UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE ARTES E COMUNICAÇÃO
LETRAS - BACHARELADO

JOÃO GABRIEL PEREIRA DA SILVEIRA

LEGENDAGEM, TRADUÇÃO AUTOMÁTICA E RECEPÇÃO: um estudo empírico-experimental

Recife
2022

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE ARTES E COMUNICAÇÃO
LETRAS - BACHARELADO

JOÃO GABRIEL PEREIRA DA SILVEIRA

LEGENDAGEM, TRADUÇÃO AUTOMÁTICA E RECEPÇÃO: um estudo empírico-experimental

Trabalho de Conclusão de Curso apresentado ao Curso de Letras - Bacharelado do Centro de Artes e Comunicação da Universidade Federal de Pernambuco como requisito parcial para a obtenção do título de Bacharel em Letras.

Orientadora: Profa. Dra. Arlene Koglin

Recife
2022

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Silveira, João Gabriel Pereira da.
Legendagem, tradução automática e recepção: um estudo empírico-experimental / João Gabriel Pereira da Silveira. - Recife, 2022.
53 : il.

Orientador(a): Arlene Koglin
Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de Pernambuco, Centro de Artes e Comunicação, Letras - Bacharelado, 2022.
Inclui referências, apêndices, anexos.

1. Legendagem. 2. Tradução automática. 3. Pós-edição. 4. Recepção. I. Koglin, Arlene. (Orientação). II. Título.

400 CDD (22.ed.)

JOÃO GABRIEL PEREIRA DA SILVEIRA

LEGENDAGEM, TRADUÇÃO AUTOMÁTICA E RECEPÇÃO: um estudo empírico-experimental

Trabalho de Conclusão de Curso apresentado ao Curso de Letras - Bacharelado do Centro de Artes e Comunicação da Universidade Federal de Pernambuco como requisito parcial para a obtenção do título de Bacharel em Letras.

Aprovado em: 20/05/2022.

BANCA EXAMINADORA

_____

Profa. Dra. Arlene Koglin (Orientadora)
Universidade Federal de Pernambuco

_____

Prof. Dr. José Luiz Vila Real Gonçalves (Examinador Externo)
Universidade Federal de Ouro Preto

Ao meu querido avô, por absolutamente tudo.

AGRADECIMENTOS

Após quatro longos anos de intensa dedicação, enfim concluo meu percurso enquanto estudante de graduação do Curso de Bacharelado em Letras da Universidade Federal de Pernambuco (UFPE). Essa jornada acadêmica, porém, não foi trilhada apenas por mim, mas também por diversas pessoas queridas que, de uma forma ou de outra, me acompanharam nesse longo percurso. Por isso, nominalmente agradeço:

Ao meu avô Lenival Silveira, pelas inumeráveis conversas sobre o universo acadêmico e por todo o suporte oferecido desde muito antes do início da graduação; e ao meu estimado pai Vinícius Guerra, a quem igualmente agradeço pelo apoio e por sempre acreditar no meu potencial. Vocês são os meus maiores exemplos de vida!

À minha companheira Cleysi Araújo, pelo carinho, amparo emocional, pela paciência e pelas “broncas” (essenciais) de sempre.

À minha orientadora Profa. Dra. Arlene Koglin, por ter me introduzido ao mundo apaixonante da Tradução Audiovisual (TAV) e ao fazer científico nesta área. Muitíssimo obrigado por todos os ensinamentos, pelas ótimas parcerias em trabalhos e eventos e, claro, pela brilhante orientação neste TCC.

Ao Prof. Dr. José Luiz Vila Real Gonçalves, por ter aceitado o convite para compor a Banca Examinadora e pela avaliação minuciosa do texto.

A todos os participantes que, de maneira gentil e voluntária, colaboraram para a realização da coleta dos dados e para a construção deste trabalho.

Ao Prof. Dr. Yuri Caribé, por ter me proporcionado o meu primeiro contato com o universo tradutório através da disciplina *LE1023 – Estudos da Tradução*; e ao Prof. Dr. Antônio Carlos Xavier, pelos três semestres de parceria na disciplina *LE735 – Linguística I: Fundamentos Teóricos* possibilitada por meio do

Programa de Monitoria da universidade, experiência esta que abriu muitas portas e me propiciou muitas alegrias.

Ao Grupo de Estudos e de Pesquisa em Tradução e Tecnologia (GETRADTEC), por todos os encontros de intensos debates sobre a área. Agradeço, em especial, a Morgana Matos, a Vitória Tassara e a Willian Moura, que têm contribuído diretamente para o meu amadurecimento enquanto pesquisador.

À Coordenação de Gestão Editorial e Impacto Social (CGEI) da UFPE, órgão do qual me orgulho por ter atuado como estagiário e por ter me proporcionado uma rica experiência no tocante à redação e à revisão de textos. Meus sinceros agradecimentos a Adriano Dias, Artur Villaça, Andressa Lira, Fernanda Barbosa, Isabel Padilha, Pedro Arruda e Widma Sandrelly pela parceria incrível e por todo o aprendizado que tive com vocês.

À Revista Ao Pé da Letra (APL), pelos dois anos de experiências no âmbito da revisão textual.

À minha querida turma de Letras-Bacharelado 2018.1, pela união e por todo o apoio durante todos esses anos. Agradeço, especialmente, ao meu *squad* – Bárbara Vicente, Lucas Aguiar e Maria Luiza –, pelos momentos de desabafo e pela nossa amizade sincera e descontraída.

Aos queridos professores Edu Silva e Tiago Licarião e à querida amiga Brenna Kelley, por terem despertado em mim o anseio pelas leituras no âmbito das Letras e das Ciências Humanas. A influência de vocês foi o “pontapé inicial” desta minha jornada e eu jamais poderia esquecer disso.

À minha querida tia Lenira Silveira, pelas maravilhosas conversas e pelo acolhimento em sua casa.

À UFPE, por todas as oportunidades a mim concedidas durante os quatro anos de graduação, e a cada docente do Departamento de Letras que contribuiu para o desenvolvimento de meu percurso acadêmico enquanto graduando.

Por fim, mas não menos importante, agradeço a cada familiar, amigo, professor e funcionário que, direta ou indiretamente, auxiliou no meu percurso escolar desde a Educação Básica até o Ensino Superior. Se hoje me torno um Bacharel em Letras, é porque vocês me incentivaram desde o início e me apoiaram em diversas etapas da minha vida escolar e pessoal. Por isso, meu muito obrigado.

"*The translation is a joint project undertaken by the translator's body and the author's spirit*" (ROBINSON, 2001, p. 36).

RESUMO

Esta pesquisa tem a finalidade de investigar a recepção de legendas interlinguais traduzidas automaticamente e pós-editadas no par linguístico EN/PT-BR através de um experimento piloto realizado com 8 participantes. Mais especificamente, objetiva determinar qual tipo de legenda (disponibilizada pela Netflix ou pós-editada) se mostra mais adequada para o telespectador e analisar a aceitabilidade da audiência frente às legendas pós-editadas. O experimento baseou-se na metodologia de triangulação de dados proposta por Alves (2001) e teve como material-base o trailer do filme Missão no Mar Vermelho, lançado em 2019 pela Netflix. O desenho experimental contou com dois grupos de participantes: a) o grupo um (G1), que ficou responsável por assistir ao trailer com legendas oficiais disponibilizadas na plataforma da Netflix; e b) o grupo dois (G2), que assistiu ao trailer com legendas traduzidas automaticamente e posteriormente pós-editadas. Para a realização da coleta, três instrumentos foram utilizados: um questionário prospectivo de levantamento de perfil, uma escala Likert e um protocolo verbal retrospectivo. Esses instrumentos foram aplicados individualmente via Google Meet e os dados foram tabulados, analisados e interpretados a partir das respostas dos participantes. Os dados provenientes do questionário prospectivo e do protocolo verbal retrospectivo foram discutidos qualitativamente por meio da apresentação e discussão das respostas e os dados quantitativos coletados a partir da escala Likert foram triangulados às informações provenientes dos instrumentos de coleta qualitativa, sendo igualmente discutidos. Os resultados obtidos a partir da análise dos dados coletados por meio da escala Likert e do protocolo verbal retrospectivo apontaram para uma melhor aceitabilidade dos espectadores no que se refere às legendas pós-editadas, concordando com estudos sobre o potencial uso da pós-edição na legendagem (DÍAZ-CINTAS; REMAEL, 2021; JI; OAKES, 2019; VOLK, 2008).

Palavras-chave: Legendagem. Tradução automática. Pós-edição. Recepção. Experimento piloto.

ABSTRACT

This piece of research aims to investigate the reception of machine-translated post-edited interlingual subtitles in the EN/PT-BR language pair through a pilot experiment conducted with 8 participants. More specifically, it aims to determine which type of subtitle (provided by Netflix or post-edited) is more adequate for the viewer and to analyze the audience's acceptability of post-edited subtitles. The experiment was based on Alves' (2001) data triangulation methodology and the material was the movie trailer entitled The Red Sea Diving Resort, released in 2019 by Netflix. The experimental design had two groups of participants: a) the group one (G1), who watched the trailer with official subtitles from Netflix platform; and b) the group two (G2), who watched the machine-translated and post-edited trailer subtitles. Moreover, three instruments were used for data collection: a prospective questionnaire to profile the participants, a five-point Likert scale, and a verbal protocol. These instruments were applied individually through Google Meet and the data were tabulated, analyzed, and interpreted from the participants' answers. The data from the prospective questionnaire and the think-aloud protocol were qualitatively discussed through the presentation and discussion of the answers, and the data collected from the Likert scale were triangulated with the information from the qualitative collection instruments and also discussed. The results derived from the data collected analysis through the Likert scale and the verbal protocol suggest that better acceptability of post-edited subtitles by the viewers, which is consistent with studies on the potential use of post-editing in subtitling (DÍAZ-CINTAS; REMAEL, 2021; JI; OAKES, 2019; VOLK, 2008).

Keywords: Subtitling. Machine translation. Post editing. Reception. Pilot experiment.

LISTA DE ABREVIATURAS E SIGLAS

AD	Audiodescrição
CC	Closed caption (legenda oculta)
GETRADTEC	Grupo de Estudos e de Pesquisa em Tradução e Tecnologia
G1	Grupo um
G2	Grupo dois
LC	Língua de chegada
Libras	Língua Brasileira de Sinais
LP	Língua de partida
LSE	Legendagem para surdos e ensurdecidos
NMT	Neural machine translation (tradução automática de base neural)
RBMT	Rule-based machine translation (tradução automática baseada em regras)
SE	Subtitle Edit
SEDCMRJ	Sindicato das Empresas Distribuidoras Cinematográficas do Município do Rio de Janeiro
SMT	Statistical machine translation (tradução automática de base estatística)
TA	Tradução automática
TAV	Tradução audiovisual
TC	Texto de chegada
TCLE	Termo de Consentimento Livre e Esclarecido
TP	Texto de partida
UFPE	Universidade Federal de Pernambuco

SUMÁRIO

1 INTRODUÇÃO 13

2 FUNDAMENTAÇÃO TEÓRICA 16

2.1 LEGENDAGEM 16

2.2 ESTUDOS DE RECEPÇÃO 18

2.3 TRADUÇÃO AUTOMÁTICA (TA) 18

2.4 PÓS-EDIÇÃO (PE) 21

3 METODOLOGIA 24

3.1 MATERIAL 24

3.2 PARTICIPANTES 26

3.3 INSTRUMENTOS DE COLETA DE DADOS 26

3.3.1 Questionário prospectivo de levantamento de perfil 27

3.3.2 Escala Likert 27

3.3.3 Protocolo verbal retrospectivo 28

3.4 DESENHO EXPERIMENTAL 29

3.5 PROCEDIMENTOS PARA COLETA DOS DADOS 30

3.6 METODOLOGIA DE ANÁLISE DOS DADOS 31

4 ANÁLISE E DISCUSSÃO DOS DADOS 33

5 CONSIDERAÇÕES FINAIS 41

REFERÊNCIAS 43

APÊNDICE A – Questionário prospectivo de levantamento de perfil 48

ANEXO A – Termo de Consentimento Livre e Esclarecido (TCLE) 50

ANEXO B – Protocolo verbal retrospectivo 52

1 INTRODUÇÃO

A tradução audiovisual (TAV), um dos vários ramos de prática de pesquisa pertencentes à grande área dos Estudos da Tradução, tem se consolidado ao longo dos anos não apenas como campo de investigações científicas, mas também como um importante espaço mercadológico e de inclusão. Tendo a acessibilidade como característica principal, a TAV conta com várias subáreas que são igualmente voltadas para a inclusão linguística e social dos indivíduos, a exemplo da dublagem, da interpretação, do *voice-over*, da audiodescrição (AD) e da legendagem.

A legendagem, uma dessas subáreas, tem igualmente se destacado como um frutífero campo de pesquisas e de atuação profissional (GAMBIER, 2009; GUILLOT, 2018), sobretudo no que diz respeito à legendagem interlingual. Apesar de as pesquisas nesse escopo serem predominantemente voltadas para o produto, ou seja, para a legenda em si e suas características internas, as investigações acerca da recepção de legendas vêm cada vez mais conquistando espaço nos diversos centros de pesquisa em tradução espalhados pelo mundo, uma vez que, com o surgimento das plataformas de *streaming*, as exigências do público no que se refere aos parâmetros técnicos e linguísticos das legendas passaram a ganhar uma visibilidade maior.

A partir dessa necessidade de levar em consideração a opinião da audiência acerca das legendas de uma dada obra audiovisual e dos avanços tecnológicos recentes, inúmeras ferramentas de auxílio à tradução foram sendo desenvolvidas e utilizadas no escopo da legendagem, tais como o uso de dicionários e glossários *on-line*, de bancos de dados terminológicos, de *softwares* de legendagem e de sistemas de tradução automática (TA). Esses avanços, impulsionados pela globalização, trouxeram uma gama de benefícios também para o processo tradutório em si, destacando-se dentre eles a diminuição do tempo dispendido na tarefa e a redução do trabalho manual.

A TA, uma dessas ferramentas, tem sido utilizada com frequência e despertado o interesse tanto de pesquisadores como de profissionais do mercado de tradução, uma vez que promove a automatização do processo tradutório de uma língua de partida (LP) para uma língua de chegada (LC) por meio de sistemas executados por uma máquina. Tais sistemas, por serem constituídos de arquiteturas distintas, costumam gerar como resultado uma tradução “crua” com necessidade de

intervenção humana – denominada pós-edição – em certos trechos do texto traduzido automaticamente (MARTINS, 2008).

Essa tradução crua tem como mérito a redução do tempo de tradução de um texto traduzido por um dado sistema de TA, que, além de poupar o tradutor humano de problemas básicos, o deixaria livre para se ocupar apenas de casos mais complexos de difícil resolução por parte da máquina (MARTINS, 2008). Devido a essas características de otimização de tempo e trabalho manual, a pós-edição de textos traduzidos automaticamente tornou-se foco do mercado de tradução e de centros de pesquisa em escala global, sendo inclusive utilizada com mais frequência em certos nichos mercadológicos em comparação à tradução humana, a exemplo da indústria de localização (GUERBEROF, 2009).

Entretanto, apesar desse crescente interesse por parte de pesquisadores da área, observa-se ainda uma grande lacuna e carência de estudos no âmbito da *recepção de legendas interlinguais traduzidas automaticamente e posteriormente pós-editadas por tradutores*, lacuna esta observada não apenas no âmbito da recepção de legendas (ORREGO-CARMONA, 2015), mas também na área de forma geral. Contudo, por ser relativamente fácil obter dados paralelos com legendas de qualidade, o uso de legendas automaticamente traduzidas em experimentos tem crescido nos últimos anos, impulsionando diversas pesquisas nesse escopo (JI; OAKES, 2019).

Dito isso, este estudo tem como objetivo geral investigar a recepção de legendas interlinguais no par linguístico inglês-português (EN/PT-BR) com o auxílio da TA. A partir desse propósito geral, tem-se como objetivos específicos: (a) determinar qual tipo de legenda (disponibilizada pela Netflix ou pós-editada) se mostra mais adequada para o telespectador; e (b) analisar a aceitabilidade pelos telespectadores frente às legendas pós-editadas.

A consecução dos objetivos propostos é norteada pelas seguintes perguntas de pesquisa: (a) qual tipo de legenda (disponibilizada pela Netflix ou pós-editada) possui melhor aceitabilidade pelo telespectador?; e (b) quais as expectativas/necessidade da audiência brasileira em termos de parâmetros linguísticos nas legendas?.

Dito isso, esta monografia apresenta uma estrutura composta por cinco capítulos no total, os quais abrangem esta Introdução e, em seguida, a Fundamentação Teórica, subdividida em quatro seções: Legendagem, Estudos de

recepção, Tradução automática (TA) e Pós-edição (PE). O capítulo seguinte é o da Metodologia, que se subdivide em seis seções – Material, Participantes, Instrumentos de coleta de dados, Desenho experimental, Procedimentos para coleta dos dados e Metodologia de análise dos dados –, seguido da Análise e discussão dos dados, e, por fim, das Considerações finais.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo será apresentado o arcabouço teórico utilizado como base para o desenvolvimento desta pesquisa. Para fins de organização, o capítulo é composto por 4 seções, as quais teorizam sobre a legendagem, os estudos de recepção, a tradução automática (TA) e a pós-edição (PE).

2.1 LEGENDAGEM

Em meados da segunda metade do século XX, os Estudos da Tradução conquistaram um espaço significativo dentro das instituições de Ensino Superior e se estabeleceram enquanto campo científico. Com a sistematização dos estudos sobre a tradução e o consequente crescimento do número de publicações acadêmicas, a área responsável pela descrição, teorização e análise dos processos, contextos e produtos resultantes do ato tradutório passou a manifestar um caráter inter e multidisciplinar em seu escopo teórico, impulsionando o surgimento de novas metodologias de pesquisa (WILLIAMS; CHESTERMAN, 2002).

A partir dessa evolução teórico-metodológica e com o auxílio dos avanços tecnológicos cada vez mais progressivos, novos campos de atuação profissional foram surgindo e se consolidando no âmbito acadêmico e no mercado, a exemplo da Tradução Audiovisual (TAV). De acordo com Rebollo-Couto, Silva e Silva (2017), A TAV é a área destinada à tradução de textos que circulam nos mais variados ambientes multimídia, tais como a televisão e o cinema, e é subdividida em algumas modalidades básicas, a saber: a dublagem, a legendagem interlingual[1], o *voice-over,* a interpretação simultânea, a janela de Língua Brasileira de Sinais (Libras), a legendagem para surdos e ensurdecidos (LSE), a legenda oculta – ou *closed caption* (CC) – e a audiodescrição (AD).

A legendagem interlingual, uma dessas modalidades, é categorizada como uma das subáreas da TAV que deram origem a uma vasta quantidade de pesquisas nos últimos anos (GAMBIER, 2009; GUILLOT, 2018; MARTINEZ, 2007). Trata-se de

[1] De acordo com Jakobson (2007), uma tradução do tipo *interlingual* – ou tradução propriamente dita – diz respeito à intepretação de signos verbais de uma língua (chamada “língua de partida”) por meio de uma outra língua (denominada “língua de chegada”). Em uma legendagem interlingual no par linguístico EN/PT-BR, por exemplo, a tradução é feita partindo-se do áudio da língua de chegada (inglês) para a transposição escrita na língua de partida (português).

um tipo de tradução que envolve a mudança do código oral para o código escrito e a transposição de um ou vários idiomas para outro ou para dois outros idiomas, como no caso da legendagem bilíngue (GAMBIER, 2009).

De acordo com Gottlieb (2005), a legendagem inclui quatro canais semióticos que, por vezes, são expostos simultaneamente em filmes e programas de televisão: (a) o canal auditivo verbal, que inclui diálogos, vozes de fundo e letras musicais; (b) o canal auditivo não-verbal, incluindo músicas, sons naturais e efeitos sonoros; (c) o canal visual verbal, que compreende títulos sobrepostos e sinais escritos na tela; e (d) o canal visual não-verbal, que abarca a composição e o fluxo das cenas. O autor também pontua que a legendagem possui como aspectos técnicos gerais a colocação de uma ou duas linhas na parte inferior da tela com uma média máxima de 35 caracteres por linha, podendo a legenda estar centralizada ou alinhada à esquerda (GOTTLIEB, 2005).

Sendo a legendagem, portanto, um tipo de tradução que considera texto escrito, sons e imagens como aspectos estruturais, as legendas não devem ser consideradas apenas sob uma perspectiva linguística, uma vez que envolvem imagens e sons que, aliados ao aspecto linguístico, constituem um texto polissemiótico (CHILE, 1999). Além dessas questões estruturais, há dois pontos que são igualmente relevantes para a tradução de legendas enquanto campo de pesquisas e de atuação profissional: a acessibilidade e a recepção do público-alvo.

Com o advento das tecnologias de auxílio à tradução, a legendagem tornou-se mais rápida e financeiramente acessível, sobretudo em comparação à dublagem. Questões como a disponibilidade instantânea de legendas em plataformas e dispositivos digitais, a multiplicidade de opções de visualização e a variedade de idiomas, fontes, textos e tamanhos, por exemplo, tornaram-se “peças-chave” para a comercialização de legendas nos dias atuais, principalmente no que diz respeito às legendas presentes nas plataformas de *streaming*[2] (GUILLOT, 2018).

Dada essa característica versátil, as capacidades e expectativas dos espectadores acerca das legendas são variáveis em virtude da considerável diversidade de público-alvo, que, por sua vez, pertence a contextos socioculturais

[2] Adão (2006, p. 21, grifo do autor) define *streaming* como um “[...] processo de distribuição de conteúdos, via Internet, em que o utilizador inicia a sua visualização sem necessidade de fazer *download* dos ficheiros que constituem os conteúdos, permitindo o início da visualização num curto espaço de tempo e exibindo o conteúdo sequencialmente, à medida que este vai chegando ao computador do utilizador”.

distintos e possui diferentes idades, níveis de escolaridade, status social, hábitos de visualização e meios de processamento de informações (GUILLOT, 2018). Sendo assim, é de suma importância investigar as caraterísticas, impressões e demais aspectos percebidos pelos espectadores ao consumirem uma dada obra audiovisual, tarefa esta que, por sua vez, já é desempenhada por um campo de pesquisas que vem cada vez mais conquistando espaço no escopo da TAV: os estudos de recepção.

2.2 ESTUDOS DE RECEPÇÃO

Os estudos de recepção se configuram enquanto campo multidisciplinar dotado de diversos paradigmas teóricos, conceitos e metodologias nas ciências humanas e sociais. Esses paradigmas, ancorados em abordagens comportamentais e sociológicas, são teorizados com o intuito de compreender como o público avalia um determinado produto e como a recepção deste produto se encontra atrelada às nuances socioculturais da audiência (HILL, 2018).

Nos estudos da tradução audiovisual, especificamente no escopo da legendagem interlingual, as pesquisas que buscam investigar a recepção de legendas são recentes nas instituições acadêmicas e ainda não contam com uma bibliografia robusta, dada sua complexidade (NIKOLIĆ, 2018). Apesar disso, a legendagem tem recebido forte apoio de ferramentas tecnológicas e o uso de equipamentos de última geração tem revolucionado as investigações empíricas nos estudos de recepção de legendas, a exemplo do rastreador ocular (GUILLOT, 2018).

Aliada a essa escassez de estudos que investiguem a recepção da tradução de legendas, também é possível notar os impactos da globalização nas demandas do público, que apresenta cada vez mais sinais de impaciência em relação ao consumo de legendas e, devido a isso, parece estar mais disposto a ignorar barreiras linguísticas na legendagem de obras audiovisuais. Isso ocorre porque o público quer apreciar jogos, filmes e séries logo após o lançamento, e, consequentemente, a audiência pode se mostrar disposta a sacrificar a qualidade de uma tradução especializada (DÍAZ-CINTAS; PIÑERO, 2015).

Em virtude dessa impaciência e da predisposição da audiência em sacrificar a qualidade das legendas, é importante observar quais aspectos técnicos e linguísticos influenciam na recepção de uma dada obra audiovisual legendada.

Plataformas de *streaming* como a Netflix e a Amazon, por exemplo, são ambientes considerados propícios para esse tipo de investigação, uma vez que esses espaços têm sido cruciais no tocante à produção e distribuição de conteúdos audiovisuais legendados, concentrando espectadores de várias partes do Brasil e do mundo (CAMPOS; AZEVEDO, 2020).

De acordo com Nikolić (2018), existem diversos modos de se conduzir uma pesquisa de recepção no escopo da legendagem interlingual, havendo estudos que promovem análises qualitativas (por meio de questionários e protocolos verbais), quantitativas (através de instrumentos de cunho estatístico) ou uma combinação dos dois métodos. Qualquer que seja o método utilizado, variáveis como a idade, o sexo, as afinidades linguísticas e o nível de escolaridade da audiência, por exemplo, são elementos que têm um impacto considerável na maneira como as legendas são recebidas e avaliadas, uma vez que os resultados podem diferenciar de acordo com as características etnográficas de cada grupo de espectadores.

Apesar de haver essa variedade de modos de se conduzir pesquisas empíricas sobre a recepção de legendas interlinguais, essa prática não conta com grandes investimentos e um alto número de publicações acadêmicas no Brasil. Em virtude dessa carência, faz-se necessário investigar a recepção da audiência em termos de parâmetros técnicos e linguísticos de legendas, sobretudo no que diz respeito à recepção de legendas traduzidas automaticamente e posteriormente pós-editadas por tradutores humanos.

2.3 TRADUÇÃO AUTOMÁTICA (TA)

A tradução automática (TA) é o processo por meio do qual a conversão entre línguas de um texto é efetuada por uma máquina, mesmo que esta tenha seu bom funcionamento garantido por um humano antes, durante ou depois do processo de tradução (KENNY, 2020). Tendo surgido e sido impulsionada na década de 1940, com o início da Guerra Fria, a TA foi sendo patrocinada e amplamente utilizada pelos governos americano e inglês com o intuito de obter informações à distância da inteligência soviética de maneira mais rápida do que o habitual (ALFARO; DIAS, 1998).

Apesar de ter sido alvo de desconfiança por parte de cientistas e políticos – que desejavam resultados tão eficientes quanto os gerados pela tradução humana –

e ter sofrido oscilações que variavam entre cortes e investimentos, a TA “resistiu” aos transtornos e se consolidou como uma importante ferramenta de auxílio ao tradutor humano (ALFARO; DIAS, 1998). Além disso, a TA se estabeleceu como um profícuo campo de estudos científicos dentro dos Estudos da Tradução, o que impulsionou o surgimento de novos núcleos de pesquisa e de novas metodologias na área.

Com o passar dos anos e o crescimento contínuo de investimentos em pesquisas sobre o funcionamento da TA, alguns sistemas foram desenvolvidos com o intuito de aprimorar o desempenho da máquina. Desses, três são utilizados em larga escala nas pesquisas envolvendo tradução por máquina: o sistema de tradução automática baseada em regras – ou *rule-based machine translation* (RBMT) –, o sistema de tradução de base estatística – ou *statistical machine translation* (SMT) – e o sistema de tradução de base neural – ou *neural machine translation* (NMT).

De acordo com Kenny (2011), a tradução automática baseada em regras (RBMT) utiliza aspectos formais das línguas para representar as estruturas gramaticais e palavras normalmente 'permitidas' em um determinado texto de partida (TP) e texto de chegada (TC). Mais recentemente, a RBMT também passou a utilizar 'regras de transferência' bilíngues com o intuito de promover um mapeamento entre as palavras e estruturas das línguas envolvidas.

Diferentemente da RBMT, a tradução automática de base estatística (SMT) não se apoia em gramáticas e léxicos dispendiosos. Em vez disso, ela tem como base um corpus paralelo já existente de TPs e suas traduções humanas, sendo que, quanto maior for o corpus paralelo em que um sistema SMT é treinado, melhor será o modelo estatístico que o sistema pode elaborar no tocante à tradução dos idiomas envolvidos (KENNY, 2011). Em outras palavras, a SMT depende da tradução humana não apenas para que o sistema seja alimentado, mas também para que a própria SMT tenha sua legitimidade garantida, já que os *corpora* paralelos utilizados pela SMT contêm boas respostas para problemas de tradução (KENNY, 2011).

A tradução automática neural (NMT), por sua vez, depende de técnicas de aprendizagem automática para elaborar modelos de tradução baseados em *corpora* paralelos pré-existentes, tal como ocorre na tradução automática de base estatística (KENNY, 2020). Porém, na NMT, esse processo é realizado por meio de métodos computacionais diferentes, uma vez que esse tipo de tradução utiliza redes neurais

artificiais por meio das quais os estados de ativação de grandes conjuntos de neurônios artificiais, interconectados entre si, podem ser entendidos como representações distribuídas de palavras soltas e seus respectivos co-textos com os quais aparecem nos dados de treinamento (KENNY, 2020).

Qualquer que seja o sistema de TA utilizado, o resultado gerado pelo sistema carecerá de correções. De acordo com Martins (2008), a tradução crua (ou *raw translation*) é o tipo de tradução automática que produz resultados intencionalmente parciais para que esses resultados possam servir como ponto de partida para posteriores correções, normalmente realizadas por revisores e/ou tradutores especializados. Esse processo é denominado pós-edição e apenas recentemente passou a ser utilizado no mercado de tradução e em experimentos da área, conforme será discutido na seção seguinte.

2.2 PÓS-EDIÇÃO (PE)

A pós-edição (PE) é o processo por meio do qual é feita a revisão de uma tradução produzida por um sistema de TA, ainda que o sistema necessite da interação humana antes ou depois do processamento automático do texto (KOVAČEVIĆ, 2014; MESA-LAO, 2013). A principal diferença entre a PE e a revisão humana é que a PE é a revisão de um texto traduzido automaticamente, ao passo que a revisão é a etapa posterior da tradução humana.

O objetivo da PE é reduzir o tempo gasto na tradução, aumentar a produtividade e solucionar problemas básicos para o tradutor humano, deixando para este a tarefa de corrigir apenas alguns casos específicos de difícil resolução por parte da máquina, a exemplo de metáforas e expressões idiomáticas (MARTINS, 2008). Sendo assim, a tarefa do pós-editor é realizar correções e modificações pontuais no texto “cru” – o texto traduzido por um sistema de TA (KOGLIN, 2015b) –, diferenciando essa atividade profissional da atividade exercida por revisores de textos.

Nos últimos anos, a pós-edição de textos traduzidos automaticamente tem despertado o interesse não apenas dos centros de pesquisa em TA, mas também do mercado de tradução. O motivo desse interesse crescente se baseia no fato de que os tradutores têm se mostrado motivados com o uso da PE em virtude da melhora na qualidade dos sistemas de TA, que, por sua vez, geram ganhos na produtividade

que podem variar de 30 a 300%, segundo dados da Associação de Tradução Automática nas Américas – no original *Association for Machine Translation in the Americas* (AMTA) (MACHINE..., [210-?]).

Em estudo que buscou investigar o esforço cognitivo despendido na pós-edição de metáforas em comparação à tradução manual de metáforas, Koglin (2015a) ressalta que as ambiguidades, frases longas e complexas, referências anafóricas, metáforas e outros aspectos textuais e discursivos do TP representam dificuldades para os sistemas de TA. Em consequência disso, a PE tem sido utilizada sobretudo em textos de natureza técnico-científica que possuem uma ocorrência menor de metáforas e menos variação lexical, tais como relatórios, manuais e artigos científicos.

Ainda que a PE esteja sendo utilizada com mais frequência em comparação à tradução humana em certos nichos mercadológicos, a exemplo da indústria de localização (GUERBEROF, 2009), ela ainda não é utilizada em larga escala pela indústria de legendas fílmicas. Porém, de acordo com Ji e Oakes (2019), a relativa facilidade de se adquirir dados paralelos de legendas de qualidade foi o incentivo-chave para que o uso de legendas traduzidas automaticamente passasse a ser aplicado em pesquisas, a exemplo dos experimentos realizados pelos projetos *SUbtitling by MAchine Translation (SUMAT)* e *EU-BRIDGE*.

O *SUMAT* (2011-2014), teve como foco a construção de *corpora* de legendas alinhadas para treinar motores de tradução automática de base estatística (SMT) em diversos pares linguísticos. O objetivo final do projeto foi produzir legendas automáticas, que seriam posteriormente pós-editadas, com vistas a incrementar a produtividade, reduzir custos e aumentar a qualidade do produto final (JI; OAKES, 2019). O *EU-BRIDGE*, por sua vez, foi conduzido no mesmo período e possuía objetivos semelhantes ao primeiro. Entretanto, o objetivo central foi testar a possibilidade de reconhecimento de fala para a posterior legendagem automática de vídeos (JI; OAKES, 2019).

Dito isso, é perceptível que a legendagem já incorporou o uso de tecnologias que auxiliam o processo de tradução no que se refere às suas especificidades técnicas, a exemplo da duração da legenda na tela, do número de linhas e caracteres e da sincronia entre o diálogo e a legenda em si. Ou seja, com o surgimento de *softwares* de legendagem, esses aspectos técnicos passaram a ser

realizados semi-automaticamente na interface dos programas, o que facilitou drasticamente o trabalho manual do tradutor.

Aliado a isso, as exigências quase que “imediatas” dos espectadores no que se refere às legendas de filmes, séries e *games* parecem ter um potencial de ser atendidas com a pós-edição de legendas, levando-se em conta a redução do tempo e o aumento da qualidade das legendas (DÍAZ-CINTAS; REMAEL, 2021; VOLK, 2008). Sendo assim, é de suma importância investigar não apenas a qualidade desse tipo de legendas, mas também a aceitabilidade do público no que se refere ao produto final.

Isso posto, o próximo capítulo abordará o quadro metodológico referente a um experimento conduzido sobre a recepção da legendagem auxiliada por TA.

2 METODOLOGIA

Neste capítulo serão apresentados os aspectos metodológicos relacionados a um experimento piloto conduzido pelo Grupo de Estudos e de Pesquisa em Tradução e Tecnologia da Universidade Federal de Pernambuco (GETRADTEC-UFPE) no âmbito da TAV. O experimento foi originalmente conduzido a partir do método da triangulação de dados (cf. ALVES, 2001) e envolveu instrumentos de coleta quantitativa e qualitativa que são normalmente utilizados nesse tipo de metodologia.

Nos Estudos da Tradução, o método da triangulação de dados é utilizado em pesquisas empírico-experimentais que buscam estabelecer parâmetros de intersubjetividade em busca de resultados mais confiáveis e passíveis de generalização por meio do cruzamento de dados qualitativos e quantitativos. Sendo assim, a triangulação se apresenta como uma alternativa metodológica que objetiva descrever com objetividade as características processuais do processo tradutório sem desprezar, porém, sua natureza subjetiva (ALVES, 2001).

Para fins de organização, este capítulo foi dividido em 6 seções, as quais descrevem o material, os participantes, os instrumentos de coleta de dados, o desenho experimental, os procedimentos para coleta dos dados e a metodologia de análise dos dados da pesquisa.

3.1 MATERIAL

O material que serviu de suporte para a realização desta pesquisa foi o trailer do filme Missão no Mar Vermelho (no original *The Red Sea Diving Resort*), um longa-metragem de suspense e espionagem dirigido pelo diretor israelense Giden Raff e lançado pela Netflix no ano de 2019:

Figura 1 – Capa do filme Missão no Mar Vermelho (no original *The Red Sea Diving Resort*) em PT-BR e EN

Fonte: Missão... (2019).

O filme foi inspirado na Operação Moisés, ocorrida na década de 1980, que foi realizada secretamente com o intuito de resgatar judeus etíopes que haviam fugido para o Sudão devido à fome na Etiópia. Nessa operação, cerca de dez mil judeus etíopes alojados em campos de concentração sudaneses foram transportados para Israel por meio de diversas ações de resgate empreendidas pelo país (CARVALHO, 2013).

O trailer conta com dois minutos e trinta e um segundos no total. Os critérios para a seleção do trailer foram os seguintes:

a) estar disponível legendado na Netflix no par linguístico EN/PT-BR;
b) estar disponível para livre acesso na internet o mesmo trailer, porém sem legendas, para que pudesse ser legendado com o auxílio da tradução automática no par linguístico EN/PT-BR.

A escolha por legendas oriundas da Netflix baseou-se no fato de, no cenário brasileiro atual, o consumo de produtos legendados estar concentrado nesse meio, ainda que, por outro lado, exista uma procura crescente por obras dubladas em

comparação às obras legendadas tanto na TV por assinatura quanto nas salas de cinema (KOGLIN, 2011). De acordo com pesquisa do Instituto Datafolha realizada em 2012 para o Sindicato das Empresas Distribuidoras Cinematográficas do Município do Rio de Janeiro (SEDCMRJ), 56% do público consultado prefere o consumo de obras audiovisuais dubladas e 35% prefere prestigiar conteúdos legendados (ARANTES, 2013). Essa tendência, já percebida pelo mercado, impulsionou muitos canais a providenciarem versões dubladas de filmes como procedimento padrão, tais como a Fox, a TNT e a Megapix (ARANTES, 2013).

3.2 PARTICIPANTES

No que se refere aos participantes do experimento piloto, foram aplicados os seguintes critérios no recrutamento para garantir a homogeneidade da amostra:

a) ser falante nativo de português brasileiro;
b) ter conhecimento de Língua Inglesa (mínimo B1)[3];
c) ser estudante de graduação em Letras (bacharelado ou licenciatura) na UFPE a partir do 2º período;
d) ter a preferência por obras audiovisuais legendadas.

Os participantes da amostra foram recrutados via *e-mail* e todo o experimento foi realizado na modalidade virtual, por meio do Google Meet. Tal alternativa permitiu que todos pudessem agendar o dia e horário da coleta de acordo com sua preferência e disponibilidade de tempo, garantindo, portanto, certa flexibilidade para ambas as partes (participante e pesquisador).

3.3 INSTRUMENTOS DE COLETA DE DADOS

Sobre os instrumentos de coleta de dados utilizados para a realização do experimento piloto, foram aplicados, individualmente, um questionário prospectivo de levantamento de perfil, uma escala Likert e um protocolo verbal retrospectivo. Os três instrumentos serão individualmente tratados a seguir.

[3] O Quadro Europeu Comum de Referência para as Línguas (CEFR) foi o meio padrão utilizado como parâmetro de proficiência em língua estrangeira (COUNCIL OF EUROPE, 2020).

3.3.1 Questionário prospectivo de levantamento de perfil

O questionário prospectivo é um instrumento de viés predominantemente qualitativo utilizado para a realização de um levantamento de perfil de cada participante do experimento. As questões nele contidas são, portanto, pensadas e elaboradas com o intuito de coletar informações pessoais, acadêmicas e profissionais dos participantes, tais como o nome completo, *e-mail*, sexo, a idade, a instituição a qual está vinculado e demais informações sobre sua vida acadêmica e/ou profissional.

Esse tipo de questionário é frequentemente utilizado nos estudos tradutórios em pesquisas de caráter empírico-experimental, auxiliando no levantamento de perfil de participantes diversos para a posterior realização das coletas e análises de dados (cf. ALVES, 2001; FONSECA, 2019; FONSECA; GONÇALVES; ZAMPIER, 2019; KOGLIN, 2015b). Trata-se, portanto, de um instrumento essencial nos Estudos da Tradução para a organização das variáveis de um dado experimento científico.

Especificamente para os fins desta pesquisa, cujo objetivo é o de investigar a recepção de legendas interlinguais auxiliadas por TA, algumas das questões solicitadas buscaram coletar informações sobre qual é o idioma de tradução (ou segundo idioma) do participante, seu grau de proficiência em Língua Inglesa, sua preferência no tocante à modalidade de acesso nos filmes (se prefere assistir dublado ou legendado), a frequência com que assiste a filmes legendados e em que lugar ou plataforma digital costuma assistir a filmes legendados (ver Apêndice A).

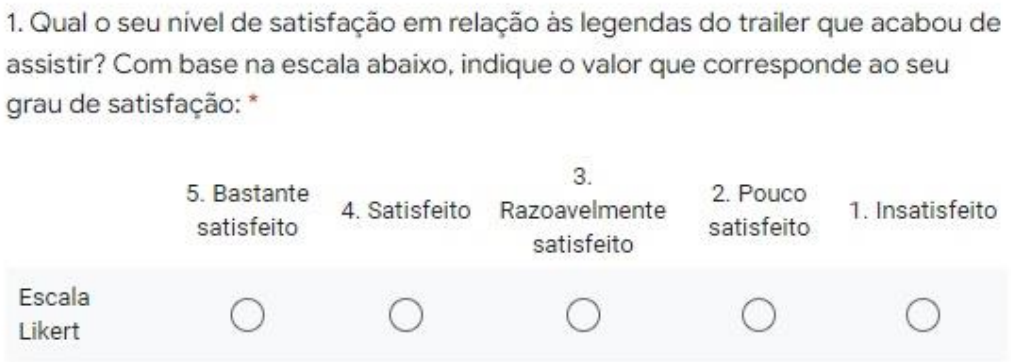
O questionário foi elaborado por meio da ferramenta Google Forms e disponibilizado aos participantes por meio de um *link* compartilhado no *chat* do Google Meet durante a realização do experimento piloto.

3.3.2 Escala Likert

A escala Likert é um instrumento geralmente utilizado em questionários que “[...] permite atribuir valores quantitativos a dados de natureza qualitativa” (KOGLIN, 2015b, p. 53). Tendo sido idealizada em 1932 pelo professor estadunidense Rensis Likert, essa escala é utilizada em diversas pesquisas com o intuito de medir o nível de concordância do público-alvo em relação a uma determinada questão ou produto.

No experimento aqui relatado, foi utilizada uma escala Likert de cinco pontos, elaborada por meio do Google Forms e disponibilizada na direção horizontal aos participantes após estes assistirem ao trailer. Os pontos buscaram solicitar o nível de satisfação dos participantes em relação às legendas do trailer assistido e apenas uma opção poderia ser assinalada dentre os seguintes níveis de satisfação apresentados em cada ponto da escala: 5. Bastante satisfeito; 4. Satisfeito; 3. Razoavelmente satisfeito; 2. Pouco satisfeito; e 1. Insatisfeito. A estrutura da escala pode ser visualizada na figura a seguir:

Figura 2 – Escala Likert de cinco pontos apresentada aos participantes durante o experimento



Fonte: O autor (2022).

Apesar da posição horizontal adotada no experimento aqui apresentado, a escala Likert também pode ser disponibilizada na direção vertical e pode ser elaborada por meio de outras plataformas de elaboração de formulários.

Os dados quantitativos provenientes da escala utilizada no experimento serão triangulados com os dados de natureza qualitativa coletados por meio do questionário prospectivo de levantamento de perfil e do protocolo verbal retrospectivo, que será abordado a seguir.

3.3.3 Protocolo verbal retrospectivo

Utilizado já há algumas décadas por áreas distintas das Ciências Humanas, tais como a Psicologia, a Antropologia e a Linguística, o protocolo verbal é um instrumento de coleta de dados de natureza qualitativa que consiste na verbalização de impressões por parte do participante de um dado experimento enquanto o sujeito

desempenha alguma tarefa orientada pelo pesquisador, fornecendo uma maior quantidade de dados linguísticos acerca do material analisado (ERICSSON; SIMON, 1980; GONÇALVES, 2001). Sendo assim, o objetivo da aplicação do protocolo é ampliar o escopo e a qualidade do produto analisado e permitir uma aproximação direcionada a parâmetros de intersubjetividade (ALVES, 2001; GONÇALVES, 2001).

No que se refere ao experimento aqui relatado, o protocolo verbal foi de cunho retrospectivo e foi composto por três perguntas (ver Anexo B) que objetivaram questionar as impressões dos participantes da amostra acerca do trailer assistido. Essas perguntas, por sua vez, foram passadas oralmente pelo pesquisador e respondidas pelos participantes da mesma maneira, sendo as respostas devidamente gravadas por meio da função *gravar* acoplada ao Google Meet e posteriormente transcritas.

Tal como ocorreu com a aplicação do questionário prospectivo de levantamento de perfil, o protocolo verbal retrospectivo também foi utilizado durante chamadas realizadas com cada participante pelo Google Meet.

3.4 DESENHO EXPERIMENTAL

Em virtude do atual contexto de restrições marcado pela pandemia de Covid-19, a coleta de dados foi integralmente realizada na modalidade virtual e todos os participantes foram individualmente contatados. A coleta contou com dois grupos de participantes:

a) o grupo um (G1), que ficou responsável por assistir ao trailer oficial da Netflix; e
b) o grupo dois (G2), que assistiu ao trailer com legendas que passaram por tradução automática e foram posteriormente pós-editadas.

Após assistirem ao trailer, foi solicitado aos participantes que preenchessem uma Escala Likert de cinco pontos e respondessem um protocolo verbal retrospectivo. O tempo inicialmente estimado para a coleta foi de 20 a 30 minutos.

A partir dessa distribuição, teve-se, portanto, um G1 e um G2, o que possibilitou a realização de uma análise intergrupos da recepção dos usuários no tocante às legendas dos trailers. Essa análise, por sua vez, foi empregada com o

intuito de observar as variações ocorridas em função das diferenças de estímulos oferecidos pelos dois tipos de legenda: a legenda oficial da Netflix e a legenda pós-editada.

Para a tradução automática e a posterior pós-edição das legendas que foram utilizadas no G2, foi utilizado o *software* Subtitle Edit (SE), que possui um leque de recursos profissionais para criar, ajustar e sincronizar as legendas e permite incorporar o uso da tradução automática ao seu sistema. As legendas foram pós-editadas por um profissional da área de pós-edição e posteriormente revisadas por um segundo profissional. Sobre isso, é importante destacar que ambos não precisaram se incomodar com o fator "pressão de tempo" para a realização da atividade, comumente empregado no mercado de tradução, e ambos receberam as devidas instruções para que o trailer com as legendas pós-editadas circulasse profissionalmente e de forma oficial.

3.5 PROCEDIMENTOS PARA COLETA DOS DADOS

Conforme dito na seção anterior, toda a coleta de dados foi realizada no formato virtual. Os participantes foram recrutados via *e-mail* e a coleta foi realizada por meio da plataforma de videoconferências Google Meet, que, por sua vez, é uma das várias ferramentas disponibilizadas pela UFPE por meio do pacote de serviços G Suite[4]. Esse leque de ferramentas foi de grande utilidade para a coleta pelo fato de apresentar um ambiente digital institucional seguro, integrado e com amplo espaço de armazenamento (UFPE, 2020).

Após o recrutamento dos participantes via formulário de pré-seleção seguindo-se os critérios já mencionados na subseção *Participantes*, foi feita a distribuição aleatória dos grupos 1 e 2. Feita essa divisão, cada participante recebeu por *e-mail* algumas instruções gerais acerca da realização do experimento seguido de um *link* do Doodle[5], uma plataforma de uso gratuito que serve para marcar datas e horários para futuras reuniões.

[4] As ferramentas do G Suite mais utilizadas para os fins desta pesquisa foram o Google Docs (para a transcrição dos protocolos verbais retrospectivos), Google Planilhas (para a tabulação de dados), Google Forms (para a elaboração de formulários e do TCLE), Google Meet (para as chamadas e gravações), Google Agenda (para marcação de horários) e Google Drive (para o armazenamento geral de arquivos).

[5] A plataforma encontra-se disponível em: https://doodle.com/en/.

Com as datas e horários de preferência dos participantes devidamente registrados na plataforma, as informações foram transferidas para o Google Agenda, e, por meio desta ferramenta, o *link* de acesso ao Google Meet foi enviado para cada participante. Além disso, também foi enviado, individualmente, o endereço de acesso ao Termo de Consentimento Livre e Esclarecido (TCLE) (ver Anexo A), para que cada participante pudesse se informar acerca do experimento e assinalar se consentia ou não em participar do experimento como voluntário, de forma anônima e não remunerada.

Na chamada via Google Meet realizada no dia e horário marcados, dava-se início à coleta de dados propriamente dita, com o participante recebendo, de início e oralmente, as informações do TCLE acerca das etapas da coleta. Após esse momento, era solicitada a permissão de gravação ao participante para que o experimento pudesse ser integralmente gravado, e, uma vez dada a permissão, a coleta era registrada para posterior tabulação e transcrição dos dados.

Na primeira etapa da coleta, o participante recebia, pelo *chat* do Google Meet, um *link* do Google Forms contendo o questionário prospectivo de levantamento de perfil para preenchimento. Na sequência, era enviado o *link* de acesso ao material (trailer legendado), e, após o participante acusar o término do trailer, um terceiro *link* com a Escala Likert era enviado. Em seguida, na etapa posterior ao preenchimento da escala, o participante respondia oralmente às questões do protocolo verbal retrospectivo, que eram igualmente verbalizadas pelo pesquisador. Por fim, uma vez concluídas as respostas do protocolo, a gravação era interrompida e a coleta era dada por encerrada.

Todos os dados de natureza qualitativa provenientes da coleta foram tabulados (questionários prospectivos e protocolos verbais retrospectivos) e transcritos (apenas os protocolos verbais retrospectivos) e serão analisados conforme detalhado na próxima seção.

3.6 METODOLOGIA DE ANÁLISE DOS DADOS

Como foi dito anteriormente, o experimento aqui relatado foi originalmente conduzido partindo-se do método da triangulação de dados (ALVES, 2001), que, por sua vez, envolve instrumentos de coleta de viés quantitativo e qualitativo. Os dados coletados foram resultantes do questionário prospectivo de levantamento de perfil,

do protocolo verbal retrospectivo e da escala Likert, sendo os dois primeiros instrumentos de viés predominantemente qualitativo e o último de viés quantitativo.

Os dados provenientes do questionário prospectivo foram coletados e organizados em uma planilha Excel com o intuito de promover um levantamento de perfil dos participantes com base em informações pessoais, linguísticas e profissionais, além de filtrar as respostas consideradas pré-requisitos para participar do experimento. As questões que solicitavam o grau de proficiência em Língua Inglesa e a preferência por filmes dublados ou legendados foram utilizadas como critério de seleção dos dados a serem analisados. Os dados do questionário foram apresentados no capítulo seguinte e, a partir deles, foi feita uma análise descritiva.

Além disso, a análise também se respaldou nos dados coletados por meio do protocolo verbal retrospectivo, que, por meio de três perguntas feitas oralmente, requereram informações acerca do motivo pelo qual foi marcado o valor na escala Likert, se houve algum aspecto técnico ou linguístico das legendas do trailer que teriam afetado na compreensão ou apreciação do participante e se este gostaria de fazer algum comentário geral acerca das legendas (ver Anexo B). Esses dados foram transcritos em doc. Word, tabulados em planilha, analisados e triangulados com os dados quantitativos provenientes da escala Likert.

Os dados advindos da escala Likert, por sua vez, foram igualmente tabulados em planilha Excel para fins de análise. Os valores numéricos (de 1 a 5) atribuídos a cada nível de satisfação do telespectador no tocante às legendas do trailer foram triangulados com os dados qualitativos oriundos do questionário e dos protocolos verbais e, a partir disso, foi estabelecida uma comparação entre os grupos 1 e 2 partindo-se das respostas coletadas por meio dos instrumentos.

Por fim, constata-se que o tratamento dos dados oriundos dos três instrumentos de coleta permitiu uma análise descritiva das diferenças entre as variáveis analisadas em cada grupo (1 e 2), a qual será apresentada no capítulo seguinte.

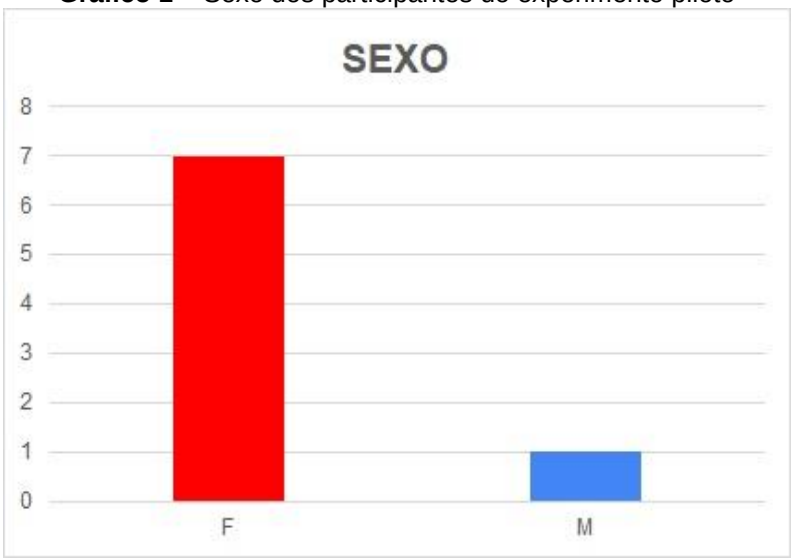
3 ANÁLISE E DISCUSSÃO DOS DADOS

Para os propósitos desta pesquisa, foram analisados os dados de um experimento piloto, realizado no GETRADTEC-UFPE, com 8 participantes acerca da recepção da legendagem auxiliada por TA. Os dados analisados foram provenientes de um questionário prospectivo de levantamento de perfil, uma escala Likert e um protocolo verbal retrospectivo, não havendo restrições de tempo para a realização das etapas do experimento por parte dos participantes.

Para garantir a confidencialidade das informações pessoais dos participantes, todos receberam denominações que vão do P01 ao P08, divididas entre grupo um (G1) e grupo dois (G2). Sendo assim, “P01_G1” indica o participante de número 1 do G1, “P01_G2” indica o participante de número 1 do G2, e assim sucessivamente.

De acordo com dados coletados por meio do questionário prospectivo de levantamento de perfil, houve uma considerável participação do público pertencente ao sexo feminino no experimento, como mostra o gráfico a seguir:

Gráfico 1 – Sexo dos participantes do experimento piloto



Fonte: O autor (2022).

Conforme apresenta o gráfico 1, dos 8 participantes recrutados, 7 foram do sexo feminino.

A média de idade dos 8 participantes é de 22,6 anos, sendo todos estudantes de Letras da UFPE a partir do segundo período com grau de proficiência em Língua

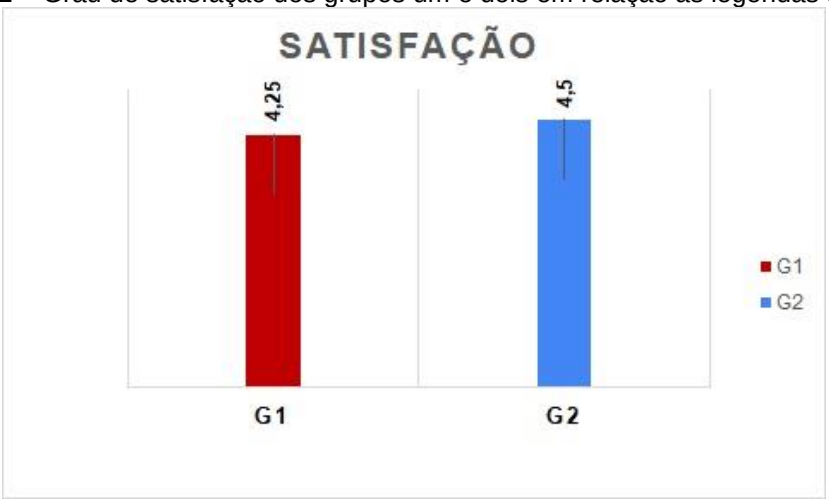
Inglesa B1 ou superior. Além disso, desses 8 participantes, 6 afirmaram estar cursando ou já ter cursado alguma disciplina de tradução[6].

Os dados provenientes do questionário prospectivo também revelaram que quase todos os participantes afirmaram preferir o consumo de filmes legendados em detrimento dos dublados, com exceção de P02_G2. No que se refere à frequência com que assiste a conteúdos audiovisuais legendados, 3 (P01_G1, P03_G1 e P03_G2) dos 8 participantes assinalaram que sempre consomem obras legendadas, enquanto que os demais participantes marcaram a opção “quase sempre”[7].

Além disso, todos os participantes do experimento assinalaram no questionário que costumam assistir a filmes legendados em plataformas de *streaming* (Netflix, Amazon Prime etc.), o que sugere uma grande concentração de espectadores nesse meio, conforme relatam Campos e Azevedo (2020). As outras opções assinaladas foram “filmes baixados da internet”, marcada por 7 participantes, “cinema”, assinalada por 6 participantes, “TV a cabo”, marcada por 2 participantes, e as duas últimas opções eram “nunca vejo filmes legendados” e “outros”, que não receberam nenhuma marcação.

No que se refere aos dados provenientes da escala Likert, percebe-se que o grau de satisfação em relação às legendas foi maior no G2, conforme apresenta o gráfico 2:

Gráfico 2 – Grau de satisfação dos grupos um e dois em relação às legendas do trailer



Fonte: O autor (2022).

[6] As disciplinas cursadas pelos participantes foram LE1023 – Estudos da Tradução (P01_G1; P02_G1; P04_G2; P04_G1), LE1037 – Ferramentas Tecnológicas de Tradução (P03_G1; P04_G1), LE1038 – Tradução de Textos Especializados (P03_G2) e LE1041 – Prática de Tradução Literária Inglês/Português (P04_G2).

[7] As demais opções passíveis de serem marcadas eram “às vezes”, “quase nunca” e “nunca”.

O gráfico 2 mostra que, no G1, a média de satisfação em relação às legendas foi de 4,25, enquanto que, no G2, a média foi de 4,5, ou seja, 0,25 décimos maior em comparação ao G1. A partir disso, percebe-se que os dados provenientes da escala apontam para uma melhor aceitabilidade das legendas traduzidas automaticamente e pós-editadas em comparação às legendas disponibilizadas pela Netflix, ainda que esse resultado não possa ser generalizado em virtude do número pequeno de participantes.

Essa tendência segue a mesma direção dos resultados de experimentos realizados pelo *Subtitling by Machine Translation (SUMAT)* sobre o potencial uso da pós-edição na legendagem, uma vez que, na ocasião do SUMAT, a partir da construção de *corpora* de legendas para treinar sistemas de tradução automática de base estatística (SMT) em diversos pares linguísticos, obteve-se um aumento da produtividade e da qualidade da tradução de legendas traduzidas automaticamente e pós-editadas (JI; OAKES, 2019).

Além disso, é importante salientar que a aceitabilidade da audiência no que se refere às legendas pós-editadas converge com as teorizações de Volk (2008) sobre a aplicabilidade da TA em legendas. Isso porque, com base no autor, a SMT se apresenta como uma ferramenta útil para o aumento da qualidade das legendas, sobretudo quando utilizada a partir de um conjunto robusto de *corpora* paralelos (VOLK, 2008).

Os dados provenientes do protocolo verbal retrospectivo, por sua vez, corroboram as marcações dos participantes na escala Likert. A princípio, quando comparados os tempos de transcrição dos grupos G1 e G2[8], é possível constatar que o G1 obteve um maior tempo de gravação em comparação ao G2, uma vez que os participantes do G1, quando questionados se houve aspectos técnicos e linguísticos das legendas que teriam afetado na *compreensão* ou *apreciação* do trailer, se mostraram mais incomodados com essas questões estruturais das legendas disponibilizadas pela Netflix em comparação aos participantes do G2, que responderam ao mesmo questionamento, assistiram às legendas traduzidas

[8] Todas as respostas dos grupos G1 e G2 coletadas e gravadas pela ferramenta de gravação do Google Meet foram posteriormente transcritas e tabuladas. Do G1, foram transcritos exatos vinte minutos e vinte segundos (00:20:20) contando com as respostas dos 4 participantes, enquanto que, no G2, foram transcritos doze minutos e vinte segundos (00:12:20). O tempo total de transcrições foi de trinta e dois minutos e quarenta segundos (00:32:40).

automaticamente e pós-editadas e, por terem apresentado menos incômodo, apontaram menos aspectos na gravação.

No G1, foram apontados alguns aspectos no que se refere aos parâmetros técnicos das legendas do trailer oficial da Netflix. Quando indagados se houve algum aspecto dessa natureza que tenha afetado a compreensão e apreciação do trailer, um participante apontou que houve falha de sincronização (P01_G1), um verbalizou que a fonte da legenda estava em um tamanho grande (P02_G1) e um apontou que houve partes em que as legendas tiveram uma duração curta (P02_G1), relatando que, de certa forma, isso o impedia de concluir a leitura.

Sobre o relato de P02_G1 acerca da duração curta das legendas, Martinez (2007) destaca que as limitações físicas do espaço da legenda na tela e o ritmo do discurso falado exigem uma redução considerável no texto, uma vez que o que determina o tempo de permanência de uma legenda na tela é o áudio original. Além disso, a autora ressalta que, no momento que identifica um enunciado oral na língua de partida, o telespectador direciona o olhar para a parte inferior da tela em busca da tradução do enunciado, o que explica o parâmetro de que a legenda deve surgir na tela cerca de 1/4 de segundo depois do início da fala original – tempo que o cérebro humano leva para captar o áudio e guiar o olhar para o local onde se localiza a legenda (KARAMITROGLOU, 1998 *apud* MARTINEZ, 2007, p. 36-37).

Entretanto, a questão da leiturabilidade das legendas não está atrelada apenas aos parâmetros técnicos da legendagem, mas também às camadas sociais e preferências da audiência no que diz respeito a obras audiovisuais. Acerca disso, Martinez (2007) ressalta que, no Brasil, apesar de a dublagem e a legendagem serem duas modalidades aceitas, a dublagem exerce um papel hegemônico na TV aberta, uma vez que o nível geral de instrução da população é baixo e, devido a esse fator socioeconômico, é possível que, a partir desse fator, o espectador realize um esforço cognitivo grande para ler as legendas com rapidez (MARTINEZ, 2007). Esses fatores apontados pela autora, por estarem diretamente atrelados também às preferências da audiência, podem influir na recepção de legendas.

Por outro lado, os parâmetros linguísticos das legendas, por sua vez, foram os mais pontuados no G1. Quando questionados se houve algum aspecto referente à tradução em si das legendas que teria afetado na compreensão e apreciação do trailer, um participante apontou a omissão de informações em virtude do tamanho das legendas (P01_G1), um mencionou a literalidade na tradução em certas partes

da legenda, o que causou certo estranhamento na leitura (P02_G1), e um relatou certa falta de intensidade em partes das legendas (P03_G1).

Sobre a questão da literalidade apontada pelo P02_G1, era esperado que esse aspecto fosse apontado no G2, uma vez que, neste grupo, as legendas do trailer foram elaboradas com o auxílio da TA. Por ter sido a literalidade apontada no G1, supõe-se que o profissional responsável pelas legendas oficiais do trailer tenha procurado atender às expectativas da audiência e, a partir disso, tenha optado por traduções que mais se aproximam da língua fonte (KOGLIN, 2008), ainda que nem todos os espectadores esperam uma tradução literal.

Ainda sobre os aspectos linguísticos das legendas relatados pelos participantes do G1, é importante mencionar, também, que 3 participantes (P01_G1; P02_G1; P04_G1) perceberam e relataram no protocolo verbal uma leve quebra de sentido em uma parte específica da legenda. No minuto 00:26 do trailer, um personagem pergunta ao outro o que ele sabe acerca do que está acontecendo na Etiópia (no original: "*What do you know about what's happening in Etiopia?"*), sendo que, na legenda em PT-BR, a mensagem passada foi "O que sabe sobre a Etiópia?". A resposta dada pelo outro personagem foi que o que está acontecendo lá "É um genocídio, e não fazem nada porque é na África" (no original: "*It's another bloody genocide but nobody gives a shit because it's in Africa*").

Quando perguntado sobre se houve algum aspecto linguístico das legendas que teria afetado na compreensão ou apreciação do trailer, o P01_G1 relatou o seguinte sobre o minuto 00:26:

> Que eu lembro agora é quando ele pergunta... 'o que você sabe sobre...' o lugar, eu esqueci agora. É... ele (tu #) perguntou pra pessoa 'você sabe sobre tal lugar?' que eu não lembro qual lugar agora, mas enfim, é... e aí... na verdade é # isso é o que foi colocado na legenda, né, mas é... ela, ele tinha perguntado o que = a [...] o que = a pessoa sabia sobre o que estava acontecendo naquele momento, naquele lugar, e não necessariamente sobre o lugar, né, e... acho que ficou, é... generalizou de certa forma a informação assim (P01_G1)[9].

O mesmo questionamento foi feito ao P02_G1, que, por sua vez, relatou:

[9] Há, nas transcrições, alguns sinais gráficos que são próprios do modelo de transcrição utilizado nesta pesquisa. São eles: "#", empregado para sinalizar as palavras que não foram compreendidas a partir da gravação; "[...]", utilizado para palavras e frases sem significado relevante; "=", utilizado para representar palavras que foram ditas mais de duas vezes; "()", palavras entre parênteses que representam a intervenção do entrevistador; "...", usado para indicar pausas; e "(?)", utilizado para sinalizar dúvidas de entendimento por parte do entrevistador.

> Teve um trecho só que me chamou a atenção, que era... muito ruim assistir uma vez só! Mas ele falou... 'você sabe... o que... você sabe o que está acontecendo com a Etiópia?', e a legenda dizia 'o que você sabe sobre a Etiópia?'. E aí a resposta era 'um genocídio', num sei o quê... que aí fica... na legenda fica completamente... fica bem estranho, né, afinal... 'o que você sabe sobre a Etiópa?', 'um genocídio'. Era o que tinha na legenda, quando a pergunta era '[...] o que você sabe sobre o que está acontecendo na Etiópia?', né... o sentido fica... quebrado aí (P02_G1).

O P04_G1 também notou a quebra de sentido no referido minuto do trailer e declarou:

> Por exemplo, teve uma parte que ele diz, é... 'você sabe sobre a Etiópia', só que... na legenda. Só que no áudio original ele pergunta 'você sabe o que está acontecendo agora', algo assim, 'na Etiópia?' (P04_G1).

Essa percepção acerca da quebra de sentido da legenda mencionada foi, no G1, a mais pontuada em termos de frequência nos relatos dos participantes, que, por terem domínio do par linguístico EN-PT-BR, puderam associar o áudio original à legenda empregada e, a partir dessa associação, apontaram o desvio semântico. Essa percepção dos participantes do G1 traz novamente à tona a concepção equivocadamente implícita do espectador de que a legendagem deve necessariamente ser literal, caracterizada pela transposição de significados "fechados" e desprovidos de contexto (KOGLIN, 2008).

No que se refere à cor branca empregada nas legendas, não houve queixas. O P01_G1 relatou que a cor da legenda facilitou a visibilidade e a leitura. Além disso, também não foram apontados incômodos em relação ao número de caracteres e de linhas empregado nas legendas do trailer.

Os participantes do G2, por sua vez, apresentaram menos incômodo no tocante aos aspectos técnicos e linguísticos das legendas traduzidas automaticamente e posteriormente pós-editadas por um tradutor humano. Especificamente sobre os parâmetros técnicos, um participante relatou incômodo em relação à fonte das legendas (P01_G2) e um participante acusou falhas de sincronização nas legendas em certos momentos (P02_G2).

Quando questionado sobre se houve algum aspecto técnico das legendas que teria afetado na compreensão ou apreciação do trailer, o P01_G2 afirmou ter tido uma "sensação estranha" com o formato das legendas, conforme relato a seguir:

> Humm... não, eu sou acostumada assim a ver legendas, mas eu acredito que o formato da legenda em si poderia ser um pouquinho diferente. Eu não lembro agora exatamente, tava... sei lá, em itálico, mas às vezes parecia estar... deu essa sensação estranha, mas foi... tranquila, assim... o... design digamos assim, da = legenda e a quantidade de linhas bem tranquila também e a duração boa, deu pra = associar, né, ouvir o áudio e ler, foi muito tranquilo (P01_G2).

Não foi possível depreender com clareza o que teria ocasionado essa sensação de estranhamento em virtude da quebra de ideias na exposição verbal do participante, mas supõe-se que o relato tenha se referido à fonte das legendas, uma vez que “formato” remete a “forma”, ou, em questões de legendagem, à fonte utilizada. Além disso, não houve, nas legendas do trailer assistido pelo G2, nenhum trecho que tenha recebido marcação em itálico.

Quando questionados se houve algum aspecto linguístico das legendas que teria afetado na compreensão ou apreciação do trailer, os participantes do G2 não reportaram incômodos, com exceção do P04_G2, que relatou:

> Eu marquei cinco, como ‘bastante satisfeito’. Eu acho que, de uma forma geral, foi bem satisfatório. Teve só uma vez eu acho que eu fiquei ‘poxa, a intenção da fala do personagem não foi muito bem descrita’, que é na parte que o personagem fala... como se fosse um palavrão entre aspas em inglês, e aí a legenda suavizou um pouco isso (P04_G2).

É percebível que, a partir do relato supracitado, o P04_G2 ressaltou que, em certa parte das legendas, houve a atenuação de linguagem de baixo calão, que, por sua vez, apresenta-se como uma modalidade cujo processo tradutório é desafiador e repleto de restrições (KOGLIN, 2009). Supõe-se, portanto, que o P04_G2 utilizou a ideia da suavização como uma norma geral da legendagem, já que o ato de amenizar a linguagem de baixo calão nas legendas se trata de uma característica marcante utilizada pelas empresas de produção audiovisual em contexto nacional.

À parte desses relatos, os participantes do G2 demonstraram um grau de satisfação um pouco maior em relação às legendas do trailer. Além disso, não houve relatos de incômodo em questões de omissão de informações, literalidade ou falta de intensidade das legendas, tal como foi relatado por participantes do G1.

A partir do exposto, percebe-se que tanto os dados provenientes da escala Likert quanto as respostas coletadas por meio do protocolo verbal retrospectivo apontam para uma maior aceitabilidade das legendas pós-editadas em comparação às legendas oficiais da Netflix.

No G1, os participantes mostraram-se incomodados com muitos parâmetros técnicos (sincronização, tamanho e duração) e linguísticos (omissão de informações, literalidade, falta de intensidade e quebra de sentido) das legendas, o que foi imediatamente perceptível a partir do tempo e conteúdo das transcrições dos protocolos. Aliado a isso disso, é importante ressaltar que apenas no G1 foi marcado um valor “3” (razoavelmente satisfeito) na escala Likert.

No G2, por sua vez, foram poucos os relatos de incômodo por parte dos participantes em questões de parâmetros técnicos (fonte e sincronização das legendas) e linguísticos (omissão de informação) das legendas e a maioria considerou que as legendas estavam satisfatórias. Isso indica, portanto, que as legendas pós-editadas foram bem recebidas em termos de parâmetros técnicos e linguísticos em comparação às legendas oficiais da Netflix.

Sendo assim, a partir da análise e interpretação dos dados provenientes do questionário prospectivo, da escala Likert e do protocolo verbal retrospectivo, o experimento aqui relatado tende a mostrar a importância de se testar a viabilidade de legendas auxiliadas por TA e posteriormente pós-editadas com um número maior de participantes, além de se aproximar das novas tendências tecnológicas aplicadas à legendagem no mercado de tradução.

5 CONSIDERAÇÕES FINAIS

O experimento relatado nas seções anteriores buscou investigar a recepção de legendas interlinguais no par linguístico EN/PT-BR auxiliadas por TA, além de analisar, a partir desse propósito geral, qual tipo de legenda se mostra mais adequada para a audiência e qual o nível de aceitabilidade dos espectadores no tocante às legendas pós-editadas.

Com base nos resultados provenientes dos dados da Escala Likert e do protocolo verbal retrospectivo, percebe-se que o experimento piloto tende a apontar para uma melhor aceitabilidade do público no que se refere às legendas traduzidas automaticamente e posteriormente pós-editadas. Sendo assim, tal resultado tende a sugerir o uso da tradução automática no contexto da legendagem, mas apenas com os resultados finais extraídos da análise dos dados da coleta definitiva – realizada com um número amostral maior – é que essa tendência poderá ser confirmada e as perguntas de pesquisa poderão ser respondidas em sua totalidade.

Mesmo com a impossibilidade de se alcançar respostas concretas às perguntas de pesquisa devido ao fato de o experimento ter um número baixo de participantes, é importante destacar que o experimento se manteve associado ao objetivo geral estabelecido na Introdução, que foi o de investigar a recepção de legendas interlinguais no par linguístico IN/PT-BR com o auxílio da TA. Da mesma maneira, a partir dos resultados obtidos e analisados, constatou-se que as legendas pós-editadas se mostraram mais adequadas para o telespectador em termos de aceitabilidade, confirmando, portanto, a consecução dos objetivos específicos.

Apesar disso, é importante ressaltar que a pesquisa aqui apresentada enfrentou algumas limitações não apenas por se tratar de um experimento piloto realizado com o intuito de prover ajustes metodológicos para uma posterior coleta definitiva, mas também em virtude do atual contexto marcado pela pandemia de Covid-19. O fator comunicação, por exemplo, foi um desses entraves, uma vez que o recrutamento e a realização das etapas de coleta foram realizados 100% *on-line* e, em consequência disso, não houve contato físico, que é de suma importância para se ter melhores impressões sobre o comportamento dos participantes frente aos instrumentos de coleta.

Ademais, outra limitação da pesquisa diz respeito a questões de infraestrutura, em partes também relacionadas ao contexto pandêmico. Isso porque,

originalmente, a coleta de dados a partir da triangulação se constitui partir de dados qualitativos e quantitativos provenientes de questionários prospectivos, escalas Likert, protocolos verbais, programas de informática (a exemplo do Translog) e de rastreadores oculares. Porém, em virtude da ausência de alguns instrumentos na UFPE e da pandemia de Covid-19, apenas três instrumentos poderiam e foram utilizados na coleta: o questionário prospectivo, a escala Likert e o protocolo verbal retrospectivo.

Apesar dessas limitações, o experimento realizado no GETRADTEC-UFPE e apresentado nesta monografia trouxe à luz uma perspectiva teórico-metodológica ainda pouco investigada nos estudos de recepção de legendas interlinguais no Brasil, uma vez que a TA, apesar de ser bastante abordada em pesquisas de diferentes naturezas, ainda não é amplamente utilizada em experimentos empíricos sobre a recepção de legendas e no mercado de tradução em contexto nacional. Daí surge a necessidade de se promover pesquisas nesse escopo, pois, tal como ocorre com textos escritos, a TA também se apresenta como uma ferramenta de grande utilidade para a legendagem, podendo, inclusive, contribuir para o aumento da qualidade das legendas e para a boa recepção por parte da audiência.

Sendo assim, espera-se que o entendimento acerca da expectativa e da aceitabilidade do público em relação a legendas pós-editadas e de seu potencial uso no mercado possa ser ampliado, uma vez que esse tipo de legenda se mostra bastante promissor em termos de qualidade e recepção. Isso, porém, só poderá ser confirmado com a realização de futuras pesquisas empírico-experimentais sobre a recepção de legendas auxiliadas por TA.

Por fim, concluiu-se que o experimento contribuiu para o entendimento acerca da expectativa e da aceitabilidade do público em relação a legendas pós-editadas e do seu potencial uso no mercado de tradução, além de ter colaborado para a ampliação do leque de pesquisas referentes aos estudos de recepção de legendas interlinguais auxiliadas por TA, área pouco investigada no Brasil.

REFERÊNCIAS

ADÃO, C. M. C. J. *Tecnologias de streaming em contextos de aprendizagem*. 2006. 181 f. Dissertação (Mestrado em Sistemas de Informação) – Escola de Engenharia, Universidade do Minho, Braga, 2006. Disponível em: http://repositorium.sdum.uminho.pt/handle/1822/6400. Acesso em: 21 abr. 2022.

ALFARO, C.; DIAS, M. C. P. Tradução automática: uma ferramenta de auxílio ao tradutor. *Cadernos de Tradução*, Santa Catarina, v. 1, n. 3, p. 369-390, 1998. Disponível em: https://dialnet.unirioja.es/servlet/articulo?codigo=4925639. Acesso em: 22 abr. 2020.

ALVES, F. A triangulação como opção metodológica em pesquisas empírico-experimentais em tradução. *In*: PAGANO, A. (org.). *Metodologias de pesquisa em tradução*. Belo Horizonte: Faculdade de Letras da UFMG, 2001. p. 69-92.

ARANTES, S. Brasileiro prefere dispensar legendas e ver filmes dublados na TV por assinatura. *Folha de S. Paulo*, São Paulo, 7 jul. 2013. Disponível em: https://m.folha.uol.com.br/ilustrada/2013/07/1306753-brasileiro-prefere-dispensar-legendas-e-ver-filmes-dublados-na-tv-por-assinatura.shtml. Acesso em: 22 fev. 2022.

CAMPOS, G. C.; AZEVEDO, T. A. Subtitling for streaming platforms: new technologies, old issues. *Cadernos de Tradução*, Florianópolis, v. 40, n. 3, p. 222-243, 2020. DOI: https://doi.org/10.5007/2175-7968.2020v40n3p222. Disponível em: https://periodicos.ufsc.br/index.php/traducao/article/view/2175-7968.2020v40n3p222. Acesso em: 20 abr. 2022.

CARVALHO, F. M. Judeus etíopes: origens e destinos. *Cadernos de Língua e Literatura Hebraica*, São Paulo, n. 11, p. 1-21, 2013. DOI: https://doi.org/10.11606/issn.2317-8051.cllh.2013.83490. Disponível em: https://www.revistas.usp.br/cllh/article/view/83490. Acesso em: 20 abr. 2022.

CHILE, D. M. I. The sitcom revisited: the translation of humor in a polysemiotic text. *Cadernos de Tradução*, Florianópolis, v. 1, n. 4, p. 167-204, 1999. Disponível em: https://antigo.periodicos.ufsc.br/index.php/traducao/article/view/5531/4990. Acesso em: 2 abr. 2022.

COUNCIL OF EUROPE. *Common European framework of reference for languages*: learning, teaching, assessment. Strasbourg: Council of Europe Publishing, 2020. Disponível em: https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4. Acesso em: 19 mar. 2022.

DÍAZ-CINTAS, J.; PIÑERO, R. B. Audiovisual translation in a global context. *In*: DÍAZ-CINTAS, J.; PIÑERO, R. B. *Audiovisual translation in a global context*: mapping an ever-changing landscape. Londres: Palgrave Macmillan, 2015. p. 1-10.

DÍAZ-CINTAS, J.; REMAEL, A. *Subtitling*: concepts and practice. Londres/Nova Iorque: Routledge, 2021. 292 p.

ERICSSON, K. A.; SIMON, H. A. Verbal reports as data. *Psychological Review*, Washington, v. 87, n. 3, p. 215-251, 1980. DOI: https://doi.org/10.1037/0033-295X.87.3.215. Disponível em: https://content.apa.org/record/1980-24435-001. Acesso em: 20 abr. 2022.

FONSECA, N. B. L. Analysing the impact of TAPs on temporal, technical and cognitive effort in monolingual post-editing. *Perspectives*: Studies in Translation Theory and Practice, Amsterdã/Filadélfia, v. 27, n. 4, p. 552-588, 2019. DOI: https://doi.org/10.1080/0907676X.2019.1597909. Disponível em: https://www.tandfonline.com/doi/full/10.1080/0907676X.2019.1597909. Acesso em: 20 abr. 2022.

FONSECA, N. B. L.; GONÇALVES, J. L. V. R.; ZAMPIER, P. Investigando o esforço cognitivo, o conhecimento sobre tradução e a satisfação na tradução libras-português. *Cadernos de Tradução*, Florianópolis, v. 39, n. 3, p. 20-49, 2019. DOI: https://doi.org/10.5007/2175-7968.2019v39n3p20. Disponível em: https://periodicos.ufsc.br/index.php/traducao/article/view/2175-7968.2019v39n3p20. Acesso em: 20 abr. 2022.

GAMBIER, Y. Challenges in research on audiovisual translation. *In*: PYM, A.; PEREKRESTENKO, A. (ed.). *Translation research projects 2*. Tarragona: Intercultural Studies Group, 2009. p. 17-25.

GONÇALVES, J. L. V. R. Pesquisas empírico-experimentais em tradução: os protocolos verbais. *In*: PAGANO, A. (org.). *Metodologias de pesquisa em tradução*. Belo Horizonte: Faculdade de Letras da UFMG, 2001. p. 13-39.

GOTTLIEB, H. Subtitling. *In*: BAKER, M. (org.). *Routledge encyclopedia of translation studies*. Londres/Nova Iorque: Routledge, 2005. p. 244-248.

GUERBEROF, A. Productivity and quality in MT post-editing. *In*: MACHINE TRANSLATION SUMMIT – BEYOND TRANSLATION MEMORIES: NEW TOOLS FOR TRANSLATORS WORKSHOP, 12., 2009, Ottawa. *Anais* […]. Ottawa: Machine Translation Summit, 2009. p. 1-9. Disponível em: https://aclanthology.org/2009.mtsummit-btm.7.pdf. Acesso em: 16 abr. 2022.

GUILLOT, M. N. Subtitling on the cusp of its futures. *In*: PÉREZ-GONZÁLEZ, L. (ed.). *The Routledge handbook of audiovisual translation*. Londres/Nova Iorque: Routledge, 2018. p. 31-47.

HILL, A. Media audiences and reception studies. *In*: DI GIOVANNI, E.; GAMBIER, Y. (ed.). *Reception studies and audiovisual translation*. Amsterdã/Filadélfia: John Benjamins Publishing Company, 2018. p. 3-20.

JAKOBSON, R. Aspectos linguísticos da tradução. *In*: JAKOBSON, R. *Linguística e comunicação*. 24. ed. São Paulo: Editora Cultrix, 2007. p. 63-72.

JI, M.; OAKES, M. (ed.). *Advances in empirical translation studies*: developing translation resources and technologies. Cambridge: Cambridge University Press, 2019. 285 p.

KENNY, D. The ethics of machine translation. *In*: NEW ZEALAND SOCIETY OF TRANSLATORS AND INTERPRETERS ANNUAL CONFERENCE, 20., 2011, Auckland. *Anais* [...]. Auckland: New Zealand Society of Translators and Interpreters, 2011. p. 1-9. Disponível em: http://doras.dcu.ie/17606/1/The_Ethics_of_Machine_Translation_pre-final_version.pdf. Acesso em: 10 abr. 2021.

KENNY, D. Machine translation. *In*: BAKER, M.; SALDANHA, G. (ed.). *Routledge encyclopedia of translation studies*. 3. ed. Nova Iorque: Routledge, 2020. p. 305-310.

KOGLIN, A. *A tradução de metáforas geradoras de humor na série televisiva Friends*: um estudo de legendas. 2008. 99 f. Dissertação (Mestrado em Estudos da Tradução) – Universidade Federal de Santa Catarina, Florianópolis, 2008. Disponível em: https://repositorio.ufsc.br/bitstream/handle/123456789/90889/250634.pdf?sequence=1. Acesso em: 7 maio 2022.

KOGLIN, A. A tradução da linguagem de baixo calão em obras fílmicas: um estudo baseado em corpus. *In*: II JORNADAS INTERNACIONALES DE TRADUCTOLOGÍA, 2., 2009, Córdoba. *Anais* […]. Córdoba: Centro de Investigación en Traducción, 2009. p. 1-8.

KOGLIN, A. O papel da competência leitora na compreensão da linguagem metafórica em obras legendadas. *In*: CONGRESO INTERNACIONAL DE LA ALFAL, 16., 2011, Alcalá de Henares. *Anais* [...]. Alcalá de Henares: Asociación de Linguística y Filología de América Latina, 2011. p. 2929-2934. Disponível em: http://www.mundoalfal.org/alfal2011/index.html#/pdf/333alfal.pdf. Acesso em: 21 fev. 2022.

KOGLIN, A. An empirical investigation of cognitive effort required to post-edit machine translated metaphors compared to the translation of metaphors. *The International Journal for Translation & Interpreting Research*, [*S. l.*], v. 7, n. 1, p. 126-141, 2015a. Disponível em: http://www.trans-int.org/index.php/transint/article/view/369/188. Acesso em: 20 abr. 2021.

KOGLIN, A. *Efeitos cognitivos e esforço de processamento de metáforas em tarefas de pós-edição e de tradução humana*: uma investigação processual à luz da Teoria da Relevância. 2015. 196 f. Tese (Doutorado em Estudos Linguísticos) – Universidade Federal de Minas Gerais, Belo Horizonte, 2015b. Disponível em: https://repositorio.ufmg.br/bitstream/1843/MGSS-A9ZS2Y/1/tese_arlene_koglin.pdf. Acesso em: 23 fev. 2022.

KOVAČEVIĆ, M. Post-editing of machine translation output with and without source text. *Hieronymus*, Zagreb, n. 1, p. 82-104, 2014. Disponível em: http://www.ffzg.unizg.hr/hieronymus/wp-content/uploads/2014/09/Hieronymus_1_2014_Kovacevic.pdf. Acesso em: 21 abr. 2022.

MACHINE Translation and Professional Translators Community. *AMTA*, Washington, [210-?]. Disponível em: https://amtaweb.org/tag/translators/. Acesso em: 16 abr. 2022.

MARTINEZ, S. L. *Tradução para legendas*: uma proposta para a formação de profissionais. 2007. 100 f. Dissertação (Mestrado em Letras) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2007. Disponível em: https://www.maxwell.vrac.puc-rio.br/colecao.php?strSecao=resultado&nrSeq=10689@1. Acesso em: 5 maio 2022.

MARTINS, R. T. Tradução automática. *Todas as Letras* – Revista de Língua e Literatura, São Paulo, v. 10, n. 2, p. 148-169, 2008. Disponível em: http://editorarevistas.mackenzie.br/index.php/tl/article/view/455/271. Acesso em: 21 maio 2022.

MESA-LAO, B. *Introduction to post-editing* – the CasMaCat GUI. Copenhague: SEECAT Project, 2013. Disponível em: http://bridge.cbs.dk/projects/seecat/material/hand-out_post-editing_bmesa-lao.pdf. Acesso em: 28 abr. 2022.

MISSÃO no Mar Vermelho. Direção: Gideon Raff. Santa Clara: Netflix, 2019. Disponível em: https://www.netflix.com/br/title/80240537. Acesso em: 21 fev. 2022.

NIKOLIĆ, K. Reception studies in audiovisual translation – interlingual subtitling. *In*: DI GIOVANNI, E.; GAMBIER, Y. (ed.). *Reception studies and audiovisual translation*. Amsterdã/Filadélfia: John Benjamins Publishing Company, 2018. p. 179-198.

ORREGO-CARMONA, D. *The reception of (non)professional subtitling*. 2015. 317 f. Tese (Doctorate in Translation Studies) – Department of English and German Studies, Universitat Rovira i Virgili, Tarragona, 2015.

REBOLLO-COUTO, L.; SILVA, L. P. N.; SILVA, C. G. Tradução audiovisual: estratégias pragmáticas e conversacionais americanas e europeias na legendagem das formas de tratamento nominais. *Caracol*, São Paulo, n. 14, p. 274-307, 2017. DOI: https://doi.org/10.11606/issn.2317-9651.v0i14p274-307. Disponível em: https://www.revistas.usp.br/caracol/article/view/131712. Acesso em: 20 abr. 2022.

ROBINSON, D. The divine inspiration of translation. *In*: ROBINSON, D. *Who translates?* Translator subjectivities beyond reason. Nova Iorque: State University of New York Press, 2001. p. 36-66.

UNIVERSIDADE FEDERAL DE PERNAMBUCO (UFPE). O que é o G Suite. *UFPE*, Recife, 2020. Disponível em: https://www.ufpe.br/gsuite. Acesso em: 5 mar. 2022.

VOLK, M. The automatic translation of film subtitles: a machine translation success story? *In*: NIVRE, J.; DAHLLÖF, M.; MEGYESI, B. (org.). *Resourceful language technology*: festschrift in honor of Anna Sågvall Hein. Uppsala: Zurich Open Repository and Archive, 2008. p. 202-214. Disponível em: https://www.zora.uzh.ch/id/eprint/8817/2/Volk__MT_of_SubtitlesV.pdf. Acesso em: 29 abr. 2022.

WILLIAMS, J.; CHESTERMAN, A. *The map*: a beginner's guide to doing research in translation studies. Manchester: St. Jerome Publishing Company, 2002. 150 p.

APÊNDICE A – Questionário prospectivo de levantamento de perfil

Questionário Prospectivo de Levantamento de Perfil

(Adaptado de KOGLIN, 2015b)

1. Nome completo: *

2. E-mail: *

3. Sexo: *

- ○ Feminino
- ○ Masculino
- ○ Outro

4. Idade: *

5. Qual curso universitário você está frequentando? *

- ○ Letras-Bacharelado
- ○ Letras-Licenciatura em Português
- ○ Letras-Licenciatura em Inglês
- ○ Letras-Licenciatura em Espanhol
- ○ Letras-Licenciatura em Francês

6. Você está cursando ou já cursou alguma disciplina da área de tradução? *

- ○ Sim
- ○ Não

7. Se você respondeu "sim" à questão anterior, especifique a(s) disciplina(s):

8. Qual é o seu idioma de tradução ou sua segunda língua? *

- ☐ Espanhol
- ☐ Inglês
- ☐ Francês
- ☐ Italiano
- ☐ Alemão
- ☐ Outros

9. Se você marcou "Outros" na opção anterior, por favor, indique qual idioma:

10. Indique o seu grau de proficiência em Língua Inglesa: *

- ○ A1 - Beginner

- ○ A2 - Elementary
- ○ B1 - Intermediate
- ○ B2 - Upper Intermediate
- ○ C1 - Advanced
- ○ C2 - Proficient

11. Você prefere filmes legendados ou dublados? *

- ○ Legendados
- ○ Dublados
- ○ Não tenho preferência

12. Com que frequência você assiste a filmes legendados? *

- ○ Sempre
- ○ Quase sempre
- ○ Às vezes
- ○ Quase nunca
- ○ Nunca

13. Onde costuma assistir a filmes legendados? *

- ☐ Cinema
- ☐ Streaming (Netflix, Amazon Prime etc.)
- ☐ TV a cabo
- ☐ Filmes baixados da internet
- ☐ Nunca vejo filmes legendados
- ☐ Outros

ANEXO A – Termo de Consentimento Livre e Esclarecido (TCLE)

Termo de Consentimento Livre e Esclarecido (TCLE)

Caro(a) participante,

Convidamos você a participar como voluntário(a) da pesquisa “Recepção da legendagem auxiliada por tradução automática: um estudo empírico-experimental”, que está sob a responsabilidade da pesquisadora Arlene Koglin. Seu endereço é Av. da Arquitetura, S/n - Campus Universitário, Recife - PE, CEP: 50740-550 - Telefone: +55 (81) 2126-8307 / +55 (48) 99963-4028, E-mail: arlene.koglin@ufpe.br.

Também participa desta pesquisa o pesquisador: João Gabriel Pereira da Silveira. Seu endereço é Rua Lígia Gomes, 331 - Ouro Preto - Olinda-PE. Telefone para contato: +55 81 99508-3349, E-mail: joao.gabrielsilveira@ufpe.br, que está sob a orientação de: Arlene Koglin, Telefone: +55 (48) 99963-4028, e-mail: arlene.koglin@ufpe.br.

O objetivo deste estudo é avaliar a viabilidade da legendagem auxiliada por tradução automática. Se concordar em participar deste estudo, que tem previsão de duração de 15-20 minutos, você será solicitado a realizar as seguintes tarefas: 1) responder a um questionário que aborda dados pessoais e/ou profissionais; 2) assistir um trailer; 3) preencher uma Escala Likert de 5 pontos e 4) responder um protocolo verbal com 3 perguntas. Todo material coletado será catalogado com um número de referência para preservar a confidencialidade de seus dados pessoais e será analisado de acordo com os fundamentos teóricos e os métodos de análise desta pesquisa.

Todas as informações desta pesquisa serão confidenciais e serão divulgadas apenas em eventos ou publicações científicas sem identificação dos voluntários, a não ser entre os responsáveis pelo estudo, sendo assegurado o sigilo sobre a sua participação. Os dados coletados nesta pesquisa (questionários, protocolos verbais (gravados) e escalas de Likert) ficarão armazenados em computador pessoal e na nuvem (Google Drive Institucional), sob a responsabilidade da pesquisadora profa. Arlene Koglin, no endereço https://drive.google.com/, pelo período mínimo de 5 anos após o término da pesquisa.

Sua participação é voluntária, anônima e gratuita, e não apresenta riscos à sua saúde física. Independentemente do motivo, você tem direito de retirar seu consentimento para participar da pesquisa em qualquer momento, inclusive sem justificativa específica. Não estão previstos benefícios diretos, porém sua participação terá benefícios indiretos como a contribuição para a formação de tradutores e para a acessibilidade.

Em caso de dúvidas relacionadas aos aspectos éticos deste estudo, você poderá consultar o Comitê de Ética em Pesquisa Envolvendo Seres Humanos da UFPE no endereço: (Avenida da Engenharia s/n – 1º Andar, sala 4 - Cidade Universitária, Recife-PE, CEP: 50740-600, Tel.: (81) 2126.8588 – e-mail: cepccs@ufpe.br).

Declaro que li o Termo de Consentimento e compreendi o objetivo do estudo. A explicação que recebi menciona os riscos e benefícios, bem como entendi que sou livre para interromper minha participação a qualquer momento, sem justificar minha decisão e sem qualquer prejuízo para mim. Portanto:

o (SIM), manifesto meu livre consentimento em participar como voluntária/o, de forma anônima e não remunerada do projeto acima mencionado.

o (NÃO), não consinto em participar como voluntária/o, de forma anônima e não remunerada do projeto acima mencionado.

Se quiser receber os resultados da pesquisa, informe seu e-mail aqui:

ANEXO B – Protocolo Verbal Retrospectivo

Protocolo verbal retrospectivo

1. Coleta com estudantes (grupos um e dois)

1. Você poderia explicar o porquê assinalou este valor na Escala Likert?
2. Houve algum aspecto técnico (número de linhas, cor da legenda, tamanho da legenda, duração da legenda) ou linguístico (tradução) que afetou a sua compreensão e a apreciação do trailer? Explique.
3. Gostaria de fazer algum outro comentário em relação às legendas do trailer?