito, na medida em que as legendas auxiliam na compreensão das produções audiovisuais em língua estrangeira que elas acompanham, elas passam a compor as estratégias de comunicação da obra, garantindo um maior acesso ao conteúdo legendado. Nesse sentido, a língua mobilizada e trabalhada pelas legendas passa a fazer parte do evento no qual a produção é assistida, compondo um todo essencialmente multimodal.

Como todo evento comunicativo, a legenda se insere em diversos domínios discursivos, a depender do tipo de produção que ela acompanha. No caso das legendas de séries e filmes ficcionais, é possível encaixá-las no domínio ficcional, proposto por Marcuschi (2008), uma

vez que elas são preparadas, em geral, no âmbito da criação de conteúdos para o entretenimento. Por outro lado, legendas que acompanham produções documentais, por exemplo, apesar de também serem roteirizadas em algum nível, não poderiam compor o mesmo domínio, por apresentarem elementos factuais e terem como objetivo não só entreter, mas, também, informar, de modo que estariam mais próximas dos domínios jornalístico ou instrucional. Assim, a legenda como um todo se mostra, mais uma vez, um gênero híbrido, que não se encaixa perfeitamente em um único domínio, podendo flutuar entre eles na medida em que reflete as características da produção que ela traduz.

Nesses domínios discursivos, a legenda possui contextos de produção e de recepção específicos, que também variam de acordo com o tipo de legenda. A elaboração de legendas feitas por profissionais se dá no seio da indústria do entretenimento, podendo envolver estúdios, distribuidoras e empresas de tradução em sua concepção. Essas legendas são recebidas pelos telespectadores em prazos relativamente dilatados, chegando ao público de canais de televisão a cabo, serviços de *streaming* e cinemas. Em contrapartida, as legendas feitas por fãs envolvem grupos de legendagem *online*, nos quais fãs de uma produção se reúnem para legendá-la de forma voluntária, dividindo-se entre as funções de tradutor, sincronizador e revisor. Nesse caso, o público consiste em comunidades de fãs que consomem os filmes e séries rotineiramente, sendo essencial que sua distribuição ocorra rapidamente por meio da internet.

Além disso, as legendas, como todo gênero, também possuem conteúdo temático, estilo e construção composicionais próprios. Pensando no primeiro elemento, cada legenda possui uma temática específica, que pode se alterar juntamente às cenas e segue o roteiro da série ou filme, sendo essencial para as escolhas linguísticas a serem feitas em sua elaboração. Já o estilo da legenda é definido a partir da relação entre o original e a tradução, das recomendações estilísticas elencadas pelos guias de legendagem, e pelo nível de familiaridade entre o legendista e o público. De maneira geral, como explicitado na subseção 4.2.1, a língua empregada nas legendas tem um estilo mais curto, com omissões recorrentes, sendo elaborada com a intenção de emular a fala original, que é tida como natural. Por fim, quanto à construção composicional, a legenda possui uma estrutura bastante característica, com quebras de linha determinadas, marcações de diálogo sobreposto, formatações distintas para textos que aparecem em tela, controle do número máximo de CPL, CPS, duração entre outros traços ligados aos parâmetros técnicos que guiam sua criação.

Todas essas características operam conjuntamente na composição da legenda como um gênero profundamente multimodal, sendo necessário voltar um olhar mais apurado para os seus

aspectos estruturais, estilísticos e situacionais, a fim de compreender como eles podem afetar o modo como a língua se manifesta nesse gênero.

6.2.1 O gênero legenda: aspectos estruturais

Na subseção 4.2.1, dedicada à apresentação da legenda a partir de referências do campo da TAV, foram elencados parâmetros técnicos importantes para esse gênero e que impactam diretamente em seu estilo: o número e a posição das linhas, a duração e os limites de CPL e de CPS. Daqui em diante, cada um desses parâmetros será retomado, buscando contrastar o que de fato foi encontrado nas legendas profissionais e nas *fansubs* do *corpus* com as recomendações de seus guias de legendagem.

Começando pelo número e pela posição das linhas, ambos os guias consultados – o do *Amazon Prime Video* e o da equipe InSUBs – definem um máximo de duas linhas por legenda, devendo-se utilizar apenas uma linha sempre que possível. Tal regra é categoricamente respeitada, não tendo sido verificada nenhuma legenda com mais de duas linhas em todo o *corpus*. Da mesma maneira, todas as legendas consideradas estavam centralizadas na parte inferior da tela, mesmo que, para as *fansubs*, não houvesse uma orientação explícita quanto à posição adequada. Trata-se, portanto, de dois parâmetros bastante estáveis, que funcionam na estruturação da legenda como um todo.

Passando à duração, são poucas as diferenças entre as recomendações feitas para os dois tipos de legenda. De acordo com o guia do *Amazon Prime Video*, a duração mínima de uma legenda deve ser de 1 segundo e a máxima, de 7 segundos; já o guia da InSUBs indica uma duração mínima de 1,4 segundos, não determinando, porém, uma duração máxima para suas legendas. No geral, esses limites foram respeitados nas legendas que compõem as duas amostras desta pesquisa, havendo, entretanto, exceções no que diz respeito à duração mínima das legendas. Com efeito, dentre as profissionais, as legendas mais curtas observadas possuíam uma duração de 0,834 segundo e, dentre as *fansubs*, de 0,986 segundo, ou seja, em ambas as amostras foram encontradas legendas mais curtas do que o recomendado.

Contudo, é importante mencionar que, aparentemente, a utilização de legendas com duração menor do que o tempo estabelecido foi mais recorrente nas legendas profissionais do que nas *fansubs*. Para as legendas do primeiro grupo, 7 dos 9 episódios considerados apresentaram legendas que fugiram ao padrão. Já nas legendas feitas por fãs, esse número caiu para 4, com as legendas para os outros 5 episódios cravando 1,4 segundos como duração mínima, o que sugere uma tentativa mais enfática de se manter dentro do padrão definido.

Quanto ao número de CPL, são bastante diversas as orientações verificadas nos guias de legendagem examinados. Enquanto o *Amazon Prime Video* define um máximo de 42 CPL para as legendas de sua plataforma, esse limite é de apenas 35 CPL para a equipe InSUBs. Tal diferença pode ser resultado de uma expectativa distinta quanto ao suporte no qual a legenda será projetada: enquanto as legendas da plataforma de *streaming* são feitas, a princípio, para exibição em televisões, aquelas feitas por fãs são consumidas principalmente por meio de computadores, de forma que há menos espaço na tela para comportar as legendas.

De todo modo, em ambas as amostras de legendas, o número máximo de CPL verificado foi de 35, ou seja, não foram encontradas legendas que desobedeciam ao limite imposto pelos guias de legendagem. Assim, o número de CPL permitido nas legendas também se projeta como um parâmetro técnico relevante, sendo essa uma restrição característica do gênero, que, por ser amplamente obedecida, traz consequências diretas para o estilo empregado nas legendas.

Por fim, o número de CPS se configura como outro parâmetro cujas recomendações parecem se aproximar. No guia de legendagem do *Amazon Prime Video*, é definida uma velocidade de leitura ideal de 17 CPS, podendo o máximo de CPS chegar a 21 caso necessário. Por outro lado, as orientações da equipe InSUBs não traçam considerações acerca desse parâmetro, presumindo que o controle da duração das legendas e do número de CPL seria suficiente para garantir que o espectador tenha tempo suficiente para ler confortavelmente a legenda. Apesar disso, ao examinar a versão do programa de legendagem *Subtitle Workshop* (Spiridonov, 2013) configurada com os padrões da equipe e disponibilizada para seus membros, é possível observar uma tolerância de 20 CPS, posto que, ultrapassando esse valor, o programa passa a sinalizar um erro na legenda.

Ao voltar o olhar para as legendas em si, porém, verificou-se um máximo de 34 CPS nas legendas profissionais e de 28 para as *fansubs*, indicando que, por vezes, os tradutores ultrapassam os limites estabelecidos, principalmente ao legendar falas nas quais as personagens falam rapidamente. Todavia, as legendas que não se adequam às orientações estabelecidas são minoria diante do todo, no qual prevalecem legendas mais curtas e com número de CPS bem abaixo do que o determinado. Imagina-se, portanto, que, assim como no caso dos CPL, essa restrição das legendas tenha um impacto sobre a língua nelas utilizada, afetando as escolhas do tradutor.

A Tabela 6 sumariza as recomendações para cada um dos parâmetros técnicos apresentados tanto para legendas profissionais como para *fansubs*, ao lado do que foi constatado por meio do contato com as legendas que compõem o *corpus* desta pesquisa.

Tabela 6 – Recomendações dos guias de legendagem para os parâmetros técnicos e resultados da análise das legendas

Parâmetro	Orientação do guia de legendagem		Resultado da análise	
	<i>Amazon Prime Video</i>	Equipe InSUBs	Profissionais	<i>Fansubs</i>
Nº de linhas	máximo de 2	máximo de 2	máximo de 2	máximo de 2
Posição das linhas	centralizada na parte inferior da tela	-	centralizada na parte inferior da tela	centralizada na parte inferior da tela
Duração mínima	1 segundo	1,4 segundos	0,834 segundo	0,986 segundo
Duração máxima	7 segundos	-	6,5 segundos	6 segundos
CPL	42	35	35	35
CPS	21	20	34	28

Fonte: Elaborada pela autora

Além desses parâmetros, outras orientações dadas pelos guias de legendagem também afetam a dimensão estrutural da legenda, sobretudo no que concernem às estruturas linguísticas empregadas, isto é, ao estilo. Começando pelas legendas profissionais, a primeira indicação relacionada a isso aparece na seção intitulada “Formalidade”, na qual se lê: “Use formas de tratamento formais ou informais com base na relação entre as personagens, e mantenha-as consistentes ao longo da legenda (a não ser que ocorra uma mudança na relação). Se tiver dúvida, use formas formais.” (Amazon, 2019, p. 6, tradução nossa)⁷⁰. Essa indicação dá ao legendista a liberdade de, a partir de sua interpretação da cena e da situação comunicativa representadas, escolher qual seria a forma de tratamento mais adequada, especialmente nos casos em que a língua do original nem sempre utiliza estratégias marcadas pela formalidade. É o que acontece na passagem do inglês para o português, por exemplo, quando, ao se deparar com o pronome “*you*”, o legendista é forçado a escolher se traduz o termo como “você”, “tu”, ou até mesmo “o senhor/a senhora”.

Essa questão é expandida na seção seguinte, voltada à ortografia, em que, além de considerações sobre a escrita das palavras, há, também, as seguintes orientações:

⁷⁰ “Use formal or informal terms of address based on the relationship between characters, and keep them consistent throughout (unless there is a change in the relationship). When in doubt, use formal terms.”

15.3 Tanto os estilos da norma culta como da linguagem coloquial são permitidos, desde que eles se adequem à categoria/gênero do programa. Por exemplo, uma série que se passa em uma escola de ensino médio usaria o estilo coloquial e um documentário histórico usaria a norma culta.

15.4 Segunda pessoa do singular: tanto você como tu são aceitáveis, desde que a frase esteja gramaticalmente correta (Amazon, 2019, p. 7, tradução nossa)⁷¹.

Como pode ser observado, o guia de legendagem do *Amazon Prime Video* permite a utilização da norma culta e/ou da linguagem coloquial, porém sem indicar o que a empresa entende por esses dois termos. Dessa maneira, fica a cargo dos tradutores selecionar qual será a linguagem efetivamente empregada, o que passa tanto por sua compreensão do tom da produção como por suas convicções acerca de língua e adequação. Ao se deparar com uma frase como “*I found him*”, para dar um exemplo relacionado ao fenômeno variável analisado nesta pesquisa, o legendista poderia facilmente traduzir o trecho como “Encontrei ele”, caso julgasse adequado aplicar o pronome lexical, comumente associado a situações de menor monitoração e informalidade, ou “Encontrei-o”, se imaginasse que a conversa deveria seguir um tom mais formal e mais próximo do que define o padrão.

Ademais, é preciso considerar que uma ressalva como “[...] desde que a frase esteja gramaticalmente correta [...]” (Amazon, 2019, p. 7, tradução nossa) pode ser entendida como uma necessidade de se obedecer, sempre, às regras das gramáticas tradicionais. Se um legendista profissional entende gramática como o conjunto de regras trazidas nesses compêndios, essa instrução o levará, naturalmente, a manter o uso do mesmo pronome ao longo do filme ou série, empregando, também, seus respectivos pronomes acusativos e possessivos. Por exemplo, se optasse pelo uso de “você”, uma fala como “*I helped you*” só poderia ser traduzida como “Eu ajudei você” ou “Eu o/a ajudei”, e nunca como “Eu te ajudei”. Assim, pode ser que sejam encontrados mais clíticos associados ao uso do “você” nessas legendas, em uma clara tentativa de seguir os postulados das gramáticas tradicionais, o que não só levaria a um estilo visto como mais formal, mas impactaria nas taxas de aplicação dessa variante.

Ainda sobre o tópico da construção do estilo nas legendas profissionais, outro ponto importante a ser considerado são as instruções quanto a repetições. Para esse tipo de legenda, a orientação apresentada é: “se uma palavra ou expressão é repetida várias vezes em seguida pelo mesmo falante, sincronize a legenda com o áudio, mas traduza apenas uma vez” (Amazon,

⁷¹ “15.3 Both the educated norm and the colloquial language styles are allowed if they match the category/genre of the program. For example, a show set in a high school would use the colloquial style and a historic documentary would use the educated norm.

15.4 Second person singular: Both você and tu are acceptable as long as the sentence is grammatically correct.”

2019, p. 9, tradução nossa)⁷², podendo essa regra ser ignorada quando a repetição for utilizada para ênfase ou para efeitos cômicos. Portanto, por mais que seja necessário emular a fala das personagens, há limites para o que deve ser transposto, especialmente em virtude do tempo e do espaço disponíveis para cada legenda. Como resultado, chega-se a uma legenda marcada pela brevidade, traço reforçado pelo controle do número de CPL e CPS, discutidos acima.

Por outro lado, o guia de legendagem das *fansubs* analisadas, elaborado pela equipe InSUBs, apresenta bem menos considerações sobre estilo ou norma, dedicando-se a apontar estratégias para a elaboração das traduções. De fato, o único momento em que é mencionado qualquer apontamento acerca da língua que deve construir as legendas é no trecho inicial, quando se afirma que a “[...] legendagem é um trabalho minucioso, no qual: ortografia, concordância, o bom português, tempo certo de sincronia, tudo faz a diferença” (InSUBs, [2012?], p. 1)⁷³. Só por essa indicação, já se pode notar que a orientação para as *fansubs* é se aproximar da norma-padrão, posto que, no geral, o “bom português” é visto como aquele que mais se aproxima dos postulados das gramáticas tradicionais.

Tal percepção ganha mais força ao se observar outras considerações trazidas pelo manual, especialmente acerca do que a equipe chama de concordância. Mais à frente, o guia da equipe InSUBs orienta seus legendistas da seguinte maneira: “Cuidado com a concordância. Se está usando ‘você’, não use ‘te’” (InSUBs, [2012?], p. 3), apresentando, em seguida, um exemplo de inadequação, utilizando o pronome átono de segunda pessoa, e outro de uma boa tradução, que emprega o acusativo de terceira pessoa. Desse modo, o documento defende a utilização de uma estrutura mais próxima da norma-padrão em detrimento de uma forma absolutamente comum e que já se tornou parte do vernáculo, unicamente por não ser ela preconizada pelos compêndios gramaticais. Percebe-se, assim, uma clara preocupação com a correção das legendas elaboradas, talvez pelo fato de o emprego do “bom português” ser visto por muitos como um atestado da qualidade de um trabalho, ponto importante para legendistas que fazem seu trabalho de maneira informal.

Além disso, outra questão importante discutida no manual da equipe InSUBs diz respeito à brevidade das legendas. Partindo do acrônimo “KISS”, que significa “*Keep it simple, stupid*”, os legendistas defendem a ideia de que

⁷² “If a word or a phrase is repeated multiple times in a row by the same speaker, time the subtitle to audio, but only translate once.”

⁷³ O ano provável da referência é 2012, pois foi esse o ano em que o guia de legendagem foi disponibilizado na internet, no mesmo endereço em que pode, hoje, ser acessado (<https://www.mediafire.com/?vc3w6864e5xtllw>).

É sempre importante reduzir as falas o quanto puder, sem, é claro, perder o sentido. Acaba sendo um trabalho de bom senso, mas que se justifica, pois quem assiste tem que ter tempo para ver as imagens, e não ficar lendo o tempo todo, cada proposição, cada palavrinha que o texto apresenta (InSUBs, [2012?], p. 1)

Para tanto, a equipe indica uma série de procedimentos que podem ser utilizados na redução das legendas, dentre os quais se destacam: (i) traduzir a fala sem vocativos, desde que o personagem seja conhecido e esteja claro a quem a frase se dirige; (ii) omitir os pronomes “eu” e “você” sempre que possível, caso se saiba a quem se destina a fala; (iii) evitar a repetição de palavras, traduzindo termos repetidos apenas uma vez e alterando a ordem da frase se necessário; (iv) não traduzir expressões e/ou gírias desnecessárias como “*but*”, “*well*” e “*like*”; e (v) ignorar, sempre que possível, as interjeições utilizadas pelas personagens. Mais uma vez, renuncia-se à reprodução completa da fala no original com vistas a facilitar o entendimento do espectador, que é capaz de completar as informações omitidas enquanto assiste à produção audiovisual.

Dessa forma, estrutura composicional e estrutura linguística se misturam na concretização da legenda, criando um todo facilmente identificável, com características e restrições próprias, e que só funciona efetivamente em conjunção com as outras linguagens em jogo durante sua exibição – a imagem e o som.

6.2.2 O gênero legenda: aspectos situacionais

A fim de atingir uma descrição mais ampla e detalhada da legenda audiovisual como um gênero, é imprescindível considerar seus aspectos situacionais, dado que eles se relacionam diretamente com seus contextos de produção e de recepção, e, por conseguinte, afetam a materialidade linguística desses textos. Com isso em mente, esta subseção se ocupa da caracterização da dimensão situacional da legenda, baseando-se, conforme anunciado na subseção 5.6, na proposta de Biber e Conrad (2009) acerca da descrição e comparação de gêneros textuais-discursivos. Daqui em diante, serão retomados os sete fatores elencados pelos autores – participantes, relações entre participantes, canal, condições de produção, cenário, propósito comunicativo e tópico –, demonstrando como eles se aplicam à legenda e se há diferenças entre a sua manifestação em traduções profissionais e em *fansubs*.

Como já foi descrito na subseção 2.2, o primeiro desses fatores, os participantes, diz respeito ao(s) interlocutor(es) envolvido(s) na produção e na recepção de um gênero, podendo-se dividi-los de modo mais geral entre emissores e destinatários. Pensando no emissor, ele pode

ser único, mais de um, institucional ou não identificado, de modo que nem sempre é possível recuperar suas características sociais. Já para o destinatário, é possível que ele seja único, mais de um ou não enumerado, consistindo em outro indivíduo ou no próprio autor do texto. Ainda, é necessário, de acordo com os autores, considerar se há espectadores observando o evento comunicativo, mesmo que não sejam eles os destinatários finais.

No caso das legendas profissionais, produzidas por tradutores que são contratados por empresas ou estúdios, há um único emissor que nem sempre é identificado, já que é usual que o legendista não receba os créditos pela elaboração da legenda publicamente. Em decorrência disso, é difícil recuperar as características sociais desse tradutor, estando elas virtualmente inacessíveis a quem lê a legenda. Quanto ao destinatário dessas legendas, pode-se dizer que ele é não enumerado, posto que não há como saber com exatidão a quantas pessoas a legenda chegará. Apesar disso, tem-se que o público-alvo dessas traduções são os espectadores das produções legendadas, podendo ele consistir em assinantes de televisão a cabo, assinantes de serviços de *streaming* ou pessoas que vão ao cinema, de modo que é possível traçar algum controle sobre quem consome esse material.

De maneira diversa, para as *fansubs* há vários emissores, já que elas são elaboradas por um grupo de fãs legendistas que se distribui entre as funções de tradutores, sincronizadores e revisores das legendas. Novamente, não se pode acessar as características sociais desses legendistas apenas por meio das legendas, uma vez que eles são costumeiramente identificados por *nicknames* (apelidos), sendo necessário outro tipo de contato – por meio das redes sociais, por exemplo – para que se possa saber quem foram as pessoas envolvidas na elaboração de cada tradução. Pelo mesmo motivo que o apontado para as legendas profissionais, o destinatário das *fansubs* também se caracteriza com não enumerado, havendo, porém, uma diferença marcante no que diz respeito ao seu público-alvo geral. Não tendo qualquer ligação com a distribuição comercial das produções audiovisuais que elas acompanham, as *fansubs* são sempre feitas pela e para a comunidade de fãs de algum filme ou série, sendo esse público geralmente mais restrito.

Abordando, agora, as relações entre participantes, Biber e Conrad (2009) definem quatro parâmetros a serem observados: interatividade, papéis sociais, relação pessoal e conhecimento compartilhado. Ao analisá-los, a ideia é avaliar a possibilidade de se estabelecer ou não contato direto entre emissor e destinatário, o *status* relativo ou poder de um participante sobre o outro, o tipo de ligação entre os participantes, e o quanto eles compartilham conhecimento sobre algum tema, considerando sempre que esses aspectos implicam necessidades e escolhas diferentes.

Voltando o olhar à legenda, percebe-se que ela se caracteriza como um gênero de baixa interatividade entre os indivíduos nos contextos de produção e de recepção, já que não é possível que os espectadores respondam à tradução elaborada de maneira imediata. Pensando na questão dos papéis sociais, é apenas por meio do trabalho dos legendistas que é possível ter acesso à produção audiovisual em língua estrangeira, de modo que, comparativamente, os tradutores possuem um *status* relativo mais alto do que o de seu público. Além disso, os emissores e os destinatários são, comumente, estranhos uns aos outros, especialmente no caso das legendas profissionais. Para as *fansubs*, é possível que eles se conheçam, caso tenham uma presença ativa na comunidade de fãs da qual fazem parte. Finalmente, existe entre os participantes um conhecimento compartilhado sobre o filme ou série legendado, que pode assumir níveis distintos. Com efeito, ao passo que o tradutor profissional pode conhecer a produção apenas no momento de realizar seu trabalho, o fã legendista apresenta uma profunda familiaridade com o conteúdo audiovisual a ser legendado, equiparável ao dos fãs que consomem o filme ou série.

Quanto ao canal, Biber e Conrad (2009) destacam, primeiramente, a modalidade por meio da qual um gênero se materializa, isto é, se se trata de um gênero falado, escrito ou sinalizado, a fim de averiguar como isso afeta a elaboração do texto. Em seguida, detalham o meio específico utilizado, podendo ele ser permanente – se for gravado, transcrito ou impresso, por exemplo – ou transitório, como em uma conversa presencial ou pelo telefone, no rádio e na televisão.

Aplicando essa proposta à legenda, tem-se, então, que a modalidade através da qual a legenda se concretiza é a escrita, apesar de ela só tomar forma em conjunção com outras linguagens multimodais como a imagem e o som da produção audiovisual (cf. subseção 4.2.1). Nesse contexto, a legenda se projeta como um gênero de cunho transitório, uma vez que acompanha a troca das cenas da produção audiovisual que acompanha, sendo essencial que o espectador tenha tempo suficiente para realizar a sua leitura sem se perder no enredo do filme ou série.

Passando às circunstâncias de produção, quarto aspecto elencado por Biber e Conrad (2009), é importante lembrar que elas são analisadas a fim de verificar qual foi o planejamento envolvido na concepção de cada gênero. Assim, os autores diferenciam textos elaborados em tempo real, que são mais dinâmicos e pressupõem pouco planejamento, daqueles planejados ou roteirizados, que usualmente passam, ainda, por processos de revisão e edição antes de chegarem aos seus destinatários.

Com base nisso, pode-se concluir que a legenda é um gênero completamente planejado, que começa a ser elaborado após a produção ou exibição de uma produção audiovisual, seja como parte da cadeia de distribuição da obra ou para permitir o acesso de fãs a novos filmes ou episódios de série. Ademais, a legenda implica um trabalho com o roteiro da produção, que serve como ponto de partida das traduções, além de passar por sucessivos processos de revisão e edição antes de atingir o seu destinatário final.

A fim de investigar o cenário em que os gêneros se materializam, Biber e Conrad (2009) apontam, como mencionado anteriormente (cf. subseção 2.2), três questões essenciais. Deve-se, em primeiro lugar, observar se o tempo e o lugar da comunicação são compartilhados pelos participantes para, então, considerá-los de maneira mais detalhada. Sobre o lugar da comunicação, é importante tanto identificar se ele é privado ou público, como se há um cenário específico. Já para tratar do tempo, verifica-se se ele é contemporâneo ou relacionado a um período histórico, com vistas a averiguar como essas especificidades são refletidas na composição do gênero.

Pensando na legenda, percebe-se que os participantes não compartilham o tempo e o lugar da comunicação, já que a elaboração da legenda acontece antes que o espectador assista ao filme ou à série legendada. Ademais, ela tem como lugar da comunicação um cenário bastante peculiar, em que grande parte do trabalho é realizado em um ambiente privado, mas cria-se um produto público, que irá circular por meio da televisão, do cinema ou da internet. Caracteriza-se, além disso, como uma produção contemporânea, que surgiu impulsionada pela amplificação do meio audiovisual e acompanha, atualmente, a produção de novas obras.

O penúltimo aspecto a ser explorado por Biber e Conrad (2009), os propósitos comunicativos, se volta aos motivos pelos quais uma comunicação acontece, estando eles profundamente conectados à função assumida por um gênero em seu contexto de produção. Para os autores, é possível pensar em um propósito geral e em propósitos específicos, que variam amplamente de acordo com cada gênero e podem ser considerados, também, a partir do tipo de informação transmitida. Com vistas a exemplificar essas possibilidades, os teóricos indicam uma lista não exaustiva de propósitos possíveis, reproduzida na Tabela 1, localizada na página 50.

Adaptando a proposta de Biber e Conrad (2009), já que a legenda não apresenta nenhum dos propósitos gerais ou específicos elencados pelos autores, propõe-se que, nesse gênero, o objetivo geral seja dar acesso a uma produção audiovisual em língua estrangeira – de maneira mais rápida possível, no caso das *fansubs* –, e os propósitos específicos se vinculem às funções de entreter ou informar os espectadores. Além disso, a legenda consiste em um gênero factual,

que, no entanto, representa situações comumente fictícias e/ou roteirizadas, deixando espaço para que o tradutor tome decisões de maneira subjetiva.

Finalmente, Biber e Conrad (2009) destacam o tópico do texto como algo que pode afetar a composição do gênero, podendo esse parâmetro também ser separado entre geral e específico. Dentre os tópicos gerais, figuram temas recorrentes de diversos campos de atividade que, naturalmente, englobam assuntos particulares, relacionados às outras especificidades de cada situação comunicativa. Novamente, os autores criam uma relação de tópicos possivelmente abordados, que não engloba, porém, todas as possibilidades temáticas.

Assim, da mesma maneira que para os propósitos comunicativos, propõe-se, aqui, uma adição à lista dos teóricos, com vistas a torná-la ampla o suficiente para abrigar, também, a legenda. Com isso em mente, defende-se que o tópico das legendas gira em torno do enredo das produções audiovisuais que elas acompanham, sendo ele, portanto, tão versátil quanto os próprios filmes e séries legendados. Ademais os tópicos específicos acompanham o desenvolvimento da trama retratada e das personagens ou pessoas em cena, podendo até mesmo se modificar a cada cena.

A Tabela 7 reúne as informações discutidas nesta subseção, apresentando, de maneira mais visual, a aplicação dos conceitos de Biber e Conrad (2009) para a caracterização dos aspectos situacionais de um gênero às legendas.

Tabela 7 – Configuração das características situacionais em legendas profissionais e *fansubs*, a partir da proposta de Biber e Conrad (2009)

Características situacionais	Legendas profissionais	<i>Fansubs</i>
Participantes	Um emissor, nem sempre identificado	Vários emissores, identificados por <i>nicknames</i>
	Destinatário não enumerado	Destinatário não enumerado
Relações entre participantes	Baixa interatividade	Baixa interatividade
	<i>Status</i> mais alto dos legendistas	<i>Status</i> mais alto dos legendistas
	Estranhos uns aos outros	Geralmente estranhos uns aos outros
	Conhecimento compartilhado sobre a produção	Conhecimento compartilhado sobre a produção

Canal	Escrito, em conjunção com outras linguagens	Escrito, em conjunção com outras linguagens
	Transitório	Transitório
Circunstâncias de produção	Planejado	Planejado
	Roteirizado	Roteirizado
	Revisado	Revisado
	Editado	Editado
Cenário	Tempo e lugar não compartilhados	Tempo e lugar não compartilhados
	Ambiente privado, com criação de um produto público	Ambiente privado, com criação de um produto público
	Contemporâneo	Contemporâneo
Propósitos comunicativos	Dar acesso a uma produção em língua estrangeira	Dar acesso mais rápido a uma produção em língua estrangeira
	Entreter	Entreter
	Informar	Informar
Tópico	Moldado pelo enredo da produção	Moldado pelo enredo da produção

Fonte: Elaborada pela autora

7 CONSIDERAÇÕES FINAIS

A presente pesquisa visou a acrescentar uma nova descrição do objeto de estudo (o OD anafórico de terceira pessoa) aos estudos variacionistas sobre o PB, em um gênero textual-discursivo ainda pouco explorado nas investigações sociolinguísticas – empreitada que incluiu o trabalho com amostras-piloto para a escolha da variável dependente/resposta, a descrição da legenda audiovisual como um gênero e discussões acerca das relações entre variação linguística e gêneros textuais-discursivos. Nesse sentido, desenvolveu-se um trabalho descritivo do fenômeno variável em questão, que levou em conta as características da legenda levantadas e exploradas a partir do conceito de gênero.

Embasando-se nos pressupostos teórico-metodológicos da Teoria da Variação e Mudança Linguísticas (Weinreich; Labov; Herzog, 2006[1968]; Labov, 1982, 1994, 2001, 2003, 2008[1972]), em noções sobre gêneros textuais discursivos e suas características estruturais e situacionais (Bakhtin, 2016[1979]; Marcuschi, 2005, 2008, 2010; Biber; Conrad, 2009), e em pesquisas que defendem uma maior aproximação entre os conceitos de variação e de gênero (Biazolli, 2010, 2016; Vieira e Lima, 2019; Biazolli e Berlinck, 2021; Lima, 2022; Vieira, em andamento), observou-se como os acusativos anafóricos de terceira pessoa se manifestaram em legendas audiovisuais a partir de duas amostras: uma de legendas profissionais e outra de *fansubs*. Esperava-se que esses dois tipos de legenda fossem divergir na utilização da língua, dados os seus contextos de produção e de recepção distintos, principalmente no que diz respeito aos tradutores envolvidos, à motivação para a tradução e às condições de trabalho. Os dados analisados foram coleados de legendas para a série *Grey's Anatomy*, extraídas da plataforma de *streaming Amazon Prime Video*, no caso das legendas profissionais, e baixadas do *site Legendas.TV*, para as legendas feitas por fãs, elaboradas pela equipe InSUBs.

Antes da coleta e codificação das ocorrências encontradas nessas duas amostras, foi realizado um trabalho de adequação dos arquivos de legenda para que eles pudessem ser posteriormente submetidos a um programa estatístico, incluindo a remoção das marcações de tempo e das numerações de linha e a adição de códigos para identificar cada personagem. Em seguida, delimitou-se a variável dependente/resposta deste estudo, com base na literatura disponível acerca do fenômeno, a fim de que a coleta pudesse ser direcionada. Foram observados, então, os contextos linguísticos e extralinguísticos que haviam sido apontados como relevantes para a realização do OD anafórico de terceira pessoa, com vistas a incluí-los como variáveis independentes/previsores da pesquisa. Após esse passo, todos os dados foram qualificados para as variáveis escolhidas, quais sejam: natureza, traço semântico e função

sintática do antecedente, forma verbal, estrutura sintática, tópico, relação entre participantes, tipo de legenda, número de CPL e número de CPS. Como era a intenção, focou-se apenas no OD de natureza nominal, sendo que as ocorrências em contexto oracional foram descartadas.

Para o tratamento estatístico de todos os acusativos anafóricos de antecedente nominal encontrados, recorreu-se à plataforma *R* (R Core Team, 2023), através da qual foram realizados testes exatos de Fisher e regressões logísticas multinomiais com a intenção de averiguar a manifestação do objeto de estudo. Inicialmente, as quatro variantes do fenômeno foram consideradas nas projeções geradas; em um segundo momento, porém, o pronome lexical foi excluído da análise multivariada, devido ao seu baixo número de ocorrências. Os dados das duas amostras – a de legendas profissionais e a de legendas amadoras – foram explorados de maneira conjunta, uma vez que não houve diferença significativa entre eles.

De maneira geral, o clítico acusativo foi a variante mais empregada no *corpus* de legendas audiovisuais elaborado para esta pesquisa, tendo sido a forma preferida em virtualmente todos os contextos analisados. Das quatro variáveis linguísticas consideradas, três foram apontadas pelo modelo de regressão logística como correlacionadas à realização do OD anafórico nas legendas, sendo elas o traço semântico do antecedente, a forma verbal e a estrutura sintática. Para o traço semântico do antecedente, como já havia sido indicado por outras pesquisas acerca do acusativo anafórico de terceira pessoa, o objeto nulo foi favorecido em retomadas de antecedentes não animados, representando mais do que a metade de todas as ocorrências, ao passo que o clítico foi a variante mais utilizada quando se fazia referência a uma entidade animada. Também em consonância com estudos anteriores, observou-se que formas verbais com infinitivo favoreceram o uso do pronome acusativo em detrimento das outras variantes, que, por sua vez, foram mais presentes com outras formas finitas. Já no caso da estrutura sintática, também correlacionada à ocorrência do fenômeno no *corpus*, a realização do SN em sentenças de configuração S V OD + complementos foi desfavorecida, sendo que a forma dominante foi, novamente, o clítico. Apesar de a função do antecedente não ter se correlacionado de maneira significativa com a VR, constatou-se um comportamento interessante do fenômeno em relação aos antecedentes cuja função sintática não poderia ser identificada, que pareceu ser um caso intermediário entre os dois níveis da variável. Nesse contexto, houve menos clíticos do que quando o antecedente exercia função diferente, mas mais do que quando a função era igual, e mais objetos nulos do que com antecedentes de função diferente e menos do que quando o antecedente também era um OD.

Quanto às variáveis extralinguísticas, quase todos os preditores selecionados demonstraram correlação estatisticamente significativa para a expressão do acusativo anafórico

de terceira pessoa. No caso do tópico, constatou-se que a categoria vazia foi favorecida em relação ao clítico em interações acerca de aspectos profissionais, representando mais do que um terço do total de dados. Por outro lado, conversas pessoais contaram com uma presença maior do clítico, fato que foi contrário às expectativas na medida em que não se esperava encontrar taxas de utilização tão grandes da variante padrão em diálogos que, em teoria, deveriam ser menos monitorados. Passando à relação entre os participantes, evidenciou-se que, seguindo a tendência geral do *corpus*, o clítico acusativo foi a estratégia mais empregada tanto no contexto de relações simétricas como nas assimétricas. Em diálogos nos quais havia diferenças de *status* relativo/poder entre os personagens, porém, o pronome acusativo foi favorecido de maneira significativa, enquanto as outras duas variantes foram desfavorecidas.

Já no que diz respeito às variáveis relacionadas diretamente ao gênero textual-discursivo abordado nesta dissertação, tanto o número de CPL como o de CPS se correlacionaram com a realização do OD anafórico nas legendas. Nesse panorama, verificou-se que as chances de uma categoria vazia ser utilizada ao invés de um clítico acusativo diminuíram a cada caractere a mais utilizado nas legendas, enquanto, no caso da restrição temporal, a probabilidade de se encontrar um SN cresceu conforme aumentou o número de CPS. Contudo, a grande maioria dos dados foi retirada de legendas que se afastavam consideravelmente do máximo de caracteres para esses dois parâmetros técnicos, de modo que talvez não tenha sido possível aferir o que de fato acontece nos casos em que as legendas estão próximas do limite de caracteres. Além disso, os próprios processos de adaptação aplicados durante a elaboração das legendas resultam naturalmente em legendas mais curtas, o que pode ajudar a justificar, também, a concentração dos dados em contextos com uso menor de caracteres.

Por fim, os resultados obtidos para a variável tipo de legenda revelaram que, contrariamente à hipótese inicial, não foi observada uma distinção significativa entre a expressão do fenômeno em legendas profissionais e a de *fansubs*. Pelo contrário, tanto as análises das frequências e proporções como as estimativas do modelo de regressão apontaram para uma similaridade mais forte do que o previsto entre os dois tipos de legenda, o que suscitou discussões interessantes para o trabalho ao se tentar compreender esse resultado. Como para as outras variáveis, a preferência pela variante do clítico acusativo foi consistente em ambos os conjuntos de dados, representando mais da metade das ocorrências nos dois contextos, de maneira que ambas as legendas demonstraram certa adesão à norma-padrão, com poucas ocorrências de formas estigmatizadas.

Conjectura-se que esses resultados estejam diretamente conectados às recomendações de estilo presentes nos guias de legendagem e ao fato de a legenda ser alvo comum de críticas,

uma vez que é possível que, na tentativa de atender não só às exigências dos manuais como às expectativas do público, o tradutor conscientemente escolha empregar formas mais próximas do que é tido como correto. Ademais, considera-se que a proximidade entre os dois tipos de legenda pode estar relacionada, também, a outros dois fatores: (i) ao fato de a legenda não conseguir se descolar completamente da língua escrita através da qual toma forma, mesmo que exista uma tentativa de emular a fala natural, e (ii) ao movimento generalizado de padronização das fansubs ao longo dos anos desde seu surgimento, baseado nas orientações técnicas elaboradas para legendas profissionais. Assim, este estudo apontou para uma homogeneidade surpreendente entre as legendas profissionais e as fansubs, o que trouxe consequências importantes para as decisões tomadas no momento de empreender a descrição da legenda como um gênero, buscando-se observar se as semelhanças entre os dois tipos de legenda seriam mais comuns do que as diferenças.

Diante desses resultados, percebeu-se que a realização do OD anafórico apresenta um comportamento bastante particular nas legendas quando comparada a outros gêneros para os quais já há uma descrição no PB. Com relação a dados da oralidade, em que as proporções de clíticos não passam de 7% (Duarte, 1986; Marafoni, 2004; Arruda, 2012; Marques de Sousa, 2021), essa variante se mostrou muito mais expressiva, aproximando-se das porcentagens encontradas para gêneros jornalísticos mais monitorados (Freire, 2005; Lima, 2022). Para o objeto nulo, também se observou uma distribuição mais próxima daquela constatada em textos do domínio jornalístico (Freire, 2005; Lima, 2022), que em geral apresenta proporções muito menores do que aquelas aferidas na fala, em que é essa a variante predominante (Duarte, 1986; Marafoni, 2004; Arruda, 2012; Marques de Sousa, 2021). Dessa forma, no que toca ao emprego do OD anafórico de terceira pessoa, a legenda parece se assemelhar mais aos gêneros escritos do que à fala, o que a coloca em um lugar híbrido, posto que, ao menos em tese, ela deveria ser um gênero pensado para emular a fala.

Além disso, esta pesquisa trouxe *insights* interessantes sobre a variação estilística, na medida em que buscou explorar o comportamento de um fenômeno variável em um gênero no qual o estilo é determinado por uma combinação de fatores. De fato, no caso das legendas, há, ao mesmo tempo, ao menos três dimensões em jogo: a das recomendações do guia de legendagem sobre estilo, a da interpretação do legendista em face dessas orientações e a das situações comunicativas apresentadas no conteúdo a ser legendado. Ao se deparar com uma instrução de empregar a norma culta – colocada pelos guias como oposto de linguagem coloquial – caso ela seja adequada ao contexto, um legendista deve mobilizar tanto o que entende por esse termo como seu conhecimento de mundo para decidir o quão formal deve ser

a linguagem em uma cena, tendo sempre em mente que o destinatário final da tradução é o público. Nesse sentido, analisar a variação nas legendas permite que sejam discutidas questões complexas sobre esse tema ao tentar identificar os fatores que levam a uma escolha linguística ou outra, de modo que esse gênero se projeta como um fértil material para análises desse tipo.

Considerando todas as intersecções apontadas entre o fenômeno variável em estudo nesta dissertação e o gênero textual-discursivo no qual ele ocorre, pode-se dizer, também, que a pesquisa realizada contribui para a discussão de Biazolli e Berlinck (2021), uma vez que busca aplicar os postulados das autoras ao seu desenvolvimento. Com efeito, houve o cuidado de considerar o gênero legenda em todos os passos da pesquisa: na elaboração do *corpus*, pensando na possibilidade de variação entre os dois tipos de legenda; na definição do fenômeno a ser investigado, a partir do que mais chamou a atenção a partir do contato com o *corpus*; na definição e codificação das variáveis, com a tentativa de se transformar características estruturais e situacionais do gênero em variáveis a fim de medir os efeitos do gênero sobre a variação; e na análise e interpretação dos resultados, ao buscar identificar quais especificidades do gênero podem ser determinantes no emprego de uma forma ou outra. Esse esforço acabou por gerar resultados contundentes, demonstrando não só as vantagens de se empregar essa abordagem na análise da variação linguística, mas confirmando a realidade de que a variação está inescapavelmente ligada aos gêneros na medida em que “[...] a língua passa a integrar a vida através de enunciados concretos (que a realizam); [e] é igualmente através de enunciados concretos que a vida entra na língua” (Bakhtin, 2016[1979], p. 17).

Esta dissertação também traz contribuições para a TAV, ao apresentar uma descrição detalhada da legenda audiovisual a partir de textos dessa área de estudo, incluindo considerações sobre a tradução como um todo, sobre natureza da tradução de textos audiovisuais e sobre as particularidades da legenda (cf. seção 4). Ademais, ela permite que interessados na TAV adquiram uma perspectiva distinta daquela comumente trazida em trabalhos sobre tradução, levando-os a enxergar a legenda como um objeto muito mais complexo do que uma mera transposição linguística de um idioma para o outro. Ao comprovar a realidade da variação linguística nas legendas, esta pesquisa dá destaque à autoria do legendista sobre suas traduções, além de apresentá-las como um gênero textual-discursivo estruturado.

Para além de todo o trabalho com a variação linguística que envolve o OD anafórico de terceira pessoa, este estudo trouxe, ainda, uma descrição do gênero textual-discursivo legenda, a partir de suas características estruturais e situacionais (cf. subseção 6.2), a fim de fortalecer o trabalho com a interface entre variação e gênero e contribuir para o preenchimento de uma

lacuna na literatura correspondente à descrição desse gênero. Articulando os pressupostos teóricos discutidos na subseção 2.2, demonstrou-se, primeiramente, como a legenda cumpria aos requisitos básicos para que ela fosse considerada um gênero, identificando-a como um gênero essencialmente multimodal e relativamente híbrido, uma vez que é de concepção oral, por se propor a emular a fala natural, mas se materializa por meio da escrita. Em seguida, com base no que já havia sido descrito pela TAV (cf. subseção 4.2) e no que era determinado pelos guias de legendagem referentes às legendas do *corpus*, investiu-se na caracterização estrutural da legenda, buscando averiguar, além dos elementos básicos de sua composição, quais aspectos são definidores para sua estrutura linguística. Finalmente, efetuou-se um detalhamento das características situacionais da legenda, refinando a compreensão de seus contextos de produção e de recepção a partir da proposta de Biber e Conrad (2009).

Dessa maneira, outro campo que também pode se beneficiar das discussões trazidas por esta pesquisa é o do estudo dos gêneros textuais-discursivos, já que uma descrição tal como a que foi empreendida pode apontar caminhos para a caracterização de novos gêneros. Além disso, o detalhamento de um gênero audiovisual, de cunho essencialmente híbrido como a legenda, levanta questões interessantes para a área acerca dos limites entre fala e escrita, da multimodalidade, da relação entre gêneros emergentes e a mídia, e dos novos contextos de produção e de recepção surgidos ao longo dos anos.

Finalmente, vale traçar, também, alguns comentários sobre as limitações deste trabalho. Primeiramente, deve-se apontar para a escolha de uma série específica – *Grey's Anatomy* – como uma restrição, posto que, apesar de a série representar diversas interações pessoais e profissionais, elas ainda são, do ponto de vista temático, razoavelmente semelhantes, voltando-se para os pacientes que os médicos atendem e os problemas de relacionamento amoroso ou familiar que afligem as personagens. Dessa forma, seria interessante trabalhar com produções que permitiriam uma variedade temática maior e/ou com mais de um tipo de conteúdo visual. Ademais, é importante reconhecer que a pesquisa trabalhou com apenas um tipo de dado – ainda que esse tipo seja subexplorado pela Sociolinguística –, de modo que seria interessante contrastá-lo com dados advindos de outras fontes. Por fim, posto que não era seu objetivo avaliar essas questões, este tudo tratou apenas superficialmente da opcionalidade e da influência do decalque linguístico nas traduções, sendo necessário voltar um olhar mais refinado para essa dimensão.

Como encaminhamentos para pesquisas futuras, pode-se elencar, portanto, algumas possibilidades: (i) a descrição do OD anafórico de terceira pessoa em legendas de outras produções audiovisuais, proporcionando uma gama maior de temas a serem considerados; (ii)

o refinamento das análises por meio da identificação dos contextos que não apresentam opcionalidade para todas as variantes do acusativo anafórico de terceira pessoa, e do controle do decalque entre a língua original e o português; (iii) a investigação de outros fenômenos variáveis nesse mesmo material, a fim de observar se seu comportamento se aproxima daquele encontrado para a expressão do OD anafórico de terceira pessoa; (iv) o trabalho com outros tipos de dados, com vistas a se obter uma compreensão mais ampla do fenômeno sob estudo; e (v) a comparação aprofundada da realização do acusativo anafórico nas legendas e em outros gêneros.

Diante das considerações apresentadas nesta seção, conclui-se que esta pesquisa atingiu os seus objetivos iniciais, realizando a descrição e análise do OD anafórico de terceira pessoa em legendas audiovisuais e apresentando uma caracterização da legenda a partir do conceito de gênero textual-discursivo. Nesse sentido, contribuiu não só para a pesquisa sociolinguística desenvolvida no Brasil, mas, também, para os estudos acerca de gênero e para a tradução, firmando-se como um estudo tão múltiplo quanto a legenda.

REFERÊNCIAS

- AEGISUB. Version 3.2.2. [S.I.]: Aegisub Team, 2014. Disponível em <<https://www.baixaki.com.br/download/aegisub.htm>>. Acesso em: 08 jan. 2023.
- ALVES FILHO, F. **Gêneros jornalísticos: notícias e cartas de leitor no Ensino Fundamental**. São Paulo: Cortez, 2011.
- AMAZON Confidential. **Brazilian Portuguese (Brazil) Timed Text Style Guide**, [s.l.], 2019. Disponível em: <https://videocentral.amazon.com/home/help?topicId=GBKB422Q9GYC7DWE>. Acesso em: 8 fev. 2024.
- ARRUDA, N. C. **A realização do objeto direto no português brasileiro culto falado: um estudo sincrônico**. 2006. 201f. Dissertação (Mestrado em Linguística e Língua Portuguesa) – Faculdade de Ciências e Letras, Universidade Estadual Paulista, Araraquara, 2006.
- ARRUDA, N. C. **A realização do objeto direto anafórico em línguas românicas: um estudo sincrônico no português e no espanhol**. 2012. 167 f. Tese (Doutorado em Linguística e Língua Portuguesa) – Faculdade de Ciências e Letras, Universidade Estadual Paulista, Araraquara, 2012.
- AVERBUG, M. C. **Objeto direto anafórico e sujeito pronominal na escrita de estudantes**. 2000. 155 f. Dissertação (Mestrado em Língua Portuguesa) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2000.
- BAKHTIN, M. **Os gêneros do discurso**. 1. ed. São Paulo: Editora 34, 2016[1979].
- BANDEIRA, A. P. **Don't tell me what I can't do!:** as práticas de consumo e participação dos fãs de Lost. 2009. 135 f. Dissertação (Mestrado em Comunicação Social) – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2009.
- BECHARA, E. **Moderna Gramática Portuguesa**. 37. ed. Rio de Janeiro: Nova Fronteira/Lucerna, 2009[1999].
- BIAZOLLI, C. C. **Clíticos pronominais no português de São Paulo: 1880 a 1920 - uma análise sócio-histórico-linguística**. 2010. 230 f. Dissertação (Mestrado em Linguística e Língua Portuguesa) – Faculdade de Ciências e Letras, Universidade Estadual Paulista, Araraquara, 2010.
- BIAZOLLI, C. C. **Posição de clíticos pronominais em duas variedades do português: inter-relações de estilo, gênero, modalidade e norma**. 2016. 381 f. Tese (Doutorado em Linguística e Língua Portuguesa) – Faculdade de Ciências e Letras, Universidade Estadual Paulista, Araraquara, 2016.
- BIAZOLLI, C. C.; BERLINCK, R. de A. Por que investigar processos de variação e mudança linguísticas por meio de gêneros textuais-discursivos? In: BIAZOLLI, C. C.; BERLINCK, R. de A. (org.). **Gêneros textuais-discursivos no estudo de processos de variação e mudança**. Campinas: Pontes Editores, 2021. p. 13-38.

- BLAZOLLI, C. C.; SENE, M. G. de. ; AZEVEDO, L. O.; ZAMBRANO, P. C.; ALMEIDA, M. A. de. Gêneros textuais-discursivos do Jornal Folha de S.Paulo: da construção da amostra a resultados preliminares. *In: VIEIRA, S. R. et al. Variação linguística, ensino e interfaces: resultados e propostas.* São Paulo: Pimenta Cultura, 2023. p. 60-86.
- BIBER, D. **Variation across speech and writing.** Cambridge: Cambridge University Press, 1988.
- BIBER, D; CONRAD, S. **Register, genre and style.** Cambridge: Cambridge University Press, 2009.
- BOITO, F. S.; CAETANO, M. O. Humor translation in the translation process for the subtitles of the Brazilian TV series "A Diarista". **Mutatis Mutandis**, Medellín, v. 11, n. 1, p. 126-144, 2018.
- BOLD, B. The power of fan communities an overview of fansubbing in Brazil. **Tradução em Revista**, Rio de Janeiro, v. 11, n. 2, p. 1-19, 2011.
- BORTONI-RICARDO, S. M. O Português brasileiro. *In: BORTONI-RICARDO, S. M. Educação em língua materna: a sociolinguística na sala de aula.* São Paulo: Parábola, 2004, p. 51-70.
- BORTONI-RICARDO, S. M. Um modelo para análise sociolinguística do português brasileiro. *In: BORTONI-RICARDO, S. M. Nós chegemos na escola, e agora?* Sociolinguística e educação. São Paulo: Parábola, 2005, p. 45-52.
- BORTONI-RICARDO, S. M. Um modelo para a análise sociolinguística do português do Brasil. *In: BAGNO, M. (org.). Linguística da norma.* 3. ed. São Paulo: Edições Loyola, 2012. p. 301-316.
- BRAGANÇA, M. L. L. **A gramaticalização do verbo IR e a variação de formas para expressar o futuro do presente:** uma fotografia capixaba. 2008. 145 f. Dissertação (Mestrado em Estudos Linguísticos – Universidade Federal do Espírito Santo, Vitória, 2008.
- BROWN, P.; LEVINSON, S. C. **Politeness: some universals in language usage.** Cambridge: Cambridge University Press, 1987.
- BROWN, R.; GILMAN, A.; The pronouns of power and solidarity. **Style in language.** Cambridge, v. 9, n. 5, p. 253-276, 1960.
- BUGOCKI, L. Amateur Subtitling on the Internet. *In: DÍAZ-CINTAS, J.; ANDERMAN, G. (org). Audiovisual translation: language transfer on the screen.* Londres: Palgrave Macmillan, 2009. p. 49-57.
- CALAZANS, R. Mercado começa a ver com bons olhos equipes de fãs que legendam séries na internet. **O Globo**, Rio de Janeiro, 10 mai. 2010. Disponível em: <http://oglobo.globo.com/cultural/mat/2010/05/07/mmercado-comeca-ver-com-bons-olhos-equipes-de-fas-que-legendam-series-na-internet-916531486.asp>. Acesso em: 8 fev. 2024.
- CÂMARA JR, J. M. **Princípios de linguística geral.** Rio de Janeiro: Acadêmica, 1972.

CÂMARA JR, J. M. **História e estrutura da língua portuguesa**. 2. ed. Rio de Janeiro: Padrão, 1976.

CÂMARA JR, J. M. **Estrutura da língua portuguesa**. 21. ed. Petrópolis: Editora Vozes, 1992[1970].

CAMPOS, G. **Como fazer tradução**. Petrópolis: Editora Vozes, 1988.

CAROLL, M; IVARSSON, J. **Code of Good Subtitling Practice**. 1998. Disponível em: <https://www.esist.org/wp-content/uploads/2016/06/Code-of-Good-Subtitling-Practice.PDF.pdf>. Acesso em: 07 fev. 2024.

CASTILHO, A. T. de. **Nova gramática do português brasileiro**. 1. ed. São Paulo: Contexto, 2020[2010].

CHAFE, W. Integration and Involvement in Speaking, Writing, and Oral Literature. *In*: TANNEN, D. (org.). **Spoken and Written Language: Exploring Orality and Literacy**. Norwood, NJ: Ablex, 1982. p. 35-53.

CHAFE, W. Linguistic difference produced by differences between speaking and writing. *In*: OLSON, D. R.; TORRANCE, N.; HILLYARD, A. (org.). **Literacy, Language and Learning**. Cambridge: Cambridge University Press, 1985. p. 105-123.

COELHO, I. L.; GÖRSKI, E. M.; SOUZA, C. M. N. de.; MAY, G. H. **Para conhecer sociolinguística**. São Paulo: Contexto, 2015.

COSERIU, E. Los conceptos de “dialecto”, “nivel” y “estilo de lengua” y el sentido propio de la dialectología, **Linguística Española Actual**, Madrid, v. 3, p. 1-32, 1981.

COSTA, S. **O (não) preenchimento do objeto anafórico na língua portuguesa: análise diacrônica do PB e do PE dos séculos XIX e XX**. 2011. 268 f. Tese (Doutorado em Linguística) – Centro de Comunicação e Expressão, Universidade Federal de Santa Catarina, Florianópolis, 2011.

COSTA, T. M. **Um estudo diacrônico das variadas realizações do objeto direto anafórico na imprensa baiana dos séculos XIX e XX**. 2012. 251 f. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, 2012.

CUNHA, C; CINTRA, L. **Nova gramática do português contemporâneo**. 7. ed. Rio de Janeiro: Lexikon, 2017[1985].

CYRINO, S. M. L. **O objeto nulo no português do Brasil: uma mudança paramétrica?**. Campinas: Unicamp, 1990a. (Manuscrito).

CYRINO, S. M. L. **O objeto nulo no português do Brasil: uma investigação diacrônica**. Campinas: Unicamp, 1990b. (Manuscrito).

CYRINO, S. M. L. **O objeto nulo no português brasileiro: um estudo sintático-diacrônico**. 1994. 229 f. Tese (Doutorado em Ciências) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 1994.

CYRINO, S. M. L. O objeto nulo. *In*: CYRINO, S. M. L.; TORRES-MORAES, M. A. T. (org.) **Mudança Sintática do Português Brasileiro**: perspectiva gerativista. 1. ed. São Paulo: Contexto, 2018a.

CYRINO, S. M. L. Observações sobre a mudança diacrônica no português do Brasil: objeto nulo e clíticos. *In*: ROBERTS, I; KATO, M. (org.). **Português Brasileiro**: uma viagem diacrônica. 3. ed. São Paulo: Contexto, 2018b[1993]. p. 129-142.

CYRINO, S. M. L.; DUARTE, M. E. L.; KATO, M. A. Visible subjects and invisible clitics in brazilian portuguese. *In*: KATO, M. A.; NEGRÃO, E. V. (org.) **Brazilian portuguese and the null subject parameter**. Madri: Iberoamericana; Frankfurt am Main: Verveurt, 2000, p. 55-73.

DÍAZ-CINTAS, J. **La traducción audiovisual**: El subtitulado. Salamanca: Almar, 2001.

DÍAZ-CINTAS, J. Back to the Future in Subtitling. **MuTra 2005 – Challenges of Multidimensional Translation**: Conference Proceedings Saarbrücken, p. 1-17, 2005.

DÍAZ-CINTAS, J. Subtitling. *In*: GAMBIER, Y.; DOORSLAER, L. van. (org.). **Handbook of Translation Studies**. Filadélfia: John Benjamins, 2010. p. 344-349.

DÍAZ-CINTAS, J.; REMAEL, A. **Audiovisual Translation**: Subtitling. Abingdon: Routledge, 2014[2007].

DÍAZ-CINTAS, J.; SÁNCHEZ, P. M. Fansubs: Audiovisual Translation in an Amateur Environment. **Journal of Specialised Translation**, Londres, n. 6, p. 37-52, 2006.

DIONISIO, A. P. Multimodalidade discursiva na atividade oral e escrita. *In*: MARCUSCHI, L. A; DIONISIO, A. P. (org.). **Fala e escrita**. Belo Horizonte: Autêntica, 2007. p. 177-196.

DUARTE, M. E. L. **Variação e sintaxe**: clítico acusativo, pronome lexical e categoria vazia no Português do Brasil. 1986. 73 f. Dissertação (Mestrado em Ciências – Linguística aplicada ao ensino de línguas) – Pontifícia Universidade Católica de São Paulo, São Paulo, 1986.

DUARTE, M. E. L. Sociolinguística “Paramétrica”. *In*: MOLLICA, M. C.; FERRAREZI JR, C. (org.). **Sociolinguística, sociolinguísticas**. São Paulo: Contexto. 2016. p. 33-44.

ECKERT, P. **Variation, convention, and social meaning**. Plenary talking. Annual meeting of the Linguistic variation. São Francisco, 2005.

ESPERANDIO, I. B. **Legendas de seriados de tema sobrenatural**: uma abordagem terminológica para tradutores. 2015. 229 f. Dissertação (Mestrado em Teorias Linguísticas do Léxico) – Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2015.

FINALE B. Intérprete: Vários artistas. Compositor: Jonathan Larson. *In*: THE BEST of Rent: Highlights from the original cast album. Intérprete: Vários artistas. Nova York: Verve, 1999. CD, faixa 20 (2 min).

FRANCO, E. C. P.; ARAÚJO, V. S. Questões Terminológico-conceituais no campo da tradução audiovisual. **Tradução em Revista**. n.11, 2011. Disponível em: <https://www.maxwell.vrac.puc-rio.br/18884/18884.PDFXXvmi=>. Acesso em: 8 fev. 2024.

FREIRE, G. C. **Os clíticos de terceira pessoa e as estratégias para sua substituição na fala culta brasileira e lusitana**. 2000. Dissertação (Mestrado em Língua Portuguesa) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2000.

FREIRE, G. C. **A realização do acusativo e do dativo anafóricos de terceira pessoa na escrita brasileira e lusitana**. 2005. 204 f. Tese (Doutorado em Língua Portuguesa) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.

GAMBIER, Y. **(Multi) Media Translation: concepts, practices and research**. Filadélfia: John Benjamins, 2001.

GIBBON, A. de O. **A expressão do tempo futuro na língua falada de Florianópolis: gramaticalização e variação**. 2000. 126 f. Dissertação (Mestrado em Linguística) – Centro de Comunicação e Expressão, Universidade Federal de Santa Catarina, Florianópolis, 2000.

GOTTLIEB, H. Subtitling: People translating people. *In*: DOLLERUP, C.; LINDEGAARD, A. (org.). **Teaching Translation and Interpreting 2: Insights, Aims, Visions**. Amsterdã; Filadélfia: Johns Benjamins, 1994. p. 261-274.

GRIES, S. T. *Estatística com R para a linguística*. Belo Horizonte: FALE/UFMG, 2019. Disponível em: http://www.letras.ufmg.br/site/elivros/Estat%C3%ADstica_com_R_Gries_%20Mello_et%20a1.pdf. Acesso em: 7 fev. 2024.

GUY, G. **As comunidades de fala: fronteiras internas e externas**. *In*: Abralín, Fortaleza, 2001.

GUY, G. R.; ZILLES, A. **Sociolinguística Quantitativa**. São Paulo: Parábola, 2007.

HURTADO ALBIR, A. **Traducción y traductología: introducción a la traductología**. Madri: Cátedra, 2001.

INSUBS. **Tutorial para legendas**. [s.l.], [2012?]. Disponível em: <https://www.mediafire.com/?vc3w6864e5xtllw>. Acesso em: 9 fev. 2024. Acesso em: 8 fev. 2024.

INSUBS. **Da tradução à postagem: como é o processo?** Insubs blog, [s.l.], 6 nov. 2010. Disponível em: <https://insubs-blog-blog.tumblr.com/post/1498996893/da-tradu%C3%A7%C3%A3o-%C3%A0-postagem-como-%C3%A9-o-processo>. Acesso em: 8 fev. 2024.

INSUBS. **Como é o mundo das legendas?** Insubs blog, [s.l.], 22 nov. 2013. Disponível em: <https://insubs-blog-blog.tumblr.com/post/65818137930/como-%C3%A9-o-mundo-das-legendas>. Acesso em: 8 fev. 2024.

IVARSSON, J.; CARROLL, M. **Subtitling**. Simrishamn: Transedit, 1998.

JAKOBSON, R. On Linguistic Aspects of Translation. *In*: BROWER, R. A. (org.) **On translation**. Nova York: Galaxy Books, 1966. p. 232-239.

KARAMITROGLOU, F. A proposed set of subtitling standards in Europe. **Translation Journal**, v. 2, n. 2, p. 1-5, 1998.

- LABOV, W. Some Principles of Linguistic Methodology. **Language in Society**, Cambridge, v. 1, n. 1, p. 97-120, 1972.
- LABOV, W. Where does the linguistic variable stop? A response to Beatriz Lavandera. **Working Papers in Sociolinguistics**, Austin, v. 44, p. 171-182, 1978.
- LABOV, W. Building on Empirical Foundations. In: LEHMANN, W. P.; MALKIEL, Y. (org.). **Perspectives on Historical Linguistics**. Philadelphia: John Benjamins Publishing, 1982. p. 17-92.
- LABOV, W. **Principles of Linguistic Change**. Vol 1: Int ed. Cambridge: Blackwell, 1994.
- LABOV, W. **Principles of Linguistic Change**. Vol 2: Soc ed. Cambridge: Blackwell, 2001.
- LABOV, W. Some sociolinguistic principles. In: PAULSTON, C. B.; TUCKER, G. R. (org.). **Sociolinguistics: The Essential Readings**. Oxford: Blackwell, 2003. p. 234-250.
- LABOV, W. **Padrões Sociolinguísticos**. São Paulo: Parábola, 2008[1972].
- LAUAR, A. B. T. F. **Não o vejo mais em Vitória**: a substituição do clítico acusativo de terceira pessoa na fala capixaba. 2015. 110 f. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Federal do Espírito Santo, Vitória, 2015.
- LAVANDERA, B. R. Where does the sociolinguistic variable stop? **Language in Society**, Cambridge, v.7, p. 171-182, 1978.
- LAVANDERA, B. R. **Variación y significado**. Buenos Aires: Librería Hachette, 1984.
- LEVSHINA, N. **How to do Linguistics with R**. Amsterdã: John Benjamins, 2015.
- LIMA, M. D. A. de O. **Continuum de gêneros textuais jornalísticos para a descrição de norma(s) culta(s)**: o acusativo anafórico de terceira pessoa e a ordem dos clíticos pronominais. 2022. 313 f. Tese (Doutorado em Letras Vernáculas) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2022.
- LOPES, C. R. dos S.; AFONSO, L. G. P. Variantes linguísticas de segunda pessoa na tradução audiovisual: a legendagem e a dublagem como material para estudos de variação e mudança. In: BIAZOLLI, C. C.; BERLINCK, R. de A. (org.). **Gêneros textuais-discursivos no estudo de processos de variação e mudança**. Campinas: Pontes Editores, 2021. p. 131-166.
- LOUVIERE, J. J.; HENSHER, D. A.; SWAIT, J. D. **Stated Choice Methods: Analysis and Application**. Cambridge: Cambridge University Press, 2000.
- LUCCHESI, D. **Sistema, mudança e linguagem**: um percurso na história da linguística moderna. São Paulo: Parábola, 2004.
- MACIEL, L. V. de C. Os elementos constitutivos do enunciado em suas relações dialógicas: um exemplo de análise. **Linguagem em (Dis)curso**, v. 15, n. 2, p. 249-266, 2015.
- MARAFONI, R. L. **A realização do objeto direto anafórico**: um estudo em tempo real de curta duração. 2004. 112 f. Dissertação (Mestrado em Letras Vernáculas) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2004.

MARCUSCHI, L. A. Gêneros textuais: definição e funcionalidade. *In*: DIONISIO, A. P.; MACHADO, A. R.; BEZERRA, M. A. (org.). **Gêneros textuais e ensino**. 4. ed. Rio de Janeiro: Lucerna, 2005. p. 19-36.

MARCUSCHI, L. A. **Produção textual, análise de gêneros e compreensão**. São Paulo: Parábola, 2008.

MARCUSCHI, L. A. Gêneros textuais emergente no contexto da tecnologia digital. *In*: MARCUSCHI, L.A.; XAVIER, A. C. (org.). **Hipertexto e gêneros digitais: novas formas de construção de sentido**. São Paulo: Cortez Editora, 2009.

MARCUSCHI, L. A. **Da fala para a escrita: atividades de retextualização**. 10. ed. São Paulo: Cortez, 2010[2001].

MARQUES DE SOUSA, A. A. **As realizações do acusativo anafórico em variedades do português**. 2021. 201 f. Tese (Doutorado em Letras Vernáculas) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2021.

MATEUS, M. H. M.; BRITO, A. M.; DUARTE, I.; FARIA, I. H.; FROTA, S.; MATOS, G.; OLIVEIRA, F.; VIGÁRIO, M.; VILLALVA, A. **Gramática da Língua Portuguesa**. 6. ed. Lisboa: Caminho, 2003[1983].

MEURER, J. L.; BONINI, A.; MOTTA-ROTH, D. **Gêneros: teorias, métodos, debates**. São Paulo: Parábola, 2005.

MILROY, L. **Language and Social Networks**. Oxford: Blackwell, 1980.

NAVES, S. B.; MAUCH, C. ALVES, S. F.; ARAÚJO, V. L. S. (org.). **Guia Para Produções Audiovisuais Acessíveis**. Brasília: Secretaria do Audiovisual do Ministério da Cultura, 2016.

NORD, C. Loyalty and Fidelity in Specialized Translation. **Confluências**, n. 4, p. 29-41, 2006.

NUNES, E. A. T. **A legendagem da televisão por assinatura do Brasil**. 2012. 119 f. Dissertação (Mestrado em Linguística) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2012.

NUNES, J. Direção de cliticização, objeto nulo e pronome tônico na posição de objeto em português brasileiro. *In*: ROBERTS, I; KATO, M. (org.). **Português Brasileiro: uma viagem diacrônica**. 3.ed. São Paulo: Contexto, 2018[1993]. p. 161-73.

O'HAGAN, M. Evolution of User-generated Translation: Fansubs, Translation, Hacking and Crowdsourcing. **The Journal of Internationalisation and Localisation**, v.1, p. 5-32, 2009.

OLIVERA, J. M. de. **O futuro da língua portuguesa ontem e hoje: variação e mudança**. 2006. 254 f. Tese (Doutorado em Língua Portuguesa) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2006.

OLIVEIRA, S. M. DE. **Legendação de metáforas: um estudo empírico-experimental bom base no filme “La lenga de las mariposas”**. 2008. 101 f. Dissertação (Mestrado em Língua Portuguesa) – Universidade Federal de Santa Catarina, Florianópolis, 2008.

OMENA, N. P. de. **Pronome pessoal de terceira pessoa: suas formas variantes em função acusativa**. 1978. Dissertação (Mestrado em Língua Portuguesa) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 1978.

OUSHIRO, L.; Coleta de dados para pesquisas sociolinguísticas (em tempo de pandemia), 04/2021, **XXXV ENANPOLL - Encontro Nacional da ANPOLL**, vol. 1, pp.1-8, Londrina, PR, Brasil, 2020.

OUSHIRO, L. **Introdução à estatística para linguistas**. Campinas: Editora da ABRALIN, 2022.

OUSHIRO, L. **dmsocio**. v0.2.0. Disponível em: oushiro.shinyapps.io/dmsocio. 2018. Acesso em: 8 fev. 2024.

OUSTINOFF, M. **Tradução: História, teorias e métodos**. São Paulo: Parábola, 2015.

PAES, J. P. **Tradução A ponte necessária: Aspectos e problemas da arte de traduzir**. São Paulo: Editora Ática, 1990.

PAGOTTO, E. G. **posição dos clíticos em português: um estudo diacrônico**. 1992. 168 f. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 1992.

PAVLOV, I. **7-zip**. 19.00. Disponível em: <https://www.7-zip.org/>. 2023. Acesso em: 8 fev. 2024.

PETERSON, M. S. **A ordem dos clíticos pronominais em lexias verbais simples e complexas em cartas de leitor: uma contribuição da Sociolinguística Variacionista**. 2010. 210 f. Dissertação (Mestrado em Língua Portuguesa) - Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2010.

PINTO, C. A. V. **Variação do objeto anafórico acusativo na fala de Florianópolis**. 2015. 164 f. Dissertação (Mestrado em Linguística) – Centro de Comunicação e Expressão, Universidade Federal de Santa Catarina, 2015.

PIVETTA, V. **Objeto direto anafórico no português brasileiro: uma discussão sobre a importância dos traços semântico-pragmáticos - animacidade/especificidade vs. gênero semântico**. 2015. 131 f. Dissertação (Mestrado em Teoria e Análise Linguística) – Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2015.

R CORE TEAM. **R: A language and environment for statistical computing**. Vienna: R. Foundation for Statistical Computing, 2023. Disponível em: <http://www.R-project.org/>. Acesso em: 8 fev. 2024.

REISS, K. Type, Kind and Individuality of Text: Decision Making in Translation. **Poetics Today**, v. 2, n. 4, p. 121–131, 1981.

REISS, K.; VERMEER, H. J. **Towards a general theory of translational action: skopos theory explained**. Londres; Nova York: Routledge Taylor & Francis Group, 2013[1984].

REMAEL, A. Audiovisual translation. In: GAMBIER, Y.; DOORSLAER, L. van. (org.). **Handbook of Translation Studies**. Amsterdã; Filadélfia: John Benjamins, 2010. p. 344-349.

ROCHA LIMA, C. H. da. **Gramática normativa da língua portuguesa**. 49. ed. Rio de Janeiro: José Olympio, 2011[1972].

RODRIGUES, L. da S. **O caso acusativo nos pronomes pessoais de terceira pessoa do português brasileiro e europeu**. 2018. 165 f. Tese (Doutorado em Linguística) – Centro de Humanidades, Universidade Federal do Ceará, 2018.

ROJO, R. Gêneros do discurso e gêneros textuais: questões teóricas e aplicadas. *In*: MEURER, J. L.; BONINI, A.; MOTTA-ROTH, D. (org.). **Gêneros**. Teorias, métodos, debates. São Paulo: Parábola, 2005. p. 184-207.

ROJO, R. H. R.; BARBOSA, J. P. **Hipermodernidade, multiletramento e gêneros discursivos**. São Paulo: Parábola, 2015.

ROMAINE, S. **Socio-Historical Linguistics**. Cambridge: Cambridge University Press, 2009[1982].

ROMERO-FRESCO, P. Accessible filmmaking: Joining the dots between audiovisual translation, accessibility and filmmaking. **Journal of Specialized Translation**, n. 20, p. 201-223, 2013.

RÓNAI, P. **A tradução vivida**. 4. ed. Rio de Janeiro: José Olympio, 2012.

SANKOFF, D.; TAGLIAMONTE, S. A.; SMITH, E. **GoldVarb X**: a variable rule application for Macintosh and Windows. 2005.

SANTANA, J. M. C. P. de. **Diagnose e ensino de pronomes**: um estudo sobre a retomada anafórica do objeto direto de terceira pessoa no Português brasileiro. 2016. 205 f. Dissertação (Mestrado em Letras Vernáculas) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2016.

SANTANA, J. M. C. P. de; PISSURNO, K. C. da S.; LIMA, M. D. A. de O. Variação estilística das estratégias de preenchimento do acusativo anafórico de terceira pessoa. *In*: VIEIRA, S. R.; LIMA, M. D. A. de O. (org.). **Variação, gêneros textuais e ensino de português**: da norma culta à norma-padrão. 1. ed. Rio de Janeiro: Letras UFRJ, 2019. p. 17-41.

SARDINHA, T. B. Linguística de corpus: histórico e problemática. **Revista D.E.L.T.A.**, v. 16, n. 2, p. 323-367, 2000.

SILVA, J. M. V. da. Construindo corpora de legendas: passo a passo metodológico para pesquisas baseadas em corpus. **Texto Livre: Linguagem e Tecnologia**, v. 12, n. 3, p. 168-195, 2019.

SILVA, V. F. F. da. **Fansubbing e legendagem profissional**: um estudo comparativo da tradução inglês-português de referências culturais na série Family Guy. 2021. 185 f. Dissertação (Mestrado em Tradução) – Faculdade de Letras, Universidade de Lisboa, Lisboa, 2021.

SPINELLI, A. C. **Pronomes e sua ausência: por um tratamento unificado da omissão e da expressão de sujeitos e objetos diretos pronominais de 3ª pessoa**. 2018. 80 f. Dissertação

(Mestrado em Estudos da Linguagem) – Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2018.

SPIRIDONOV, A. **Subtitle Workshop**. Version 6.0a. [S.I.];, 2013.

SPOLIDORIO, S. **Comunidades online e legendas de fãs: novas formas de produzir e consumir legendas**. 2017. 136 f. Dissertação (Mestrado em Linguística Aplicada) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2017.

TARALLO, F. L. **Relativization strategies in Brazilian Portuguese**. 1983. 270f. Dissertação (Doutorado em Linguística) – University of Pennsylvania, Pennsylvania, 1983.

TARALLO, F. L. **A pesquisa sociolinguística**. São Paulo: Editora Ática, 1986.

TARALLO, F. L. Por uma sociolinguística românica ‘paramétrica’: fonologia e sintaxe. *In: Ensaios de Linguística*, n.13, p. 51-83, 1987.

TESCH, L. M. **A expressão do tempo futuro no uso capixaba: variação e gramaticalização**. 2011. 192 f. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2011.

VIEIRA, S. R. **Colocação pronominal nas variedades europeia, brasileira e moçambicana: para a definição da natureza do clítico em português**. 2002. 441 f. Tese (Doutorado em Língua Portuguesa) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2002.

VIEIRA, S. R. (Coord.) **Projeto Pró-norma plural: do *continuum* fala-escrita para a norma-padrão (em andamento)**.

VIEIRA, S. R.; LIMA, M. D. A. de O. (org.). **Variação, gêneros textuais e ensino de português: da norma culta à norma-padrão**. 1. ed. Rio de Janeiro: Letras UFRJ, 2019.

WEINER, E. J.; LABOV, W. Constraints on the agentless passive. *Journal of Linguistics*, v. 19, p. 29-58, 1983.

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

ZABALBEASCOA, P. The nature of audiovisual text and its parameters. *In: DÍAZ-CINTAS, J. (org.). The didactics of audiovisual translation*. Amsterdã; Filadélfia: John Benjamins, 2008. p. 21-37.

APÊNDICE A
Lista de identificação das personagens

A1	Alex Karev	CP1	Charles Percy
A1PC11	Acompanhante 1 PC11	CP2	Casey Parker
A1PC15	Acompanhante 1 PC15	D1	Derek Shepherd
A1PC2	Acompanhante 1 PC2	D1N	Derek Shepherd (narração)
A1N	Alex Karev (narração)	D2	Andrew DeLuca
A2	April Kepner	D3	Dahlia Qadri
A2PC11	Acompanhante 2 PC11	E1	Enfermeira 1 Kathleen
A2PC15	Acompanhante 2 PC15	E2	Enfermeira 2 Andrea
A2PC2	Acompanhante 2 PC2	E3	Enfermeira 3
A3	Amelia Shepherd	E4	Enfermeiro 4
ACP18	Acompanhante PC18	E5	Enfermeiro 5
AD1	Advogado 1	E6	Enfermeira 6
AM1	Amanda	E7	Enfermeira 7
APC1	Acompanhante PC1	E8	Enfermeira 8
APC10	Acompanhante PC10	EC1	Elizabeth Chen
APC13	Acompanhante PC13	EH1	Evelyn Hunt
APC14	Acompanhante PC14	EI1	Enfermeira(o)
APC23	Acompanhante PC23	EL1	Eliza Minnick
APC3	Acompanhante PC3	FG1	Figurante 1
APC4	Acompanhante PC4	FG2	Figurante 2
APC5	Acompanhante PC5	FT1	Fisioterapeuta 1
APC7	Acompanhante PC7	G1	Guarda 1
AZ1	Arizona Robbins	GC1	Gary Clark
AZ1N	Arizona Robbins (narração)	GI1	Graciella Interna
B1	Miranda Bailey	H1	Heather Brooks
B1N	Miranda Bailey (narração)	HE1	Herrera
B2	Ben Warren	I1	Izzie Stevens
B3	Betty Nelson	I1N	Izzie Stevens (narração)
BB1	Bombeiro 1 Aaron	IN1	Interno T14E13
BB2	Bombeiro 2	J1	Jackson Avery
BB3	Bombeiro 3	J2	Jo Wilson
BB4	Bombeiro 4	JE1	Jimmy Evans
BB5	Bombeiro 5	K1	Knox
BBS1	Bombeiros	K2	Tom Koracic
C1	Callie Torres	L1	Lexie Grey
C1N	Callie Torres (narração)	L2	Leah Murphy
C2	Cristina Yang	LI1	Leo Interno
C2N	Cristina Yang (narração)	LJ1	Larry Jennings
C3	Carina DeLuca	LO1	Louise O'Malley
CA1	Catherine Avery	LS1	Loudspeaker
CB1	Candidato Buck	M1	Meredith Grey
CD1	Candice Warren	M1N	Meredith Grey (narração)
CI1	Claire Interna	M2	Matthew
CM1	Candidata Mindy	M3	Megan Hunt

M4	Maggie Pierce	PD1	Padre 1
MB1	Maura Brooks	PI1	Pierce Interno
MC1	Médico cirurgião 1	PM1	Paramédico 1
MC2	Médico cirurgião 2	PM2	Paramédica 2
MC3	Médico cirurgião 3	PM3	Paramédica 3
MG1	Miles Green	PM4	Paramédico 4
MI1	Megan Interna	PM5	Paramédica 5
MI1	Ministra 1	PM6	Paramédica 6
MK1	Mark Sloan	PO1	Policial 1
MK1N	Mark Sloan (narração)	PO2	Policial 2
MT1	Médico transplante 1	PO3	Policial 3
N1	Nathan Higgs	PO4	Policial 4
N2	Noiva	PO5	Policial 5
N3	Noivo	Q1	Qadri
N4	Naomi Evans	R1	Richard Webber
O1	Owen Hunt	R1N	Richard Webber (narração)
O1N	Owen Hunt (narração)	R2	Reed
OH1	Olivia Harper	RE1	Repórter 1
P1	Pedreiro 1	RE2	Repórter 2
P2	Pastor	RE3	Repórter 3
PC1	Paciente 1	RE4	Repórter 4
PC10	Paciente 10	RE5	Repórter 5
PC10	Paciente 10	RE6	Repórter 6
PC11	Paciente 11 Max	RO1	Residente oncológico 1
PC12	Paciente 12	RP1	Ray Pelletier
PC13	Paciente 13	S1	Stephanie Edwards
PC14	Paciente 14	S2	Shane Ross
PC15	Paciente 15	S3	Levi Schmitt
PC16	Paciente 16	SA1	Sam interna
PC17	Paciente 17	SG1	Seattle Grace 1
PC18	Paciente 18	SI1	Steve Interno
PC19	Paciente 19	SO1	Sofia
PC2	Paciente 2	SP1	Sobrepasto
PC20	Paciente 20	T1	Teddy Altman
PC21	Paciente 21	T2	Taryn Helm
PC22	Paciente 22	TE1	Terapeuta Owen
PC23	PC 23 Kimmie	TF1	Terry Funcionário 1
PC3	Paciente 3	TI1	Tilden Bombeiro
PC4	Paciente 4	TT1	Texto em tela
PC5	Paciente 5	TT2	Texto em tela (resp)
PC6	Paciente 6	V1	Vikram Roy
PC7	Paciente 7	WP1	Wedding planner
PC8	Paciente 8 Sasha	Z1	Zelador 1
PC9	Paciente 9 Oscar	Z2	Zola