



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE HUMANIDADES
DEPARTAMENTO DE LETRAS ESTRANGEIRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS DA TRADUÇÃO

SAMUEL LEVI SILVA DE OLIVEIRA

**A INFLUÊNCIA DA TRADUÇÃO DOS EFEITOS SONOROS DURANTE A
RECEPÇÃO DE LEGENDAS PARA SURDOS E ENSURDECIDOS (LSE) POR
ESPECTADORES SURDOS EM UM FILME DE AÇÃO**

FORTALEZA

2022

SAMUEL LEVI SILVA DE OLIVEIRA

A INFLUÊNCIA DA TRADUÇÃO DOS EFEITOS SONOROS DURANTE A RECEPÇÃO
DE LEGENDAS PARA SURDOS E ENSURDECIDOS (LSE) POR ESPECTADORES
SURDOS EM UM FILME DE AÇÃO

Defesa de Mestrado apresentado ao Programa de Pós-graduação em Estudos da Tradução (POET) da Universidade Federal do Ceará (UFC) como requisito parcial à obtenção do título de Mestre em Estudos da Tradução. Área de concentração: Processos de Retextualização.

Orientadora: Profa. Dra. Patrícia Araújo Vieira
Coorientadora: Profa. Dra. Ana Katarinna Pessoa do Nascimento

FORTALEZA

2022

SAMUEL LEVI SILVA DE OLIVEIRA

A INFLUÊNCIA DA TRADUÇÃO DOS EFEITOS SONOROS DURANTE A RECEPÇÃO
DE LEGENDAS PARA SURDOS E ENSURDECIDOS (LSE) POR ESPECTADORES
SURDOS EM UM FILME DE AÇÃO

Defesa de Mestrado apresentado ao Programa de Pós-graduação em Estudos da Tradução (POET) da Universidade Federal do Ceará (UFC) como requisito parcial à obtenção do título de Mestre em Estudos da Tradução. Área de concentração: Processos de Retextualização.

Aprovada em: 24/01/2022.

BANCA EXAMINADORA

Prof. Dr. José Ferrari Neto
Universidade Federal da Paraíba (UFPB)

Prof. Dra. Pâmela Freitas Pereira Toassi
Universidade Federal do Ceará (UFC)

Prof. Dra. Alexandra Frazão Seoane
Universidade Estadual do Ceará (UECE)

Profa. Dra. Patrícia Araújo Vieira (Orientadora)
Universidade Federal do Ceará (UFC)

Profa. Dra. Ana Katarinna Pessoa do Nascimento (Coorientadora)
Programa de Pós-Graduação em Estudos da Tradução (POET)

AGRADECIMENTOS

Antes de qualquer agradecimento pessoal, devo ser grato àquilo que nos permite trilhar nossos caminhos tortuosos, mas brilhantes e particulares, durante tão breve tempo, a natureza. Sua vastidão, suas leis sutis e seus mistérios permitiram a nós humanos estarmos aqui desfrutando deste pequeno pálido ponto azul por gerações e mais gerações. Sou grato também a todos gigantes que nos precederam e que desbravaram o inóspito, o desconhecido e que dedicaram suas vidas a desenvolver a mais bela de nossas criações, a ciência. A mesma ciência que nos permitirá alcançar o que ainda sequer sonhamos que exista. Meu fascínio e reverência será eterno enquanto eu durar.

Agradeço de coração cheio a minha orientadora Patrícia Vieira que abraçou um forasteiro e deu a ele um norte a seguir. Sou grato pela sua dedicação, puxões de orelha e conselhos fundamentalmente necessários para que este trabalho tomasse forma e conteúdo.

Devo agradecer também a minha coorientadora Katarinna Nascimento que proveu a coluna central desta pesquisa e que, sem a qual, você não estaria lendo este trabalho. Também agradeço as suas inúmeras correções e aprimoramentos indicados para este trabalho.

Sou grato também a UFC por minhas formações acadêmicas e ao IFCE por me prover moradia intelectual e profissional.

A meu pai e mãe, seu Dirceu e dona Ruth, que proveram, mesmo nos momentos mais difíceis, a educação e o assentamento sólido no qual pude construir minha vida, carreira e personalidade. A meu irmão Lucas, camaradinha de debates e descobertas desde a infância que é, além de bom amigo, um bom conselheiro. A todos da minha família que me suportaram até aqui eu só posso dizer até mais e obrigado pelos peixes!

A Andressa Bezerra entrego meu coração e minha mais profunda paixão e carinho. É ela a minha companheira de todos os momentos, conselheira, professora e árbitra, fonte de inspiração e modelo de esforço, dedicação e integridade. A você eu devo muito mais que agradecimentos, algo que ainda não sei pôr em palavras, mas sentir.

Agradeço aos professores Ítalo Alves, a Pâmela Toassi e Marcelo Amorim pela consultoria técnica e por todos os conselhos dados para aprimorar esta pesquisa, sem o esforço e o trabalho deles esta pesquisa não teria tomado os rumos que tomou. Sou também grato ao Antônio, companheiro de coleta e que foi fundamental nesta etapa da pesquisa. Devo agradecer também aos companheiros de POET Eurijúnior e Lia Carvalho por todas as correções e ajudas providas ao longo desses anos de trabalho e pesquisa. Gostaria de

agradecer também ao Kelvis Santiago, que é muito mais que um funcionário público, mas um grande amigo para todos da POET.

A todas pesquisadoras e pesquisadores que fazem parte do grupo LEAD e são a ponta da lança na pesquisa sobre acessibilidade não só no Ceará, mas em todo o Brasil. Meu muito obrigado e minha admiração pelo grupo sintetizam-se na figura da professora Vera Lúcia Santiago, a qual li as suas pesquisas ainda na graduação e somente neste momento da vida tive o prazer de conhecer mais de perto todo seu trabalho e de seus companheiros de grupo de estudo. Vocês fazem a diferença para muita gente nesse Brasil, mesmo sem eles sequer saberem.

Por fim, mas não menos importante, gostaria de agradecer a você que pegou esta dissertação para ler. Espero que ela sirva de material de aprendizagem a sua pesquisa e contribua para o avanço das ciências.

“Diante da vastidão do tempo e da imensidão do universo, é um imenso prazer para mim dividir um planeta e uma época com você.”
(Carl Sagan, 1980)

RESUMO

A Legendagem para Surdos e Ensurdecidos (LSE) é uma modalidade de Tradução Audiovisual Acessível (TAVa) que possibilita ao público surdo o acesso aos programas audiovisuais. Os pesquisadores do grupo Legendagem e Audiodescrição (LEAD), localizado na Universidade Estadual do Ceará (UECE), têm investigado diversos aspectos técnicos e linguísticos da LSE, como a importância da velocidade da legenda e sua segmentação na recepção de obras audiovisuais por surdos (ARAÚJO, VIEIRA e MONTEIRO, 2013; MONTEIRO e DANTAS, 2017). No entanto, não há ainda no Brasil um modelo bem estabelecido que padronize os diversos aspectos da tradução dos efeitos sonoros. A partir do estudo de Nascimento (2018) sobre a tradução de efeitos sonoros na LSE em filmes de DVD, esta pesquisa averiguou de maneira experimental a influência que a tradução de ruídos e músicas tem na emoção dos espectadores surdos durante a recepção de trechos de um filme de ação. Foram selecionados 16 participantes surdos para serem estimulados com 4 trechos do filme Dunkirk (2017). Para medir a influência desses efeitos sonoros sobre as emoções dos espectadores surdos, foi realizado um estudo com um programa de reconhecimento de expressões faciais que gera, a partir delas, dados quantitativos. Portanto, nossos objetivos foram: investigar a intensidade das emoções por meio do modelo de tradução dos efeitos sonoros proposto por Nascimento (2018) e mensurar a diferença de emoções mediante o uso e não uso da tradução de efeitos sonoros na LSE, com base nos dados gerados pelo programa de quantificação de emoções Visage com os espectadores surdos. Tivemos como variáveis dependentes as emoções básicas, que são: alegria, tristeza, raiva, medo, nojo, neutro e surpresa. As variáveis independentes foram a LSE com tradução dos efeitos sonoros e sem. Os dados foram analisados estatisticamente pela Análise de Variância e pelo Modelo Linear Misto. Os resultados obtidos em cada variável dependente corroboraram para um impacto positivo da tradução de efeitos sonoros nos espectadores surdos, eles responderam com maiores intensidades de emoção aos estímulos.

Palavras-chave: estudos da tradução, legendagem para surdos e ensurdecidos, tradução de efeitos sonoros, expressões faciais das emoções.

ABSTRACT

Subtitling for the Deaf and Hearing Impaired (SDH) is a modality of Accessible Audiovisual Translation (AAT) that enables deaf audiences to access audiovisual programmes. Researchers of the Subtitling and Audio description group (LEAD), located at the State University of Ceará (UECE), have investigated several technical and linguistic aspects of LSE, such as the importance of subtitle speed and its segmentation in the reception of audiovisual works by deaf people (ARAÚJO, VIEIRA and MONTEIRO, 2013; MONTEIRO and DANTAS, 2017). However, there is still no well-established model in Brazil that standardizes the various aspects of sound effects translation. Based on Nascimento's (2018) study on the translation of sound effects in SDH in DVD movies, this research experimentally investigated the influence that the translation of noises and music has on the emotion of deaf spectators during the reception of excerpts of an action movie. Sixteen deaf participants were selected to be stimulated with four excerpts from the movie *Dunkirk* (2017). To measure the influence of these sound effects on deaf viewers' emotions, a study was conducted with a facial expression recognition program that generates, from them, quantitative data. Therefore, our objectives were: to investigate the intensity of emotions through the sound effects translation model proposed by Nascimento (2018) and to measure the difference of emotions by using and not using sound effects translation in LSE, based on the data generated by the Visage emotion quantification program with deaf spectators. We had as dependent variables the basic emotions, which are: joy, sadness, anger, fear, disgust, neutral, and surprise. The independent variables were LSE with translation of the sound effects and without. The data were statistically analysed by Analysis of Variance and Mixed Linear Model. The results obtained in each dependent variable corroborated to a positive impact of sound effects translation on deaf spectators, they responded with higher intensities of emotion to the stimuli. The effect size, although always shown to be small, was found to be significant.

Keywords: translation studies, subtitling for the deaf and the hard of hearing, translation of sound effects, facial expressions of emotions.

LISTA DE ILUSTRAÇÕES

Figura 1 – Fluxograma da árvore sintática de Karamitroglou (1998)	35
Figura 2 – Interface do programa de edição <i>Adobe Premiere Pro</i> com cena do filme <i>Dunkirk</i> (2017)	60
Figura 3 – Interface do programa de legendagem <i>Subtitle Workshop 6.0b</i> com legendas do primeiro trecho do filme sendo confeccionadas	61
Figura 5 – Interface da IDE <i>Spyder</i> com parte do algoritmo escrito e em funcionamento	68
Figura 6 – <i>Frame</i> do vídeo com o processamento das emoções em quadro	70
Figura 7 – Gráfico gerado pelo programa com a prevalência de emoções pelo tempo.....	71
Figura 8 – Tabela gerada pelo <i>Visage</i> com os dados numéricos de um teste.....	72

LISTA DE QUADROS

Quadro 1 - Formato das legendas em tela	36
Quadro 2 – Comparativo entre os blocos de legenda com e sem tradução de efeitos sonoros	65
Quadro 3 – Questionário com Escala Diferencial das Emoções de Izard (1977) adaptado para a pesquisa.....	67
Quadro 4 – Quadrado Latino com a disposição dos participantes por trechos a serem expostos	73

LISTA DE TABELAS

Tabela 1 – Taxa de caracteres por segundo de acordo com a velocidade de 145ppm em caracteres por segundo.....	29
Tabela 2 – Taxa de caracteres por segundo de acordo com a velocidade de 160ppm em caracteres por segundo.....	29
Tabela 3 – Taxa de caracteres por segundo de acordo com a velocidade de 180ppm em caracteres por segundo.....	30
Tabela 4 – Dados descritivos da variável resposta <i>angry</i>	80
Tabela 5 - Dados descritivos da variável resposta <i>disgust</i>	83
Tabela 6 - Dados descritivos da variável resposta <i>fear</i>	84
Tabela 7 - Dados descritivos da variável resposta <i>happy</i>	87
Tabela 8 - Dados descritivos da variável resposta <i>neutral</i>	90
Tabela 9 - Dados descritivos da variável resposta <i>sad</i>	92
Tabela 10 - Dados descritivos da variável resposta <i>surprise</i>	95
Tabela 11 – ANOVA da variável resposta <i>angry</i>	99
Tabela 12 – Estimativas de efeito randômico da variável resposta <i>angry</i>	99
Tabela 13 – Estimativas de efeito fixo da variável resposta <i>angry</i>	99
Tabela 14 – ANOVA da variável resposta <i>disgust</i>	101
Tabela 15 – Estimativas de efeito randômico da variável resposta <i>disgust</i>	101
Tabela 16 – Estimativas de efeito fixo da variável resposta <i>disgust</i>	101
Tabela 17 – ANOVA da variável resposta <i>fear</i>	102
Tabela 18 – Estimativas de efeito randômico da variável resposta <i>fear</i>	103
Tabela 19 – Estimativas de efeito fixo da variável resposta <i>fear</i>	103
Tabela 20 – ANOVA da variável resposta <i>happy</i>	104
Tabela 21 – Estimativas de efeito randômico da variável resposta <i>happy</i>	104
Tabela 22 – Estimativas de efeito fixo da variável resposta <i>happy</i>	105
Tabela 23 – ANOVA da variável resposta <i>neutral</i>	106
Tabela 24 – Estimativas de efeito randômico da variável resposta <i>neutral</i>	106
Tabela 25 – Estimativas de efeito fixo da variável resposta <i>neutral</i>	106
Tabela 26 – ANOVA da variável resposta <i>sad</i>	108
Tabela 27 – Estimativas de efeito randômico da variável resposta <i>sad</i>	108
Tabela 28 – Estimativas de efeito fixo da variável resposta <i>sad</i>	108
Tabela 29 – ANOVA da variável resposta <i>surprise</i>	109

Tabela 30 – Estimativas de efeito randômico da variável resposta <i>surprise</i>	110
Tabela 31 – Estimativas de efeito fixo da variável resposta <i>surprise</i>	110

LISTA DE GRÁFICOS

Gráfico 1 – <i>Boxplot</i> da variável <i>angry</i> em ambas as variáveis preditoras.....	81
Gráfico 2 – Histograma da variável <i>angry</i> totalizada e sem <i>outliers</i>	82
Gráfico 3 – <i>Boxplot</i> da variável <i>disgust</i> em ambas as variáveis preditoras.....	83
Gráfico 4 – <i>Boxplot</i> da variável <i>fear</i> em ambas as variáveis preditoras.....	85
Gráfico 5 – <i>Boxplot</i> da variável <i>fear</i> em ambas as variáveis preditoras sem <i>outliers</i>	86
Gráfico 6 – Histograma da variável <i>fear</i> totalizada.....	86
Gráfico 7 – <i>Boxplot</i> da variável <i>happy</i> em ambas as variáveis preditoras	88
Gráfico 8 – <i>Boxplot</i> da variável <i>happy</i> em ambas as variáveis preditoras sem muitos <i>outliers</i>	89
Gráfico 9 – Histograma da variável <i>happy</i> totalizada	89
Gráfico 10 – <i>Boxplot</i> da variável <i>neutral</i> em ambas as variáveis preditoras	91
Gráfico 11 – Histograma da variável <i>neutral</i> totalizada.....	92
Gráfico 12 – <i>Boxplot</i> da variável <i>sad</i> em ambas as variáveis preditoras	93
Gráfico 13 – Histograma da variável <i>sad</i> totalizada.....	94
Gráfico 14 – <i>Boxplot</i> da variável <i>surprise</i> em ambas as variáveis preditoras.....	95
Gráfico 15 – <i>Boxplot</i> da variável <i>surprise</i> em ambas as variáveis preditoras sem muitos <i>outliers</i>	96
Gráfico 16 – Histograma da variável <i>surprise</i> totalizada.....	97
Gráfico 17 – Plot da variável <i>angry</i> de SSOM para CSOM.....	100
Gráfico 18 – Plot da variável <i>disgust</i> de SSOM para CSOM.....	102
Gráfico 19 – Plot da variável <i>fear</i> de SSOM para CSOM	103
Gráfico 20 – Plot da variável <i>happy</i> de SSOM para CSOM	105
Gráfico 21 – Plot da variável <i>neutral</i> de SSOM para CSOM	107
Gráfico 22 – Plot da variável <i>sad</i> de SSOM para CSOM	109
Gráfico 23 – Plot da variável <i>surprise</i> de SSOM para CSOM.....	111

SUMÁRIO

1	INTRODUÇÃO	16
2	REFERENCIAL TEÓRICO	21
2.1	Tradução Audiovisual (TAV) e Tradução Audiovisual Acessível (TAVa).....	21
2.2	Tradução Audiovisual Acessível no Brasil.....	24
2.3	Estudos em Legendagem para Surdos e Ensurdecidos (LSE)	26
2.4	O papel do som no cinema.....	37
2.5	Pesquisa de recepção com a tradução de efeitos sonoros na LSE	42
2.6	Pesquisas no contexto do grupo Legendagem e Audiodescrição (LEAD)	44
2.6.1	<i>Projeto MOLES.....</i>	46
2.6.2	<i>Pesquisas Experimentais em Legendagem</i>	47
2.7	Teoria das Emoções Básicas e Teoria da Universalidade das Expressões Faciais ..	53
3	METODOLOGIA.....	57
3.1	Tipo de pesquisa	57
3.2	Contexto da pesquisa	58
3.3	Estímulos.....	58
3.3.1	<i>Manipulação da LSE</i>	63
3.4	Instrumento qualitativo	65
3.5	Instrumento quantitativo	67
3.5.1	<i>O programa desenvolvido - Visage</i>	67
3.6	Seleção dos participantes.....	72
3.7	Desenho experimental.....	73
3.8	Variáveis e medidas	76
3.8.1	<i>Variáveis independentes.....</i>	76
3.8.2	<i>Variáveis dependentes.....</i>	76
3.9	Análise dos dados	77

4	RESULTADOS	79
4.1	Análise descritiva dos resultados quantitativos	79
<i>4.1.1</i>	<i>Angry</i>	<i>80</i>
<i>4.1.2</i>	<i>Disgust</i>	<i>82</i>
<i>4.1.3</i>	<i>Fear</i>	<i>84</i>
<i>4.1.4</i>	<i>Happy</i>	<i>87</i>
<i>4.1.5</i>	<i>Neutral</i>	<i>90</i>
<i>4.1.6</i>	<i>Sad</i>	<i>92</i>
<i>4.1.7</i>	<i>Surprise</i>	<i>94</i>
4.2	Resultados da Análise de Variância (ANOVA) e o Modelo Linear Misto (LMM)	97
<i>4.2.1</i>	<i>Angry</i>	<i>99</i>
<i>4.2.2</i>	<i>Disgust</i>	<i>100</i>
<i>4.2.3</i>	<i>Fear</i>	<i>102</i>
<i>4.2.4</i>	<i>Happy</i>	<i>104</i>
<i>4.2.5</i>	<i>Neutral</i>	<i>105</i>
<i>4.2.6</i>	<i>Sad</i>	<i>107</i>
<i>4.2.7</i>	<i>Surprise</i>	<i>109</i>
4.3	Discussão dos resultados	111
5	CONSIDERAÇÕES FINAIS	116
	REFERÊNCIAS	119
	APÊNDICE	124

1 INTRODUÇÃO

Para consumir entretenimento audiovisual, o surdo brasileiro tem a sua disposição duas alternativas de acessibilidade: a Tradução Audiovisual em Língua de Sinais (TALS) e a Legendagem para Surdos e Ensurdidos (LSE)¹. Esta última tornou-se obrigatória em todo o território nacional a partir da Norma Complementar nº1/2006 do Ministério das Comunicações, sendo dado o prazo de 11 anos às radiodifusoras para sua conformação à lei. A norma, entretanto, não trata de especificidades da tradução na LSE.

As emissoras e produtoras brasileiras de conteúdo audiovisual se adequaram às normas, mas a preferência foi dada ao sistema norte-americano *Closed Caption* (CC), de legendagem fechada, em detrimento da LSE (ARAÚJO, 2004). Esse tipo de legendagem apresenta uma série de problemas, tanto para ouvintes como para surdos, dentre os quais podemos citar: serem exibidas em até quatro linhas, apresentadas palavra por palavra em tela e ter sentenças de difícil compreensão (ARAÚJO, 2004; 2008; 2012; 2016). O CC nada mais é que uma transcrição da fala, que não passa por nenhum processo de edição, diferente do que é feito na Legendagem para Ouvintes (LO) e Legendagem para Surdos e Ensurdidos (LSE) (ARAÚJO, 2012). O acesso do público surdo às produções audiovisuais no Brasil é, então, prejudicado pelo modelo adotado pelas emissoras brasileiras de TV (ARAÚJO, 2016; ASSIS, 2016; MONTEIRO; DANTAS, 2017).

A fim de aprimorar o acesso e a compreensão de obras fílmicas pelo público surdo, a LSE é encarada, diferentemente do CC, por seus parâmetros técnicos. Ela segue as mesmas características de uma legenda direcionada a ouvintes, com algumas pequenas diferenciações, como a adição do nome do falante nas legendas e a tradução de efeitos sonoros (ARAÚJO, 2012).

A linha de pesquisa Legendagem e Audiodescrição (LEAD), da Universidade Estadual do Ceará (UECE), presente no grupo Tradução e Semiótica, investiga a Tradução Audiovisual Acessível (TAVa) há mais de quinze anos. Utilizaremos a nomenclatura grupo LEAD para nos referirmos ao grupo Tradução e Semiótica ao longo do texto. Seus membros já investigaram diversos parâmetros da LSE, como velocidade da legenda, segmentação linguística e tradução de músicas e ruídos. Os trabalhos de Araújo (2004; 2007; 2008; 2012;

¹ Tradução estabelecida por Selvatici (2010) para “Subtitling for the Deaf and the Hard of Hearing” (SDH).

2016) estabeleceram a distinção entre LSE e o CC, apontando seus problemas e sugerindo como alternativa a adoção dos mesmos parâmetros técnicos da LO para a LSE. Também foi levantada a hipótese de a velocidade das legendas ser o fator primordial para a boa recepção do texto pelos surdos, no entanto, tal hipótese foi refutada com o seguimento das pesquisas experimentais com rastreamento ocular (VIEIRA, 2016; VIEIRA; ARAÚJO, 2017).

Os trabalhos de Araújo e Assis (2014), Vieira (2016), Assis (2016), Monteiro e Dantas (2017) e Souza (2020) são alguns desse grupo de pesquisa que investigaram a segmentação linguística e a velocidade da LSE como influência direta na recepção de legendas pelos surdos. Os resultados apresentaram um paralelo entre uma boa segmentação e a compreensão das obras audiovisuais.

Os trabalhos de Nascimento (2013; 2018), por sua vez, investigaram como são feitas as traduções de efeitos sonoros em filmes de DVD e o último propõe um modelo de tradução dos efeitos sonoros na LSE, mas que ainda não foi testado com espectadores surdos. É nesta lacuna que se insere a presente pesquisa, em que testaremos o modelo proposto por Nascimento (2018) com surdos brasileiros, a fim de investigar o papel da tradução de músicas e ruídos na recepção fílmica pelo público surdo.

O estudo da tradução de efeitos sonoros na LSE faz-se necessário porque as produções audiovisuais têm, no som, um canal de transmissão de informações. Os sons no cinema se faziam presentes mesmo quando as películas não tinham a capacidade técnica de apresentar imagem e som em sincronia, o que era conhecido como cinema mudo. O avanço tecnológico permitiu que o som nas mídias visuais fosse consolidado como um canal de informação tão importante quanto o próprio vídeo. Para Bordwell (2008), o som não pode ser mais dissociado da imagem em um filme, sendo sua função variável, desde um simples acompanhamento da cena à provocação ou realce de emoções pelo espectador.

Os sons, que vão além das músicas ou dos diálogos, estão presentes também em ruídos ao longo de toda a trama e alguns podem ser fundamentais para o realce de emoções desejadas pelo diretor da obra. Um carro que está fora de cena e aproxima-se freando perto de um personagem, por exemplo, só é notado pelo espectador se ele puder ouvir aquele ruído. Nesse exemplo, além de apresentar uma função de localização espacial que o áudio pode exercer, se de fato houver um acidente, o impacto poderá ser sentido com mais ênfase, justamente pela antecipação do áudio. As emoções que são suscitadas pelo filme, e realçadas pelo som, podem ser mensuradas e categorizadas junto ao enredo fílmico.

Neste trabalho, utilizaremos a definição da Teoria das Emoções Básicas, de Ekman (1970; 2003) e Ekman e Friesen (1978), para identificar e catalogar as emoções expressas pelos participantes. Segundo os autores, todos os seres humanos, independentemente de sua cultura ou meio social, expressam através da face seis emoções básicas. Dentre elas, há o estado neutro, no qual a face humana parte para expressar a emoção sentida no momento. Todas essas seis emoções são expressas por uma determinada ativação do conjunto dos músculos faciais no instante que o cérebro experiencia o sentimento ligado à emoção. Essa expressão facial da emoção é universal e faz-se presente tanto em humanos como em demais primatas. Os seres humanos, portanto, deverão ser capazes de expressar suas emoções por meio de suas expressões faciais (EKMAN e FRIESEN, 1978; DARWIN, 2009).

As expressões faciais são, então, um meio de se avaliar o impacto emocional de algum artefato em um humano. Essa avaliação pode se dar de maneira manual, com o pesquisador analisando a formação de cada expressão pelos músculos da face, ou por meio de programas de computador que já reconhecem e avaliam as expressões faciais de maneira automática. Estes *softwares* são capazes de executar as tarefas de captar, categorizar, separar, analisar e gerar dados numéricos a partir de um vídeo da face de uma pessoa. Nesta pesquisa, utilizamos um programa, intitulado *Visage*, com essas mesmas capacidades, a fim de avaliar as reações faciais de cada um dos participantes do estudo.

As pesquisas feitas sobre a LSE no Brasil descobriram novos paradigmas quanto à boa recepção das legendas pelo público surdo. No entanto, ainda se pode identificar uma lacuna no estudo do impacto da tradução de efeitos sonoros na LSE. Essa necessidade abre oportunidades para estudos mais aprofundados sobre a influência que a tradução de sons tem na recepção de espectadores surdos. Este trabalho, portanto, leva em consideração os estudos descritivos de Nascimento (2013; 2018) sobre a tradução dos efeitos sonoros em filmes de DVD. Assim, o objetivo geral desta pesquisa é investigar se a tradução dos efeitos sonoros na LSE de um filme de ação pode influenciar as emoções dos espectadores surdos durante sua recepção.

Para alcançar o referido objetivo, alguns objetivos específicos fazem-se necessários, sendo eles: 1) Investigar a intensidade das emoções na recepção da LSE por meio do modelo de tradução dos efeitos sonoros proposto por Nascimento (2018); 2) Mensurar a diferença sobre as emoções mediante o uso e não-uso da tradução de efeitos sonoros na LSE.

As hipóteses formuladas para esta pesquisa são: 1) a tradução dos efeitos sonoros na LSE intensifica as emoções durante a recepção de espectadores surdos em um filme de

ação; 2) a não tradução dos efeitos sonoros na LSE reduz a intensidade das emoções durante a recepção de espectadores surdos.

A presente pesquisa se valeu da participação de 16 surdos cearenses para a coleta de dados. Eles foram expostos a quatro trechos legendados do filme de ação *Dunkirk* (2017). As exibições foram divididas em dois grupos, pautados em como as legendas foram manipuladas: um com tradução de efeitos sonoros e outro sem esta tradução. Todos os participantes responderam a um questionário demográfico antes de iniciar a coleta de dados e, à medida que forem assistindo aos trechos, responderam a um questionário de autoavaliação de suas emoções. Enquanto o participante estava assistindo, sua face era gravada, com o devido consentimento, e os vídeos gerados a partir das gravações foram analisados no programa *Visage*, para estudo das expressões faciais. Os dados do *software* serão analisados, procurando confirmar ou refutar as nossas hipóteses.

Esta pesquisa é realizada em parceria entre este pesquisador e o grupo LEAD, fazendo parte, mais precisamente, do projeto de Pesquisas Experimentais em Legendagem (ExLEG). Buscamos investigar a importância da tradução de músicas e ruídos para a boa recepção do público surdo da obra fílmica. Esperamos fomentar e estabelecer bases para a execução de mais pesquisas experimentais na área da LSE com base em programas de análise facial, a fim de testar as hipóteses que o grupo LEAD vem propondo ao longo dos anos. Esperamos, também, com este trabalho, dar início a uma exploração da problemática de maneira mais aprofundada e que permita que os dados obtidos possam abrir novos caminhos de pesquisa na área da tradução de efeitos sonoros. Além dos motivos já citados, acreditamos que este trabalho poderá contribuir com a promoção da acessibilidade no consumo de produtos audiovisuais pelo público surdo e ensurdecido brasileiro.

O texto desta dissertação está dividido em cinco capítulos, a contar com esta introdução. O capítulo introdutório apresenta um preâmbulo do trabalho como um todo, além dos objetivos e das hipóteses de nosso estudo. O capítulo dois faz um apanhado geral da literatura relacionada à LSE, ao cinema e ao som e, por fim, a teoria das emoções básicas. Nesse mesmo capítulo, é feita uma descrição dos principais parâmetros técnicos da legendagem e como elas se relacionam com a LSE. Também são apresentados os diversos trabalhos do grupo LEAD e a forma como pavimentaram o caminho teórico da LSE no Brasil. Em seguida, apresentamos o estudo descritivo de tradução de efeitos sonoros de Nascimento (2018). No terceiro capítulo, apresentamos o desenho da pesquisa e todo o desenvolvimento do estudo, além dos questionários, a criação e funcionamento do programa *Visage* e o método

utilizado para analisar os dados obtidos na coleta. O quarto capítulo traz os resultados obtidos em nossa pesquisa. Apresentamos os dados numéricos e estatísticos gerados por nosso programa. O último capítulo de nosso trabalho versa sobre nossas considerações sobre a pesquisa, nossos resultados e perspectivas de trabalhos futuros.

2 REFERENCIAL TEÓRICO

Neste capítulo, apresentamos o percurso das pesquisas que fundamentam a base teórica deste trabalho. Ele conta com sete seções que explanarão a base teórico-metodológica advinda da área de Tradução Audiovisual Acessível (TAVa) e seus aspectos no Brasil, do papel do som no cinema e da Teoria Das Emoções Básicas Universais.

A primeira seção introduz alguns aspectos da tradução audiovisual e suas ramificações na acessibilidade audiovisual. O segundo ponto trata do percurso traçado no Brasil em relação à acessibilidade nas suas produções audiovisuais, seus parâmetros e legislações que promovem a inclusão neste meio, principalmente no que tange ao público surdo brasileiro. Na terceira seção, apresentamos a trajetória que as pesquisas em TAVa tiveram nas últimas décadas e os avanços que elas proporcionam para a manipulação e compreensão da LSE. A quarta seção trata do papel do som no cinema e nas diversas produções audiovisuais, destacando sua importância nas tramas e seus meios mais comuns de manifestação nas obras fílmicas. Também abordamos os trabalhos de Nascimento (2013, 2018), sobre a tradução de ruídos e músicas do cinema. Na quinta seção, apresentamos a pesquisa de Aleksandrowicz (2019), que testou modos diferentes de tradução de ruídos e músicas na LSE polonesa e apresenta algumas semelhanças com esta pesquisa. O sexto ponto trata das pesquisas desenvolvidas no contexto do grupo LEAD e as descobertas em relação à LSE, seus parâmetros técnicos e impactos na área. A sétima e última parte deste capítulo trata dos conceitos e implicações teóricas relativas à Teoria das Emoções Básicas Universais, utilizada como base para validar os dados gerados pela ferramenta quantitativa utilizada nesta pesquisa.

2.1 Tradução Audiovisual (TAV) e Tradução Audiovisual Acessível (TAVa)

O audiovisual, como negócio, tem em sua trajetória a marca da expansão para diversos mercados ao redor do mundo, sendo um expoente da globalização (BORDWELL, 2013). Este crescimento global leva consigo a língua e a cultura de seus países de origem para serem compreendidas nos lugares de chegada. O cinema e, posteriormente, a televisão foram os principais difusores desses aspectos no começo da era da comunicação.

Díaz Cintas (2003) cita o avanço tecnológico, principalmente o crescimento do digital, como principal catalisador de uma oferta mais ampla e acessível de produtos

audiovisuais. A tecnologia VHS, nos anos 1980 e 1990, proporcionou aos produtores audiovisuais os primeiros passos na ampliação de seu mercado consumidor, sendo posteriormente seguido pelo DVD e canais a cabo (DÍAZ CINTAS, 2003) e difundido por meio do *streaming* de vídeo hoje em dia.

Mediante a crescente profusão de novos produtos e mídias em diversas línguas, surgiu a necessidade de se compreender e desenvolver modelos e parâmetros para a tradução dessas obras para as línguas de chegada. Nesse ponto, as pesquisas no âmbito dos Estudos da Tradução, especificamente da Tradução Audiovisual (TAV), passam a florescer e dilatar-se com a globalização das mídias audiovisuais (SOUZA, 2020).

O crescimento, portanto, de variados e persistentes modos de se consumir uma mídia estrangeira alçou a TAV a uma área de estudo cada vez mais importante e examinada por diversas áreas de pesquisadores (SOUZA, 2020). Díaz Cintas (2003) ainda vê muito espaço para o crescimento da área de estudo, tanto pela inserção de novos modos de se consumir o audiovisual como pelo avanço dos estudos que aliam o audiovisual à acessibilidade de sujeitos surdos.

A TAV passou a ser incluída nos Estudos da Tradução a partir da classificação de Gambier (2003). Nela, a tradução é dividida em: 1) tradução interlinguística, ou seja, entre línguas diferentes, enquadrando, aqui, as legendas que traduzem um filme em inglês para o português; 2) tradução intralinguística, na qual a língua oral é traduzida para seu texto escrito, enquadrando aqui, também, a LSE, por possibilitar a tradução de produções audiovisuais aos surdos; 3) tradução intersemiótica, que seria a tradução entre meios semióticos diferentes, enquadrando a tradução dos efeitos sonoros na LSE, foco do nosso estudo, e a audiodescrição para pessoas com deficiência visual.

Há algumas outras especificidades na tradução como um todo. Perego (2008) afirma que traduções que partem da língua falada para a escrita perdem características e sutilezas próprias do código falado. É uma característica inerente ao processo de tradução. Contudo, mesmo com a perda de informações no texto escrito, ainda é possível minimizar o impacto desses marcadores linguísticos. Na TAV, mesmo os meios que mantêm o uso do canal de fala como transmissor de informações, como no *voice-over*, na dublagem, na legendagem e na audiodescrição, há alterações no texto de chegada para que se adeque a uma cultura ou a uma necessidade específica.

Diante disso, há uma variedade significativa de meios tradutórios dentro da TAV, algumas apresentando maior ou menor índice de perda de informação, mas, ainda assim,

tendo nichos de utilização bem demarcados. Destacam-se alguns dos tipos de tradução mais comumente difundidos, podendo ser utilizadas tanto numa tradução interlinguística como intralinguística (ARAÚJO, 2008; NASCIMENTO, 2013):

- *Legendagem*: Texto inserido concomitantemente com a fala do personagem que a profere em cena. Apresenta condensação de informações, se dá no canal visual e, geralmente, é posicionado no terço inferior da imagem;
- *Dublagem*: Substituição das falas dos atores e atrizes na língua de origem por textos interpretados por outros atores, conhecidos por dubladores, na língua de chegada. Utiliza-se do canal auditivo para a tradução. É comumente observado nas traduções interlinguísticas;
- *Voice-over*: Veiculação do canal auditivo de origem em volume mais baixo simultaneamente à tradução na língua de chegada em volume cheio. Geralmente, o texto falado na língua de chegada é inserido com um pequeno atraso em relação ao de origem. É mais utilizado na tradução interlinguística;
- *Audiodescrição*: Narração descritiva por meio do canal auditivo do que se vê em tela. Costuma descrever o ambiente, os atores e atrizes que estão em cena e as ações transcorridas.

A legendagem e a dublagem são os meios mais comuns e populares, por questões de costume e hábito da população consumidora de conteúdos audiovisuais, e, por isso, a área de estudo da TAV focada nesses modos apresentam grande potencial para continuar a crescer (GAMBIER, 2012).

Inserida à TAV, situa-se a Tradução Audiovisual Acessível (TAVa), termo inserido na literatura por Jiménez Hurtado e Seibel (2007) e Aderaldo (2014). Essa subárea foca em promover a acessibilidade em produções audiovisuais aos espectadores com deficiência sensorial. Tem se expandido tanto em âmbito nacional como internacional (MONTEIRO e DANTAS, 2017). Dentro da TAVa, existem diversas áreas de estudos direcionadas a públicos com necessidades diferentes: Tradução Audiovisual em Língua de Sinais, LSE e audiodescrição.

Na legendagem, a leitura é uma atividade necessária para a recepção dos produtos audiovisuais. Considerando tal atividade, Perego (2008) afirma que ela requer alto grau de atenção, além de apresentar uma carga cognitiva elevada para a decodificação do leitor. O ato de ler, portanto, depende da habilidade leitora e da velocidade de leitura de cada um, sendo ambas diferentes entre surdos e ouvintes (PEREGO, 2008). D'Ydewalle et al. (1991) explicam

que, em uma obra audiovisual legendada, a informação passada em tela está disponível em dois meios: o auditivo, com a fala; e o visual, com a legenda. Essa modalidade multissemiótica caracteriza o audiovisual como uma atividade ainda mais complexa que a simples leitura, visto que a entrada das informações sonoras aumenta a quantidade de dados fornecidos e, por consequência, o consumo da capacidade cognitiva do espectador (D'YDEWALLE et al., 1991).

Os surdos, por não terem acesso às informações sonoras, apresentam desvantagem na compreensão da obra fílmica em relação ao ouvinte (PEREGO, 2008). Mesmo que sua atenção se divida entre imagem e legenda, a falta do canal sonoro não permite que ele passe pela experiência de maneira igualitária ao ouvinte. Dessa forma, as legendas para surdos precisam conter informações que ajudem esse espectador a compreender a cena, traduzindo as informações auditivas.

A LSE, foco deste trabalho, pode ser definida como um recurso de acessibilidade que faz uso de legendas em produtos audiovisuais com o intuito de compensar a falta do canal auditivo por seu espectador surdo (ARAÚJO, 2008). Segundo Araújo (2008), as diferenças da LSE para a LO podem ser categorizadas em três principais pontos: tradução de efeitos sonoros, inserção de dados escritos que descrevem ações dependentes do canal sonoro e uso de colchetes para identificar os falantes. A tradução de efeitos sonoros nas legendas e as concepções de tradução são alguns dos principais diferenciadores dos dois tipos de legendas.

2.2 Tradução Audiovisual Acessível no Brasil

O Brasil, em busca de promover a acessibilidade para surdos e pessoas com deficiências visuais, promoveu a implementação da norma complementar 01/2006. Ela foi promulgada pelo Ministério das Comunicações no dia 27 de junho de 2006 e seu conteúdo versa sobre a obrigatoriedade dos veículos de comunicação em promover a acessibilidade em sua grade de programação. Os meios seriam LSE, TALS para o público surdo e Audiodescrição para pessoas com deficiência visual. As radiodifusoras tiveram o prazo de 11 anos para se adequar a nova norma.

A LSE, no entanto, não foi encarada como um processo de legendagem, visto que a programação audiovisual brasileira ficou restrita ao sistema de CC. Ao invés disso, ela foi tornada uma transcrição da fala, em vez de legendas destinadas ao público surdo. Tomando como foco a LO e o CC, por sua vez, estes são pensados e produzidos de maneiras distintas

(MONTEIRO e DANTAS, 2017). A LO segue os padrões estabelecidos pela legendagem aberta, como máximo de duas linhas por vez em tela e edição do texto antes de sua veiculação. O CC, no entanto, pode ser exibido em até quatro linhas e, no caso das legendas fechadas rotativas, ser apresentadas palavra por palavra. Além disso, por vezes, não apresenta sentido nas sentenças (ARAÚJO, 2004; 2008; 2012; 2016). Com esse contexto em mente, ressalta-se, ainda, que a falta de um processo de edição transforma essas legendas em meras transcrições da fala (ARAÚJO, 2012). Todas estas dificuldades impostas pelas legendas fechadas na televisão brasileira prejudicam o entendimento e o acesso dos surdos ao conteúdo televisivo, além de não poderem ser consideradas uma legendagem para surdos (ARAÚJO, 2016; ASSIS, 2016; MONTEIRO e DANTAS, 2017).

Outra diferença que aumenta a distância entre a produção do CC e LO no Brasil é a técnica para confeccionar as legendas. A legenda, quando é destinada para o ouvinte, é elaborada em um programa próprio para esta tarefa – nesta pesquisa, por exemplo, utilizaremos um *software* específico para a tarefa de manipular legendas. Essas legendas apresentam, também, uma velocidade constante e previamente definida, que pode variar de 145 a 180 palavras por minuto, incluindo os espaços em branco (ARAÚJO, 2012). Para encaixar as falas do canal auditivo no texto, que é naturalmente mais restrito, são feitas, também, suas condensações.

As legendas em CC, geralmente, são escritas com letras brancas sobre uma tarja de fundo preto e só podem ser lidas mediante o uso de um decodificador de legendas na TV e ativadas, geralmente, pela tecla *Closed Caption* no controle remoto. Essas legendas situam-se na linha 21 do intervalo da TV e podem, ainda, variar em dois tipos: *Roll-up* e *Pop-on* (ARAÚJO, 2016).

Segundo Araújo (2016), as legendas fechadas do tipo *Roll-up*, ou rotativas, são as mais comuns em programas ao vivo. Elas apresentam a característica de subir a linha inferior a cada nova sentença, podendo acumular até quatro linhas simultâneas em tela, e costumam aparecer palavra por palavra.

Ainda segundo Araújo (2016), as legendas do tipo *Pop-on* se diferenciam das rotativas por apresentarem sentenças inteiras na tela, ficando na tela o tempo necessário para apresentar sincronismo com o áudio e sendo substituídas logo em seguida por outra linha de legenda. É mais comum em programas pré-gravados e se aproximam mais das legendas abertas. São elaboradas por um profissional chamado estenotipista fazendo uso de um aparelho chamado estenótipo.

O estenótipo é uma espécie de teclado ligado a um computador, sendo o mesmo bastante comum para transcrever sessões de tribunais e assembleias do congresso (ARAÚJO, 2008; 2016). Normalmente, o estenotipista tem que digitar em média 150 palavras por minuto. Em programas ao vivo e *online*, como *streaming*, o CC é transmitido via um *software* que reconhece a voz e a transforma em texto escrito. Essa tecnologia é utilizada no CC do tipo *Roll-up*. Entretanto, os estudos de recepção sobre o CC já mostraram que esse tipo pode ser considerado ineficiente na promoção da acessibilidade à comunidade surda (ARAÚJO, 2004; 2008).

Os pesquisadores do grupo Legendagem e Audiodescrição (LEAD) (ARAÚJO, 2004; 2007; 2008; 2012; 2016; ARAÚJO e NASCIMENTO, 2011; ARAÚJO, VIEIRA e MONTEIRO, 2013; NASCIMENTO, 2013, 2018; VIEIRA, 2016; VIEIRA, ASSIS e ARAÚJO, 2020), no entanto, propõem uma abordagem diferente do que foi estabelecido pela Norma Complementar nº1/2006 no Brasil e implementada pelas emissoras de radiodifusão. O primeiro aspecto é elaborar a LSE de acordo com os mesmos princípios técnicos estabelecidos na LO: quantidade máxima de duas linhas, texto centralizado e de cor amarela ou branca, adicionando, ainda, a tradução dos efeitos sonoros e a identificação dos falantes.

Em relação aos aspectos técnicos de tempo de exposição, o grupo LEAD sugere o padrão mínimo de 1 segundo e máximo de 4 segundos (D'YDEWALLE, VAN RENSBERGEN e POLLE, 1987). Essa duração foi estabelecida com base na necessidade de tempo hábil para que o espectador possa ler a legenda e ainda tenha tempo para olhar para a imagem na tela. O tempo excedente a seis segundos provocaria releitura, ou seja, o observador terminaria a leitura da legenda e, por ela ainda estar presente em tela, a releria, prejudicando a compreensão das imagens (D'YDEWALLE, VAN RENSBERGEN e POLLE, 1987). Por outro lado, o tempo inferior a um segundo não seria o suficiente para que o espectador pudesse ter tempo de ler o texto exposto.

No Brasil, contudo, o padrão usual é de, no máximo, quatro segundos. Apesar da regra dos quatro segundos também ter sido validada por d'Ydewalle et al. (VIEIRA, ASSIS e ARAÚJO, 2020), a manipulação das legendas na presente pesquisa seguirá a recomendação brasileira e será utilizada a regra dos quatro segundos quando necessário e de, no máximo, duas linhas de legendas com até 37 caracteres cada (D'YDEWALLE, VAN RENSBERGEN e POLLE, 1987; DÍAZ CINTAS e REMAEL, 2007).

2.3 Estudos em Legendagem para Surdos e Ensurdidos (LSE)

A evolução das pesquisas na área da legendagem e, principalmente, da LSE ocorrem em todo o mundo. No âmbito mundial, destacam-se, para este trabalho, as pesquisas de d'Ydewalle et al (1987, 1991), Karamitroglou (1997), Díaz Cintas e Remael (2003) e Aleksandrowicz (2019). No Brasil, as pesquisas desenvolvidas dentro do grupo LEAD avançam por meio de pesquisas experimentais, descritivas e exploratórias que buscam estabelecer uma literatura que proponha um modelo eficiente para a manipulação de legendas para surdos.

D'Ydewalle et al. (1987) organizaram duas pesquisas, com a finalidade de descobrir o comportamento humano em relação à leitura de legendas e, principalmente, testar a regra dos seis segundos, regra comumente usada pelos legendistas na Europa. Em um estudo anterior, os pesquisadores manipularam 104 legendas de filmes em alemão, seguindo três regras distintas de duração das legendas em tela (4, 6 e 8 segundos). Entretanto, essa mistura das legendas com tempos diferentes mostradas aos participantes na mesma apresentação, não foi suficiente para estabelecer um comportamento padrão. A pesquisa revelou, também, que os espectadores gastavam mais tempo nas legendas de duas linhas que nas de uma linha, apesar de proporcionalmente, as de duas serem mais confortáveis. Uma das conclusões desta pesquisa é que os observadores conseguiam distinguir a duração de tempo de uma legenda, visto que já apresentavam o costume de ler no tempo de 6 segundos, culminando num desconforto com os demais intervalos (D'YDEWALLE et al, 1987).

D'Ydewalle et al. (1987) repetiram os mesmos trechos de filmes da pesquisa anterior, mas o diferencial foi a requisição de uma resposta imediata por parte do participante em relação à velocidade da legenda. Eles deveriam afirmar imediatamente caso sentissem que as legendas estavam muito rápidas, adequadas ou muito lentas. Como resultado, os espectadores queixaram-se mais da duração das legendas de uma linha em relação às de duas (d'YDEWALLE et al, 1987). Como conclusão, os autores reivindicaram que as legendas de uma linha são experienciadas de maneira mais rápida que a de duas, independente da velocidade em tela. Este resultado foi de encontro, à época, com a hipótese de que legendas de duas linhas tinham uma carga cognitiva maior para o espectador decodificar (d'YDEWALLE et al, 1987).

No segundo experimento, ainda da mesma pesquisa, os autores separaram três fitas de um filme em alemão, com legendas nas durações de 2, 4 e 6 segundos – os 8 segundos já não estavam mais presentes aqui. Esse experimento fez uso da ausência do canal de áudio

em um grupo de participantes, para averiguar o impacto do áudio na leitura das legendas. Pelo fato de não se ter o som para se guiar no entendimento do trecho, os espectadores eram obrigados a recorrerem às legendas para compreender a trama em tela. Ainda houve um outro grupo de colaboradores com proficiência comprovada em alemão para acompanhar os trechos.

Nesse estudo, os autores perceberam que os participantes com proficiência leem as legendas da mesma forma que os demais – desde que a regra dos 6 segundos fosse seguida – e que a presença ou ausência do áudio não interferia na duração do olhar do espectador no texto da legenda. A conclusão deste estudo foi que, mesmo com o auxílio do canal auditivo, o espectador ainda assim dará prioridade a leitura das legendas. D’Ydewalle et al. (1987) afirmaram que legendas de uma linha poderiam ficar mais tempo em tela que o seu equivalente de duas linhas. Caso respeitada a regra dos seis segundos, a leitura das legendas com duas linhas seria o melhor cenário possível para o espectador.

Mais adiante, Díaz Cintas e Remael (2007) trouxeram uma grande contribuição técnica e teórica para legendistas. Os estudos de d’Ydewalle et al. (1987) foram de grande importância para os pesquisadores estabelecerem a quantidade de caracteres por quadro de uma legenda, além de diversos parâmetros técnicos, como posicionamento, segmentação e marcação do tempo.

Para propor a quantidade de caracteres por segundo, (CpS), Díaz Cintas e Remael (2007) usaram a regra dos 6 segundos proposta por d’Ydewalle et al. (1987). Os pesquisadores estabeleceram a taxa de caracteres para três diferentes velocidades de legendas por meio de um programa de legendagem. As velocidades eram de 145, 160 e 180 palavras por minuto (ppm). A quantidade máxima de caracteres que cada bloco de legenda poderia ter era de 74, 78 e 78 CpS, respectivamente. Os autores ainda estabeleceram uma correlação entre a quantidade de *frames* (quadros) e o valor máximo de caracteres que deveriam aparecer em tela. Vale ressaltar que um filme, geralmente, possui uma taxa de 24 quadros por segundo (FPS), ou seja, a cada segundo, vemos 24 fotogramas diferentes, que nos dão a sensação de movimento na tela. Esta taxa pode variar, mais comumente, para 30 ou 60 FPS, dependendo da velocidade de gravação de cada filme (BORDWELL e THOMPSON, 2014).

As tabelas mostram a relação entre as três diferentes velocidades e a quantidade respectiva de CpS.

Tabela 1 – Taxa de caracteres por segundo de acordo com a velocidade de 145ppm em caracteres por segundo

145 palavras por minuto		Segundos(s):	Caracteres	Segundos(s):	Caracteres
		<i>Frames(f)</i>		<i>Frames(f)</i>	
		1s:00f	16	2s:00f	29
		1s:04f	17	2s:04f	32
		1s:08f	18	2s:08f	34
		1s:12f	20	2s:12f	36
		1s:16f	23	2s:16f	38
1s:20f	25	2s:20f	40		
Segundos(s):	Caracteres	Segundos(s):	Caracteres	Segundos(s):	Caracteres
<i>Frames(f)</i>		<i>Frames(f)</i>		<i>Frames(f)</i>	
3s:00f	44	4s:00f	58	5s:00f	71
3s:04f	46	4s:04f	60	5s:04f	71
3s:08f	48	4s:08f	62	5s:08f	73
3s:12f	50	4s:12f	64	5s:12f	73
3s:16f	52	4s:16f	65	5s:16f	74
3s:20f	54	4s:20f	67	5s:20f	74

Fonte: Díaz Cintas e Remael (2007, p. 97), adaptado pelo autor

Tabela 2 – Taxa de caracteres por segundo de acordo com a velocidade de 160ppm em caracteres por segundo

160 palavras por minuto		Segundos(s):	Caracteres	Segundos(s):	Caracteres
		<i>Frames(f)</i>		<i>Frames(f)</i>	
		1s:00f	17	2s:00f	31
		1s:04f	18	2s:04f	34
		1s:08f	20	2s:08f	37
		1s:12f	23	2s:12f	40
		1s:16f	26	2s:16f	42
1s:20f	28	2s:20f	44		
Segundos(s):	Caracteres	Segundos(s):	Caracteres	Segundos(s):	Caracteres
<i>Frames(f)</i>		<i>Frames(f)</i>		<i>Frames(f)</i>	
3s:00f	48	4s:00f	63	5s:00f	75
3s:04f	50	4s:04f	65	5s:04f	75
3s:08f	53	4s:08f	67	5s:08f	76
3s:12f	56	4s:12f	69	5s:12f	76
3s:16f	58	4s:16f	71	5s:16f	77
3s:20f	60	4s:20f	73	5s:20f	77
				6s:00f	78

Fonte: Díaz Cintas e Remael (2007, p. 97), adaptado pelo autor

Tabela 3 – Taxa de caracteres por segundo de acordo com a velocidade de 180ppm em caracteres por segundo

180 palavras por minuto		Segundos(s): <i>Frames(f)</i>	Caracteres	Segundos(s): <i>Frames(f)</i>	Caracteres
		1s:00f	17	2s:00f	35
		1s:04f	20	2s:04f	37
		1s:08f	23	2s:08f	39
		1s:12f	26	2s:12f	43
		1s:16f	28	2s:16f	45
		1s:20f	30	2s:20f	49
Segundos(s): <i>Frames(f)</i>	Caracteres	Segundos(s): <i>Frames(f)</i>	Caracteres	Segundos(s): <i>Frames(f)</i>	Caracteres
3s:00f	53	4s:00f	70	5s:00f	78
3s:04f	55	4s:04f	73	5s:04f	78
3s:08f	57	4s:08f	76	5s:08f	78
3s:12f	62	4s:12f	76	5s:12f	78
3s:16f	65	4s:16f	77	5s:16f	78
3s:20f	68	4s:20f	77	5s:20f	78
				6s:00f	78

Fonte: Díaz Cintas e Remael (2007, p. 97), adaptado pelo autor

A velocidade das legendas poderá variar de acordo com a rapidez das falas em cada programa, bem como a capacidade do espectador de ler as legendas e ver as imagens (DÍAZ CINTAS e REMAEL, 2007). Acreditava-se que, pelo fato de os surdos serem usuários do português como segunda língua, teriam desvantagem ao ler em velocidades mais rápidas, justamente por não estarem lendo na língua materna. Essa hipótese guiou os primeiros estudos do grupo LEAD no sentido de atribuir à velocidade o papel essencial na compreensão das legendas pelo público surdo (ARAÚJO, 2004, 2007, 2008). Com estudos posteriores, essa hipótese foi refutada, e outros elementos, como a segmentação, passaram a ter caráter mais relevante nas pesquisas do grupo (ARAÚJO, 2012; NASCIMENTO, 2013; ARAÚJO, VIEIRA e MONTEIRO, 2013; ASSIS, 2016; VIEIRA, 2016; MONTEIRO e DANTAS, 2017).

Díaz Cintas e Remael (2007) afirmam que o texto escrito em legendas, geralmente, sofre uma redução natural, quando comparado a sua partida oral, já que não é necessário que o texto contenha absolutamente todas as informações orais para que seja inteligível. Essa redução pode se dar de duas maneiras: parcial ou total. A parcial consiste na condensação de trechos orais ou abreviação de sentenças do texto de partida. Já a redução

total elimina elementos lexicais inteiros do texto da legenda; isso pode significar remoção ou omissão de certas palavras ou frases (DÍAZ CINTAS e REMAEL, 2007). Geralmente o legendista faz uso de ambos os métodos de redução para alcançar um resultado conciso, mas ainda inteligível para o espectador, sendo esse seu maior objetivo como tradutor.

Outra estratégia de manutenção da harmonização entre legendas e imagens é a condensação. A condensação é uma estratégia que pode ocorrer tanto no nível da palavra quanto da oração (DÍAZ CINTAS e REMAEL, 2007, p. 151-161), conforme os exemplos a seguir:

- Simplificar perífrases verbais. Exemplo: Eu realmente deveria estar indo embora → Eu tenho que ir;
- Generalizar enumerações. Exemplo: Você mentiu para nós, filho. Sua própria mãe e pai → Você mentiu para nós, seus pais;
- Utilizar um sinônimo menor que seja próximo ou expressão equivalente. Exemplo: Ele tem muito dinheiro → Ele é rico;
- Utilizar tempo verbal simples ao invés de composto. Exemplo: Seu pai a tinha expulsado de casa → Seu pai a expulsou de casa;
- Mudar a classe das palavras. Exemplo de troca de verbo para substantivo: Quando o General Pinochet foi preso em Londres → Depois da prisão de Pinochet;
- Quando possível, usar contrações e abreviações características da língua de chegada. Exemplo: Eu tenho que ir, eu te ligo de volta → Tenho que ir, te ligo depois.

Há, ainda, as estratégias utilizadas no nível da oração, sendo elas:

- Alterar negações ou perguntas em orações afirmativas, perguntas indiretas em perguntas diretas. Exemplo: Ok, a gente não vivia em um palácio → Ok, o lugar era pequeno;
- Simplificar indicadores de modalidade. Exemplo: Eu entendo que este pode ser o melhor resultado que podemos ter, politicamente falando, neste exato momento → Essa é a melhor solução política agora;
- Transformar o discurso direto em indireto. Exemplo: Há vezes que eu digo para mim: “Ainda bem que ela se foi, nós estamos melhor desse jeito” → Tem vezes que eu fico feliz por ela ter ido embora. As coisas ficam mais fáceis assim;

- Alterar o sujeito de uma sentença ou frase. Exemplo: Bem, eu penso que entendo o que você quer dizer, d'Ydewalle → Acho que te entendi, d'Ydewalle;
- Manipular a ordem de tema e rema. Exemplo: Lavar roupa, passar ferro, estender roupa, sua avó fazia tudo isso → Sua avó fazia todo o trabalho doméstico;
- Transformar sentenças longas ou compostas em sentenças mais simples. Exemplo: Eu só não te contei porque pensei que você ficaria com muita raiva → Eu não disse nada. Achei que você ficaria furioso;
- Tornar voz ativa em passiva ou vice-versa. Exemplo: A gente sabia há muito tempo que você era maltratada por seu marido → A gente sabe que seu marido te maltrata há muito tempo;
- Usar pronomes (pessoal, demonstrativo, possessivo) e outros dêiticos para substituir substantivos ou sintagmas nominais. Exemplo: Eu sou um cabeleireiro. A única coisa que eu sei fazer é cortar cabelo → Eu sou um cabeleireiro. É tudo que eu sei fazer;
- Unir duas ou mais frases/sentenças em uma. Exemplo: Eu quero saber o que eu posso levar comigo. Eu não quero ser acusado depois de furto → Eu quero saber o que posso levar sem ser acusado de furto.

As omissões também são uma estratégia válida para o tradutor adequar seu texto ao tempo disponível. As reduções podem ocorrer em dois níveis: palavra e oração, variando de acordo com as especificidades da língua alvo. Por parte do nível da palavra, pode-se retirar adjetivos, advérbios, palavras de função fática, cumprimentos, interjeições e repetições, cabendo ao profissional avaliar a necessidade dessas classes de palavras para a coerência de seu texto. Já a redução nas orações pode ocorrer em frases que passam poucas informações, diálogos cruzados ou frases que introduzam um tópico, mas que não contém grande carga informacional.

Diaz Cintas e Remael (2007), entretanto, sugerem que as reduções devem ser usadas em casos de extrema necessidade, principalmente as de orações, visto que, por vezes, a omissão pode atrapalhar a coesão textual e coerência da legenda. Assis (2016) corrobora para a proposição de que a boa coesão entre texto e imagem auxilia o público surdo a compreender melhor a produção audiovisual. Legendar bem, portanto, passa por levar em consideração os

parâmetros técnicos de condensação e omissão, quando se fazem necessários, para que, assim, se possa prover a acessibilidade.

Alguns aspectos técnicos são reafirmados por Diaz Cintas e Remael (2007), que os classificam em duas dimensões: temporal e espacial. A dimensão temporal envolve a marcação e duração das legendas, intervalo entre legendas consecutivas, sincronização, velocidade da legenda e seu tempo de exposição. Já a dimensão espacial apresenta os aspectos do número de linhas, posicionamento em tela, tipografia e número de caracteres por linha. Apresentaremos alguns destes tópicos a seguir.

Em relação à tipografia, é indicado utilizar fontes já conhecidas pelo público e que não possuam serifa, como Arial ou similar, e tamanho suficiente para boa leitura em tela, sendo algo entre 20 e 30 pontos o mais comum (DÍAZ CINTAS e REMAEL, 2007). Na presente pesquisa, utilizamos a fonte Tahoma, tamanho 22, visto que segue os padrões e já receberam boa receptividade em pesquisas anteriores por surdos brasileiros (ARAÚJO, 2012; MONTEIRO e DANTAS, 2017). Para sincronizar as legendas no vídeo, foi utilizado o programa *VLC Media Player*, que automaticamente une a legenda ao seu vídeo, se eles estiverem nomeados da mesma forma.

Quanto ao posicionamento, o ideal é que as legendas sejam posicionadas ao centro, com alinhamento centralizado e no terço inferior da tela. Segundo Díaz Cintas e Remael (2007), esse posicionamento é o mais comum em legendas interlinguísticas, justamente por ser o centro da tela o local onde a ação costuma ocorrer. O posicionamento da legenda alinhada à ação proporciona uma menor fadiga visual para o espectador, já que seu olho precisará viajar menos para acompanhar a leitura e as ações em tela. A despeito do padrão do CC, é comumente usado o alinhamento à esquerda. Há, também, uma variação que segue o falante em tela, parecido com o sistema europeu de legendagem. Esse padrão é mais usual em filmes veiculados pela Rede Globo (ASSIS, 2016).

Ainda como parâmetro técnico, há a marcação e a sincronização. Marcação, para Díaz Cintas e Remael (2007), é o momento em que uma legenda entra e sai de tela. Uma boa marcação leva em consideração os ritmos do filme e as falas dos atores e atrizes, as pausas e maneirismos das falas. O ideal é que a marcação comece em conjunto com a fala do personagem e finalize assim que ela termine. A sincronização, por sua vez, relaciona-se diretamente com a marcação. Ela se caracteriza como a simultaneidade entre a marcação e os diálogos do filme. Na LSE, isso ainda ocorre com os efeitos sonoros. A boa sincronização é

útil para reforçar a coerência entre os elementos sonoros da produção audiovisual e o texto traduzido em legendas.

Em relação ao intervalo entre as legendas, é indicado que se dê um tempo mínimo de pelo menos 80 milissegundos para a saída de uma legenda e a entrada de outra (DÍAZ CINTAS e REMAEL, 2007). Esse intervalo é necessário para que os olhos humanos tenham facilidade para perceber que há novo texto aparecendo em tela. Caso não haja esse intervalo, o espectador pode ignorá-la, acreditando que não houve troca de texto (DÍAZ CINTAS e REMAEL, 2007). As legendas manipuladas neste trabalho seguem o padrão do grupo LEAD de 100 milissegundos de intervalo.

Díaz Cintas e Remael (2007) concordam com os estudos experimentais de d'Ydewalle et al. (1987), que dizem que as legendas devem ter até duas linhas, por ser um padrão ideal para uma leitura confortável (DÍAZ CINTAS e REMAEL, 2007). A LSE que manipulamos está de acordo com os princípios do grupo LEAD – os mesmos parâmetros técnicos da LO interlinguística –, com o diferencial da adição de informações necessárias para os surdos.

Esta pesquisa deu prioridade ao uso de duas linhas de legenda quando o texto passar de 37 caracteres em uma linha. Quanto maior o número de palavras numa legenda até certa quantidade de caracteres, mais rápido será a leitura completa do texto (DÍAZ CINTAS e REMAEL, 2007). Caso o texto seja curto, não há a necessidade de se fazer uso de duas linhas, sendo necessária somente em textos longos. Fica a cargo do legendista quebrar o texto pequeno em duas linhas de tamanho por questões estéticas ou por estar seguindo algum parâmetro de uma empresa (ASSIS, 2016). Díaz Cintas e Remael (2007) elencam pesquisas que indicam que há um maior conforto na leitura de duas linhas e que a preferência é de quebra da legenda sempre que possível, já que o tempo gasto na leitura de cada palavra é menor na última. Para executar a quebra de linha, recorre-se a segmentação linguística.

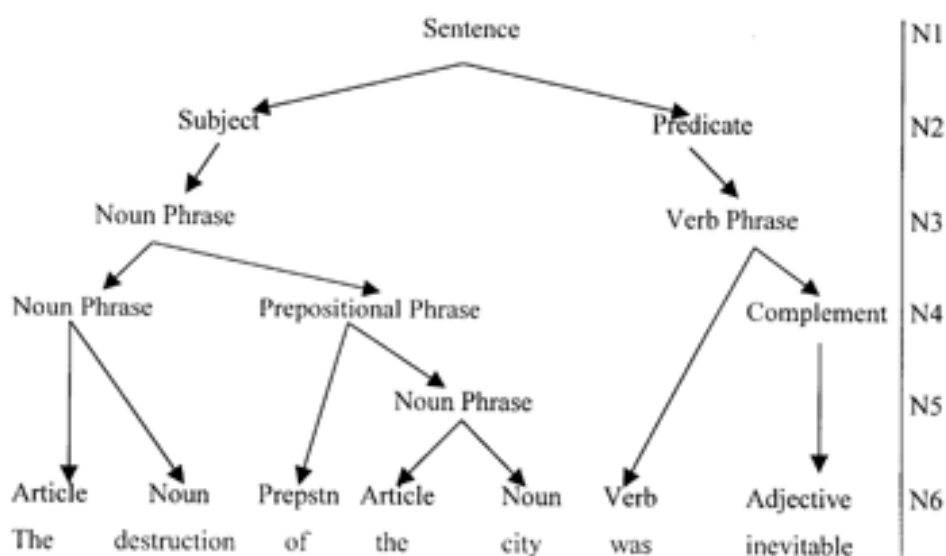
É grande o número de teóricos em Tradução Audiovisual (TAV) que destacam a relevância de uma boa segmentação linguística para a recepção e compreensão da legenda (KARAMITROGLOU, 1998; DÍAZ CINTAS e REMAEL, 2007; PEREGO, 2008). Karamitroglou (1998) afirma que o cérebro é forçado a pausar por um breve momento a decodificação da mensagem quando há a quebra de uma sentença, para que ele a retome depois quando os olhos tiverem localizado o resto dela. Uma segmentação linguística malfeita, portanto, prejudicaria o entendimento das legendas pelo espectador.

Compreende-se por segmentação a divisão da legenda em trechos de texto contínuos. Elas podem ocorrer de três maneiras distintas: visual (um corte ou mudança de cena); linguística (sintaxe, semântica, lexicogramática e fonologia-fonética); e retórica (fluxo da fala) (KARAMITROGLOU, 1998; DÍAZ CINTAS e REMAEL, 2007; ASSIS, 2016; VIEIRA, ASSIS e ARAÚJO, 2020).

Reid (1990) estabelece três critérios para se segmentar a fala na legenda. O primeiro é o linguístico, que seria baseado na sintaxe da frase. Neste caso, cada linha de legenda deve atender a completude em sua extensão sintática. O segundo critério é o retórico, que se baseia no fluxo da fala dos atores em tela. O terceiro e último é o visual, ligado diretamente à mudança de imagem em cena. Ao passo que a imagem na tela muda, a legenda deve seguir a mesma mudança e iniciar ou finalizar o texto referente a cena. É trabalho do legendista analisar e propor mudanças no texto da legenda que respeitem a sintaxe das linhas da legenda, o ritmo e fluxo da fala dos personagens, assim como a relação entre a imagem e a entrada e saída das legendas (REID, 1990).

Karamitroglou (1998) defende que a segmentação deve ocorrer no mais alto nível sintático possível. Para explicar essa proposição, ele apresenta o exemplo da Figura 1, que apresenta “nódulos sintáticos oriundos de uma descrição sintática em árvore de base gerativista” (ARAÚJO e ASSIS, 2014, p. 160).

Figura 1 – Fluxograma da árvore sintática de Karamitroglou (1998)



Fonte: Karamitroglou (1998, p. 7).

A oração em inglês “*the destruction of the city was inevitable*” é dividida em seis nódulos. O primeiro nó (N1) apresenta a oração completa. O nó N2 faz a separação dos sintagmas, que, por sua vez, são compostos, respectivamente, por nominais e verbais (N3). O nó N4 apresenta a composição de dois sintagmas: nominal (*the destruction*) e preposicional (*of the city*), para a frase “*the destruction of the city*”, enquanto que, nos demais, temos o sintagma verbal com o seu complemento. O sintagma preposicional (N4) apresenta, ainda, um sintagma nominal (*the city*) no nível N5. No último nível (N6), temos a composição léxica e suas respectivas funções na estrutura gramatical.

Segundo Karamitroglou (1998), se a frase da Figura 1 estivesse em um bloco de legenda, ela teria que ser segmentada no mais alto nível sintático, no caso “*The destruction of the city*” (sintagma nominal) na linha superior e “*was inevitable*” (sintagma verbal) na inferior.

Araújo e Assis (2014) ainda apontam para o fato de ser comum legendistas quebrarem as linhas de acordo com a quantidade máxima de 37 caracteres por linha. Essa segmentação por quantidade de caracteres pode fazer variar a forma com as quais as legendas serão apresentadas em tela, tendo três tipos distintos de formas. O primeiro tipo, por apresentar quantidade semelhante de caracteres por linha, tem uma forma retangular. Os outros dois, por ter uma leve diferença no número de caracteres, pode apresentar a forma de um triângulo equilátero (ARAÚJO e ASSIS, 2014). O Quadro 1 apresenta essa variação.

Quadro 1 - Formato das legendas em tela

Formato	Legendas
Em forma de retângulo	O guardinha me parou por causa de uma bobagem da placa que caiu!
Em forma de triângulo com a linha de cima maior	Um tutuzinho de feijão, um lombinho.
Em forma de triângulo com a linha de cima menor	[Deolinda] já imaginava, por isso fiz o tutuzinho logo hoje.

Fonte: Araújo e Assis (2014, p. 162).

Para Karamitroglou (1998), o formato em retângulo da legenda em tela deve ter preferência, visto que é o padrão que os espectadores estão já acostumados. Caso necessário, procurando se ater à quebra de linha no maior nível sintático, o legendista pode manter sua legenda no formato triangular. Araújo e Assis (2014) sugerem não haver problemas em legendas de formato triangular na sua pesquisa, assim como em anteriores do grupo LEAD

(CHAVES, 2012; CHAVES; ARAÚJO, 2014). Mais adiante, veremos outros estudos do grupo que exploraram os efeitos da boa e má segmentação na compreensão da obra audiovisual pelo público surdo, além de outros parâmetros técnicos de relevância na LSE.

Para uma boa compreensão da importância da tradução de efeitos sonoros na LSE, apresentaremos, a seguir, os conceitos e implicações que os sons têm nas obras fílmicas. O trabalho de Nascimento (2018) estabelece um estudo descritivo sobre a legendagem de sons e sua importância técnica basilar para esta pesquisa. O estudo de Aleksandrowicz (2019), por sua vez, explorou o impacto nas emoções de surdos por meio de legendas que exibem a tradução de efeitos sonoros na língua polonesa.

2.4 O papel do som no cinema

O som surge no cinema muito antes de *O Cantor de Jazz (The Jazz Singer)*, de 1927, com seus poucos diálogos e várias trilhas sonoras, ainda que a obra seja tida como pioneira na junção sincronizada de som e imagem em um filme (BORDWELL e THOMPSON, 2014). Os anos iniciais do cinema, entre 1895 e 1903, foram marcados por muita experimentação e pela formulação de uma linguagem visual (COUSINS, 2013). Os anos que se seguiram viram os produtores e técnicos do cinema testarem métodos distintos de se pôr áudio na exibição fílmica, mesmo que não pudesse estar diretamente na película. Segundo Cousins (2013), vários foram os métodos de levar o som para a sala de exibição, como orquestras, dublagem e sonorização ao vivo, além de gramofones tocando músicas para a plateia e alguns deles foram testados.

A tentativa de se aliar o áudio ao vídeo no cinema data, portanto, dos seus primeiros dias de vida. A tecnologia para isso, que impedia a fusão de áudio e vídeo à época, foi aos poucos evoluindo e modificando o cinema. Nos anos posteriores ao lançamento de *O Cantor de Jazz* (1927), o maquinário pesado e rústico ainda atrapalhava a captura de áudio in loco. A engenhosidade e criatividade driblaram os impedimentos técnicos e o cinema entrou em sua era dourada, com o som como principal revolução técnica e tecnológica (COUSINS, 2013).

Desde que a trilha de som foi inserida nas películas, nunca mais se pensou, nem se fez, cinema do mesmo modo (ALLEYNE, 2009). Lehman (2018) afirma que uma boa trilha sonora passa despercebida do público; ela deve estar tão intrínseca ao vídeo que os dois pareçam se fundir em um meio semiótico só. O que não se percebe é que a trilha sonora, junta

aos seus ruídos, guiam e manipulam o espectador em relação a como ele deve reagir à ação em tela. O som, portanto, é um agente de informação capaz de amplificar ou diminuir uma ação e, por consequência, o sentimento dos espectadores.

Bordwell (2008) categoriza os sons presentes no cinema moderno em três diferentes tipos: fala, efeitos sonoros e música. A fala é tudo que é dito, ou pensado, pelos atores e atrizes. No roteiro, as falas são demarcadas como blocos de diálogos (COMPARATO, 2018). Efeitos sonoros, ou mesmo ruídos, são todos os tipos de sons que adicionam uma atmosfera ou um feedback sonoro a alguma ação que aconteça em cena (LEHMAN, 2018). Esse tipo de som é predominante em obras audiovisuais, podendo ser de tipos variados, e se tornou fundamental para dar entendimento e verossimilhança à história (LEHMAN, 2018; LASTRA, 2000). A música, por sua vez, adota diversos usos, desde a própria ambientação até o reforço da ação presente em tela, como uma cena de romance com uma música leve ou uma cena de ação com batidas mais fortes e aceleradas (BORDWELL e THOMPSON, 2014).

Os sons, ruídos e música podem ter diversos usos narrativos em um filme. A ambientação de uma cena, por exemplo, pode ocorrer por meio de ruídos característicos deste espaço. Esse tipo de som é chamado de *ambience* (HARPER, DOUGHTLY e EISENTRAUT, 2009). Se uma cena ocorre em uma cafeteria, como as cenas iniciais e finais de *Pulp Fiction* (1994)², é natural que se ouça outras pessoas comerem, bater de talheres e conversas avulsas. A ausência desses sons característicos pode causar estranhamento ao espectador, já que aquela cena não parecerá natural. Deve-se adicionar, portanto, todos os ruídos esperados daquele lugar, para não só ambientar o espectador como dar verossimilhança à cena.

Outra função da trilha sonora como um todo é a sensação de continuidade que ela provoca. Os sons podem funcionar como uma cola entre cenas localizadas em espaços diferentes, ou até entre centros narrativos distintos (HARPER, DOUGHTLY e EISENTRAUT, 2009). Sua necessidade aqui é dada para que o ritmo se mantenha, ou até mesmo acelere, entre as cenas e locações diferentes, mantendo o espectador a par do que está acontecendo sem problemas e travas narrativas.

² *Pulp Fiction*, lançado no Brasil como *Pulp Fiction: Tempo de Violência*, é um filme estadunidense de drama policial do diretor Quentin Tarantino. Lançado em 1994, foi indicado a várias premiações, tendo ganho o Oscar de melhor roteiro original.

Há uma categorização ainda mais específica para os efeitos sonoros em relação às ações narrativas do filme: sons diegéticos e não diegéticos. Diegese é tudo que faz parte do contínuo espaço-tempo da realidade ficcional de uma obra (HARPER, DOUGHTLY e EISENTRAUT, 2009). Sons diegéticos, dessa forma, são aqueles que fazem parte da narrativa do filme, como se ele estivesse presente diante dos atores em cena e podem ser tanto *on-screen* como *off-screen*. Os sons *on-screen* são aqueles percebidos dentro do quadro da cena, fazendo parte, assim, da relação dos atores com a ação dramática. Os sons *off-screen*, por sua vez, podem ser ouvidos pelos personagens, mas não estão presentes dentro do quadro ou ambiente ficcional. Deve-se salientar que sons *off-screen* nem sempre são diegéticos, podendo fazer parte somente de um som direcionado ao espectador e que irá adicionar peso dramático à ação (HARPER, DOUGHTLY e EISENTRAUT, 2009).

Ainda temos os sons não diegéticos, que são classificados como sons que não estão presentes na realidade ficcional de uma trama, mas que o espectador pode ouvir. São exemplos de sons não diegéticos uma música tema, o *leitmotiff*, ou uma antecipação sonora de uma ação abrupta que ainda irá ocorrer com o personagem em quadro (HARPER, DOUGHTLY e EISENTRAUT, 2009).

O *leitmotiff* ainda é outro tipo de música, que funciona como um precursor sonoro de um personagem ou ação. O personagem Darth Vader, do filme *Star Wars* (1977), tem em seu tema a caracterização de um militar ameaçador e vil, que o precede para causar medo e temor no espectador. A história e personalidade do personagem são contadas por meio de seu tema musical.

Estes diversos tipos de sons e músicas são de fundamental importância na tradução para surdos, já que o cinema é a junção de imagem e som. Ambos não podem se dissociar sem que o espectador perca alguma informação (ALLEYNE, 2009).

Sobre a tradução dos efeitos sonoros, abordaremos o estudo de Nascimento (2018), que observou descritivamente essa tradução na LSE de filmes em DVD. Conforme a autora, a tradução dos efeitos sonoros na LSE auxilia a construção do enredo para os espectadores surdos. Por outro lado, a autora enfatiza que, quando o som é irrelevante para a ação, a tradução de ruídos e músicas pode desviar a atenção do espectador. Assim, Nascimento (2018) ressalta a necessidade de se estabelecer padrões para a tradução de músicas e ruídos na LSE. Para tal, ela estabeleceu um sistema de qualificação de músicas em um enredo e de tradução de efeitos sonoros mediante seu significado.

A autora se desencoraja a hipótese da tradução completa de sons em tela para as legendas e o uso do símbolo da nota musical para indicar alguma música sendo tocada. Ao invés de traduzir todos os sons, diegéticos e não diegéticos, o tradutor deve preocupar-se em traduzir os sons relevantes para a ação dramática. Após esta tradução, é de fundamental importância manter o padrão elaborado inicialmente, caso haja repetição do som ou música posteriormente. Em relação às músicas, o tradutor deve indicar o sentimento que aquela música quer passar e manter sua padronização ao longo da tradução.

O estudo de Nascimento (2018) teve como objetivo propor diretrizes e convenções na tradução das informações sonoras da LSE. Essas diretrizes podem auxiliar e facilitar o trabalho do legendista quando este for elaborar a tradução de efeitos sonoros na LSE. A direção da pesquisa, portanto, foi de simplificar o processo de transformar sons em palavras por parte do profissional da área.

A autora criou, então, um *corpus* comparável composto de 15 filmes legendados de três países diferentes: Brasil, Estados Unidos e França. Os países estrangeiros foram escolhidos por apresentarem maior tradição quanto à acessibilidade de suas produções audiovisuais. A autora categorizou as legendas de acordo com as etiquetas estabelecidas em sua dissertação de mestrado, além de analisá-las com o mesmo programa utilizado para executar a etiquetagem, o WordSmith Tools.

Na análise quantitativa, Nascimento (2018) descobriu uma ampla gama de sons possíveis de serem traduzidos e legendados. Dentre eles, têm-se os sons produzidos pelo homem, objetos, natureza e o próprio silêncio (NASCIMENTO, 2018). Em relação às músicas, há a categorização de música de fosso³ e de tela, a primeira não sendo participante da diegese fílmica, enquanto a segunda sim. As músicas ainda podem ser subcategorizadas em qualificadas, quando apresentam adjetivação; e não qualificadas, quando informam a presença da música no filme. Nascimento (2018) também aponta para a cautela na tradução dos efeitos sonoros, para não apresentar um excesso de legendas em tela que possam ser desnecessárias para o espectador surdo e, até mesmo, atrapalhá-lo. Pelo fato da pesquisa de Nascimento ser bastante abrangente, não iremos abordar todos os tipos de sons exemplificados em seu trabalho; focaremos na explanação dos tipos de sons e músicas presentes nos trechos do filme *Dunkirk* (2017), objeto de nosso estudo.

³ Música de fosso, ou não diegética, é aquela que vem de fora da tela, a qual somente o espectador pode ouvir e os personagens fictícios não (HARPER, DOUGHTLY e EISENTRAUT, 2009).

Para a tradução dos ruídos, Nascimento (2018) recomenda utilizar os caracteres entre colchetes e somente a primeira letra em maiúsculo. Há a possibilidade de fazer uso de três diferentes tipos de quantidade de informação na legenda: básico, informativo e superinformativo. O tipo básico refere-se somente à ação que ocorre no canal sonoro de maneira mais simplificada possível. A tradução da ação “passos”, para descrever o caminhar de alguém, portanto, seria escrita na legenda: “[Passos]”. No segmento informativo, há a adição de um adjetivo, para tornar o substantivo qualificado. Seguindo o exemplo anterior, a legenda poderia ser: “[Passos apressados]” ou “[Passos cautelosos]”. O tipo superinformativo fornece ao espectador todas as informações possíveis sobre o som. No mesmo exemplo, a legenda seria: “[Passos afastam-se cautelosamente]”.

Outra classe de tradução é a das músicas. Utilizaremos somente a classificação de música de fosso, já que é a única presente nos trechos do filme escolhido. Apesar de não tão padronizadas quanto os ruídos, elas se apresentam geralmente em uma oração ou locução verbal. As três principais categorias semânticas da música de fosso são: qualificada, não qualificada e implícita. Qualificada para quando se atribui algum adjetivo para a música, como: “[Música feliz]”. Não qualificada quando não se adjetiva a música, geralmente para indicar um instrumento musical ou somente a presença de música. Por fim, há a música implícita, onde se indica a alteração no estado de uma música já citada anteriormente, como “[Música feliz para]”.

Todas essas categorias semânticas foram utilizadas nas legendas dos quatro trechos do filme, visto que a música se faz presente nas cenas, sendo essencial para dar ritmo e apelo dramático a elas. As propostas de tradução dos efeitos sonoros desta pesquisa tiveram como base as convenções propostas por Nascimento (2018), e os sons mais comuns foram os produzidos pelo ser humano e as músicas de fosso. Nos trechos do filme selecionado na presente pesquisa, há a predominância de ruídos de tiro e falas, sendo, assim, sons feitos por objetos e humanos. As músicas predominantes nos extratos são as não diegéticas, ou de fosso; músicas que não estão presentes nas ações sofridas pelos personagens em tela.

Nascimento (2018) também classifica e orienta a tradução de onomatopeias. Não se notou, entretanto, nenhuma interjeição que fosse de grande necessidade para o entendimento das ações pelo espectador nos trechos manipulados para a presente pesquisa. Portanto, não se viu a obrigação do uso desta classe de legendagem sonora na presente pesquisa.

As conclusões da pesquisa de Nascimento são, primeiramente, de que os legendistas brasileiros costumam traduzir mais as informações do canal sonoro que os profissionais da França e dos Estados Unidos. Esse excesso na tradução pode significar maiores desvios da atenção dos surdos quando assistem a produtos audiovisuais. Dentre todos os tipos de sons, as categorias mais frequentes são: sons produzidos pelo homem e música de fosso. Nascimento (2018) ainda observou que esses sons possuem três estruturas básicas: apenas o substantivo, quando este já é o suficiente para traduzir o som: “[Choro]”; colocações nominais ou adjetivas, a fim de se qualificar o ruído: “[Choro forte]”; e orações, quando é necessário que a legenda exprima ação: “[Paula Chora]”.

Quanto às músicas, Nascimento (2018) observou que os norte-americanos praticamente não traduzem as músicas de fosso, sendo esse trabalho mais comum para brasileiros e franceses. Depreende-se desse fato a prioridade norte-americana em traduzir somente as informações diegéticas do filme. A importância dada aos legendistas do Brasil e da França pode ter ligação com a importância da música de fosso nessas culturas, como se fosse uma espécie de guia de expectativas no meio multissemiótico.

2.5 Pesquisa de recepção com a tradução de efeitos sonoros na LSE

Levando em consideração e aplicando boa parte do arcabouço técnico até aqui apresentado, situa-se a pesquisa de Aleksandrowicz (2019). Esse estudo experimental, com 86 participantes, sendo 60 surdos e 26 ouvintes, fez uso de dois trechos de uma *database* estabelecida por Schaefer et al. (2010) de segmentos de filmes. O intuito do estudo era verificar se diferentes abordagens da legendagem de sons e músicas em trechos de filmes suscitavam alguma resposta emocional diferente por parte dos surdos.

Os trechos selecionados dos filmes buscavam provocar sentimentos nos espectadores. A seleção dos trechos partiu de um estudo em que Schaefer et al. (2010) solicitaram a 50 especialistas em cinema que sugerissem até três excertos de filmes que pudessem provocar os seguintes sentimentos: raiva, tristeza, medo, nojo, alegria, surpresa e emoção neutra. Foi levantado um total de 827 clipes, sendo reduzidos ao total de 10 para cada emoção. Essa pesquisa ainda passou os excertos para 364 participantes, a fim de estabelecer um ranqueamento dos trechos mais poderosos no estímulo de emoções (ALEKSANDROWICZ, 2019).

Aleksandrowicz (2019) fez uso de dois cliques dessa *database* em sua pesquisa e mais um de sua escolha. O primeiro foi do filme *O Iluminado* (1980), apresentando 3 minutos e 49 segundos de duração, e estando na 2ª posição de filme que mais causava medo. O segundo trecho veio do filme *Os Últimos Passos de um Homem* (1995), de 5 minutos e 35 segundos de duração, e ranqueado em 6º lugar na suscitação do sentimento de tristeza. O clipe que não fazia parte da lista foi o trecho inicial de *O Iluminado* (1980), de 1 minuto e 40 segundos, que apresenta a trilha sonora de *Dies Irae* e nenhum diálogo. Esse trecho foi utilizado para saber se era possível provocar emoções – medo, no caso –, por meio, somente, da descrição da música, mesmo não apresentando estímulo visual para tal.

Os cliques foram escolhidos de maneira proposital, por não apresentarem grande quantidade de diálogos, já que a ação é contada mais pela imagem e pelo som, o que o pesquisador acredita facilitar o enfoque da ação pelo espectador surdo (ALEKSANDROWICZ, 2019). O outro motivo é que a maior parte dos sons são não diegéticos e seu propósito é de valorizar e amplificar os efeitos do canal visual das cenas.

Aleksandrowicz (2019) legendou os três trechos em LSE da seguinte maneira: sem informação de música, informação de humor e informação do autor e nome da trilha. As legendas apresentaram as seguintes informações para o mesmo trecho, respectivamente: nenhuma informação, “[MÚSICA ASSUSTADORA]” e “[MÚSICA: Krzysztof Penderecki, ‘De Natura Sonoris No. 2’]”. Os falantes eram explicitados em duas cores distintas de legenda – ciano e amarelo – e a velocidade não excedia os 12 caracteres por segundo, padrão polonês, onde a pesquisa foi executada. Ressaltamos que esse modo de legendar música e sons é diferente do aplicado nesta pesquisa, visto que utilizaremos o modelo de Nascimento (2018) e ele se difere em alguns aspectos do utilizado por Aleksandrowicz (2019).

Para mensurar o impacto das emoções, Aleksandrowicz (2019) fez uso do questionário qualitativo DES (*Differential Emotion Scale*) (IZARD, 1977), em polonês. Os participantes foram divididos em quatro grupos: um de ouvintes, ou controle, e três experimentais para cada tipo de LSE distinta. Todos assistiam aos trechos e, logo após verem cada excerto, preenchiam um questionário de autopercepção de suas emoções. A escala ia de 0 para “não sentiu” a 4 para “sentiu bastante”. Os resultados foram analisados via Análise de Variância (ANOVA).

No clipe 1, os participantes reportaram sentir medo e, no segundo, a emoção mais sentida foi a tristeza. Os resultados foram cruzados e o autor encontrou uma proximidade muito grande entre os três grupos de LSE diferentes, tendo desvios estatísticos não

significantes. Em relação ao terceiro trecho, em que a legenda tentava suscitar medo pela sua descrição, não houve diferença estatística significativa para as demais emoções reportadas. Entretanto, as legendas com a descrição do humor suscitaram mais interesse por parte dos surdos do que o grupo controle formado somente por ouvintes.

Como conclusão, Aleksandrowicz (2019) afirma que surdos e ouvintes percebem a intensidade das emoções de maneira muito próxima, independentemente do tipo de tradução de efeitos sonoros na LSE. Sendo assim, a hipótese derivada deste estudo é que talvez a música não tenha tanto impacto quanto os estudiosos do cinema costumam atribuir no âmbito da legendagem para surdos. Para os legendistas, Aleksandrowicz (2019) sugere que em cenas que a ação é explícita e há sons que corroboram o que é visto em tela, a tradução de efeitos sonoros pode ser omitida sem grandes impactos para os surdos. Seria até mesmo preferível a omissão nestes casos, já que a legenda poderia atrapalhar na atenção à imagem. Esse resultado vai ao encontro do proposto por Nascimento (2018), a qual também propõe que sons que já são notados pela cena não sejam adicionados na tradução pela LSE.

Apesar de apresentar algumas semelhanças com os parâmetros de Nascimento (2018), há diferenças nas escolhas tradutórias dos sons entre o modelo proposto por Nascimento e o modelo aplicado por Aleksandrowicz (2019).

O estudo de Aleksandrowicz (2019) apresentou indícios dos resultados que podemos obter em nossa pesquisa, entretanto, ele não fez uso do mesmo modelo de tradução de efeitos sonoros que esta pesquisa, além do fato que seus participantes pertencem a uma cultura diferente da nossa. Essas distinções validam a necessidade de explorar o modelo de Nascimento (2018) sobre o impacto nas emoções dos surdos, a fim de verificar se os resultados apresentarão diferenças entre os modelos.

Veremos, a seguir, as pesquisas mais relevantes para este estudo que ocorreram no âmbito do grupo LEAD-UECE. Os trabalhos serão apresentados no contexto dos projetos nos quais se alocaram: Projeto MOLES (Modelo de Legendagem para Surdos), que buscou a estabilização de um padrão para a LSE no Brasil; e ExLEG, que testou parâmetros técnicos, linguísticos e experimentais em LSE.

2.6 Pesquisas no contexto do grupo Legendagem e Audiodescrição (LEAD)

As pesquisas desenvolvidas na Universidade Estadual do Ceará (UECE) pelos pesquisadores do grupo LEAD revelaram novos parâmetros sobre a LSE, até então não tão

explorados no Brasil (FRANCO e ARAÚJO, 2003; ARAÚJO, 2004; ARAÚJO, 2008; ARAÚJO, 2012). Esses estudos revelaram a necessidade de desenvolver mais pesquisas de recepção com participantes surdos, a fim de compreender as necessidades desse público (FRANCO e ARAÚJO, 2003).

Araújo (2005) acreditava que, para um bom entendimento dos surdos da produção fílmica, seria necessária a condensação das palavras, a fim de reduzir a velocidade e manter a sincronia do texto com a fala. A autora também concluiu as legendas no modelo americano de *closed caption* não eram eficazes para a recepção confortável por parte dos espectadores surdos.

A pesquisa de 2008 tratou-se de um estudo exploratório, com a participação de 12 surdos cearenses voluntários (ARAÚJO, 2008). O objetivo dessa pesquisa era estabelecer as bases para um modelo de LSE no Brasil. Os surdos foram expostos a diferentes gêneros audiovisuais da TV brasileira – novela, documentário e humorístico – e a dois tipos de legendas distintas: as originais e outras manipuladas pela equipe de pesquisadores. No final, os participantes responderam a questões sobre o conteúdo dos vídeos. Os resultados desse estudo foram importantes para demonstrar a relevância de quatro pontos a serem analisados sobre a recepção da LSE: velocidade, formato da legenda, uso da norma culta e não discriminação de músicas em texto. Ainda nesse estudo, surdos cearenses decidiram que a velocidade ideal seria a de 145 palavras por minuto (ppm), consideradas legendas de velocidade baixa, segundo o padrão de Díaz Cintas e Remael (2007). Nos demais parâmetros técnicos sugeridos pelos surdos, Araújo (2008) conclui que as legendas deveriam ser de cor amarela em fundo transparente, o falante sempre identificado entre colchetes quando o turno de fala trocar, e que a voz em *off* deve vir em itálico. Os efeitos sonoros devem vir expressos nas legendas, mas a música deverá apresentar somente o símbolo musical de colcheia e a norma culta do português deve sempre ser respeitada.

Ressaltamos que alguns desses resultados foram refutados com os estudos a posteriori, em que, por exemplo, a velocidade mais rápida obteve melhores resultados que as mais lentas (VIEIRA e ARAÚJO, 2017). O Projeto MOLES (2012), desenvolvido pelo grupo LEAD, adicionou novas variáveis à equação da boa LSE, como a importância da segmentação linguística, para o melhor entendimento dos conteúdos audiovisuais, dentre as quais, é possível citar que a velocidade não é uma causa de ruído no entendimento.

2.6.1 Projeto MOLES

A pesquisa desenvolvida no campo do Projeto MOLES (Modelo de Legendagem para Surdos) explorou o uso de velocidades distintas com surdos de quatro regiões brasileiras, além do emprego da segmentação como fator importante para o resultado obtido, preconizada por Karamitroglou (1997), como meio de otimizar o uso dos caracteres nas legendas e na compreensão dos textos pelo público surdo (ARAÚJO e NASCIMENTO, 2011; ARAÚJO, 2012). As descobertas apontaram para a irrelevância da velocidade das legendas, desde que elas fizessem bom uso da segmentação linguística, ou seja, quebrassem a linha da legenda seguindo o mais alto nível sintático, conforme Karamitroglou (1997).

O objetivo do Projeto MOLES foi investigar dois modelos de LSE: o proposto pelo grupo LEAD-UECE e o modelo português de LSE, em que as legendas são colocadas sobre os falantes, de cor branca para fala e cor amarela para a identificação do falante e do efeito sonoro, respectivamente (ARAÚJO, 2012).

O modelo proposto pelo grupo LEAD foi gerado a partir do estudo de recepção com 12 surdos cearenses (ARAÚJO, 2008). Ambos os modelos foram postos a teste com 34 surdos de quatro regiões do Brasil, sendo de pelo menos dois estados diferentes por região. A pesquisa previa 50 participantes nas cinco regiões brasileiras, mas o Centro-Oeste não participou e somente um estado do Norte atendeu ao chamado. O perfil dos participantes foi de surdos de nascença, em torno de 25 anos de idade, universitários, sinalizados e tendo aprendido Libras com 10 anos de idade em instituições especializadas. A maioria assistia a filmes legendados, tanto no cinema quanto na TV. Na pesquisa, eles assistiram a quatro filmes diferentes de curta metragem, cada um com um parâmetro distinto a ser avaliado, sendo eles: velocidade de 145ppm, 160ppm e 180ppm e o sistema de LSE português. Apesar de não ter a participação de todas as regiões, os resultados foram considerados conclusivos.

O projeto tinha como hipótese as seguintes afirmativas: 1) quando a velocidade é 145ppm, o surdo consegue entender o conteúdo do filme e seus detalhes; 2) quando a velocidade é 160ppm, o surdo consegue entender o conteúdo do filme, mas não seus detalhes; 3) quando a velocidade é 180ppm, o surdo não consegue entender nem o conteúdo nem os detalhes do filme; 4) a recepção do modelo europeu prejudica a recepção. A primeira hipótese foi confirmada, visto que a maioria dos participantes relatou ser a velocidade ideal; já as hipóteses dois, três e quatro foram refutadas. Em relação às hipóteses que consideravam as velocidades um fator de dificuldade, os pesquisadores encontraram recepções de altas

velocidades tão eficientes quanto a velocidade mais lenta. Quanto ao modelo europeu, os surdos apresentaram recepção igual aos resultados apresentados anteriormente. A conclusão dos pesquisadores foi que a segmentação seria mais importante para a coesão textual e recepção pelo público do que a velocidade.

O fato de a velocidade mais rápida ter sido tão eficiente quanto a mais lenta, uma conclusão até então contraintuitiva, serviu para abrir os horizontes do grupo para pesquisas que abordassem outros parâmetros tanto técnicos quanto linguísticos e tradutórios na legendagem. Os resultados abriram caminho para estudos experimentais pelo próprio grupo, que se debruçou sobre a influência da segmentação na recepção das legendas.

A seguir, apresentaremos algumas pesquisas experimentais desenvolvidas pelo grupo LEAD. As pesquisas avaliaram a influência da segmentação e velocidade em dois gêneros audiovisuais distintos: propaganda política e documentário.

2.6.2 Pesquisas Experimentais em Legendagem

O projeto Pesquisas Experimentais em Legendagem (ExLEG) nasceu em 2015, dentro do grupo LEAD. Surge depois que as pesquisas do grupo apontaram para a segmentação linguística como fator determinante para a boa compreensão da LSE pelos surdos. O projeto possui diversas pesquisas experimentais em seu escopo, fazendo uso, principalmente, do *eye tracking* como ferramenta de coleta de dados relativos ao comportamento ocular durante a leitura das legendas.

Os primeiros trabalhos abordados são os de Vieira (2016) e Vieira e Araújo (2017)⁴. Neles, as pesquisadoras investigaram as influências inerentes à boa e má segmentação linguística no gênero documentário do programa Globo Repórter, veiculado pela Rede Globo de Televisão. Vieira (2016) e Vieira e Araújo (2017) selecionaram legendas provenientes do sistema fechado e do tipo *Roll-up*. As pesquisadoras, então, manipularam-nas para que se adequassem às velocidades de 145ppm (lenta) e 180ppm (rápida), seguindo os parâmetros estabelecidos por Díaz Cintas e Remael (2007).

⁴ A pesquisa de Vieira (2016) trata-se da sua tese de doutorado, em que ela investiga os parâmetros da segmentação no gênero documentário. O trabalho de Vieira e Araújo (2017) trata-se de um artigo escrito pela mesma autora, que traz os dados de sua tese de maneira resumida e concisa.

Foram selecionados ao total 16 participantes, sendo 8 surdos formando o grupo experimental e 8 ouvintes para o grupo controle. Todos os participantes se declararam espectadores de obras audiovisuais legendadas. A idade variou de 19 a 41 anos e o nível escolar era de graduandos e graduados. Todos os surdos apresentavam surdez pré-linguística, mas afirmaram serem fluentes na língua portuguesa. Nessa pesquisa, os ouvintes assistiram aos vídeos sem o auxílio do canal sonoro, a fim de mantê-los o mais próximo possível da condição dos surdos.

A pesquisadora criou quatro condições para as legendas: Lenta Bem Segmentada (LBS), Lenta Mal Segmentada (LMS), Rápida Bem Segmentada (RBS) e Rápida Mal Segmentada (RMS). Os padrões técnicos anteriormente citados foram seguidos, como cor da legenda, tamanho e fonte utilizada, em todas as condições de maneira idêntica. As únicas variáveis em relação aos parâmetros técnicos das legendas eram segmentação e velocidade. Os participantes de ambos os grupos foram expostos a cada uma das quatro condições, visto que a autora fez uso do design do quadrado latino, em que todos são expostos pelo menos uma vez a cada uma das condições da pesquisa. Foram executadas 64 gravações, compostas por 360 legendas ao total, sendo 160 lentas e 200 rápidas.

Vieira (2016) e Vieira e Araújo (2017) fizeram uso de medidas quantitativas e qualitativas. Para a análise quantitativa, as pesquisas fizeram uso de diversas variáveis dependentes ligadas ao movimento ocular, como número de fixações, deflexões e regressões. Essas variáveis foram capturadas pelo *software* de rastreamento ocular Tobii Studio 3.2. Quanto às medidas qualitativas, as autoras utilizaram-se de relatos livres e guiados. Durante os relatos livres, os participantes falavam sobre o que compreenderam do conteúdo dos vídeos. O relato guiado, por sua vez, solicitou que os participantes falassem sobre suas impressões em relação às legendas.

Como resultado, Vieira (2016) e Vieira e Araújo (2017) obtiveram constatações sobre o comportamento ocular de surdos e ouvintes, além de variações entre resultados causados pelas velocidades e pela segmentação. De início, os resultados apontaram pouca diferença no comportamento ocular entre ouvintes e surdos, apesar destes demonstrarem ler mais devagar do que aqueles. A condição RBS, no entanto, mostrou-se a mais favorável para a leitura e compreensão do grupo surdo. Seguindo o teste ANOVA para análise de dados, o resultado estatístico obtido foi que a condição RBS provou ser a melhor para ambos os grupos, sendo a má segmentação linguística mais custosa em ambos os cenários em que a velocidade permaneceu a mesma.

O fato de os surdos se adaptarem melhor a legendas mais rápidas pode se dar devido à possível capacidade aprimorada de leitura parafoveal. Essa capacidade permite a eles ter maiores indícios de quais serão as próximas palavras na legenda (VIEIRA, 2016).

Na velocidade lenta, independente da qualidade da segmentação, os surdos fizeram fixações mais longas no texto. Fixações mais longas nas legendas podem indicar que o espectador passou mais tempo olhando para o texto escrito do que para a ação em tela. A atenção do espectador sendo monopolizada na legenda pode acarretar a perda de informação visual para ele, causando, assim, menor compreensão do produto como um todo. As legendas lentas, segundo Vieira (2016), seriam mais custosas para o público surdo, indo de encontro aos primeiros estudos do grupo LEAD, em que acreditavam que a velocidade rápida seria um entrave para a recepção dos surdos (ARAÚJO, 2008).

A hipótese de que a má segmentação afeta negativamente a compreensão do espectador, seja surdo ou ouvinte, foi confirmada. Ambos os grupos apresentaram resultados piores, independentemente da velocidade, quando a variável segmentação foi comparada entre suas duas possíveis condições. As demais variáveis dependentes, como número de deflexões, regressões e releituras, permitiram conclusões semelhantes às já apresentadas na condição RBS, sendo esta condição a melhor para o público surdo, quando a questão é o gênero documentário televisivo.

Monteiro e Dantas (2017)⁵, participantes do mesmo projeto ExLEG, analisaram o comportamento ocular de surdos e ouvintes, também fazendo uso da técnica do rastreamento ocular, mas, dessa vez, com o gênero propaganda política. As conclusões de Monteiro e Dantas (2017) e, posteriormente, de Souza (2020), conversam positivamente com os resultados de Vieira (2016), por isso veremos com mais detalhes essa pesquisa a seguir.

A pesquisa de Monteiro e Dantas (2017) é também caracterizada como experimental e buscou estudar o impacto da segmentação e velocidade nas propagandas políticas do ano de 2010, veiculadas em Fortaleza, Ceará. Os autores seguiram os mesmos passos metodológicos de Vieira (2016) e manipularam as legendas nas mesmas condições: LBS, LMS, RBS e RMS. O objetivo era averiguar até que ponto a boa segmentação e a

⁵ Esta pesquisa é um recorte de Monteiro (2016) onde analisou-se estatisticamente os dados de Monteiro (2016). A pesquisa de Monteiro (2016) foi realizada concomitantemente com a pesquisa de Vieira (2016). Sendo este o motivo por apresentarem os mesmos instrumentos de coleta e participantes.

velocidade impactavam positivamente na compreensão pelo público surdo e ouvinte, bem como a má segmentação poderia prejudicar a recepção.

Foram quatro vídeos escolhidos, de políticos candidatos ao governo do estado do Ceará e à presidência do Brasil. As propagandas continham até um minuto de duração e suas legendas foram manipuladas no mesmo programa utilizado por Vieira (2016), o *Subtitle Workshop*. As velocidades também foram de 145ppm e 180ppm. Em relação à segmentação, os critérios para caracterizá-las foram: como ruim, quando a quebra dos sintagmas foi feita indevidamente; e, como boa, quando os sintagmas foram devidamente separados sem romper o aspecto semântico. Ressaltamos que a segmentação ocorrida nesta pesquisa foi entre diferentes legendas, visto que a legenda da propaganda política no Brasil é veiculada em linha única. As quatro condições da pesquisa geraram um total de 16 vídeos e 507 legendas.

Essa pesquisa também contou com oito surdos fluentes na língua portuguesa e oito ouvintes, sendo esses os mesmos participantes da pesquisa de Vieira (2016). Assim, o perfil também foi formado por participantes com ensino superior completo ou em andamento e com idades variando entre 19 e 41 anos. Os ouvintes formaram o grupo controle e os surdos o experimental. Todos responderam previamente a um questionário demográfico e afirmaram assistir a programas legendados com frequência.

Os pesquisadores utilizaram-se de uma medida quantitativa durante a coleta, o equipamento Tobii TX300 Eye Tracker, para capturar o movimento ocular, e três medidas qualitativas depois da coleta: uma chamada “relato retrospectivo livre”, outra “relato retrospectivo guiado” e um questionário pós-coleta, o mesmo utilizado por Vieira (2016). Imediatamente depois de terem assistido a cada trecho, os participantes faziam o relato livre sobre o que havia entendido do conteúdo do vídeo. Eles também tinham que responder ao questionário, mas este tinha o foco nas questões técnicas da legendagem. Após o integrante ter assistido aos quatro vídeos e respondido ao questionário e ao relato livre, ele respondia ao relato guiado. Nessa fase, os participantes assistiam ao próprio trajeto ocular, enquanto os autores questionavam-nos se algo, geralmente a má segmentação, havia prendido sua atenção naquele momento.

A fim de analisar os resultados obtidos, Monteiro e Dantas (2017) separaram em duas áreas de interesse distintas a imagem e as legendas. Para medir numericamente as variações, os pesquisadores elegeram um escopo menor que Vieira (2016), contendo somente o número de fixações e sua média de duração, dando mais peso aos resultados qualitativos, uma vez que os autores não conseguiram concluir a análise estatística. Para Monteiro e Dantas

(2017), a análise dos dados quantitativos revelou que a condição LBS, a despeito de Vieira (2016), apresentou melhor condição de leitura para surdos e ouvintes, já que o número de fixações foi menor na legenda do que na imagem. Nos relatos, os participantes consideraram a legenda como tendo boa velocidade e fácil acompanhamento, além de fornecerem detalhes do conteúdo dos vídeos.

A condição LMS, entretanto, dificultou a recepção tanto de surdos quanto de ouvintes, já que a duração média das fixações foi maior nas legendas em vez da imagem. Os participantes relataram alta velocidade da legenda, mesmo ela estando lenta. Esse fato pode ter ocorrido pelo prejuízo que a má segmentação causa na percepção e no processamento da informação pelo cérebro (MONTEIRO e DANTAS, 2017). Os surdos foram os que mais reclamaram dessa condição, sendo, também, os que menos compreenderam seus conteúdos.

Na condição RMS, os surdos apresentaram mais dificuldades do que os ouvintes. Os ouvintes chegaram a apresentar conforto na leitura das legendas, visto que suas fixações duraram mais na imagem do que nas legendas. Os surdos, por sua vez, apresentaram maior duração das fixações nas legendas do que na imagem. Nos relatos, os grupos apresentaram desconforto quanto ao conhecimento do conteúdo dos vídeos. Novamente, houve reclamações quanto à velocidade das legendas, sugerindo, mais uma vez, que a má segmentação é fator crucial para boa leitura e entendimento das legendas.

A condição RBS, por sua vez, apresentou duração maior das fixações pelos surdos quando comparados aos ouvintes. Quanto à compreensão, os dados apontam que ambos os grupos foram capazes de compreender vídeo e legenda muito bem. Os autores chegaram à conclusão de que a condição RBS é a melhor para surdos e ouvintes, já que facilita a recepção por ambos os grupos (MONTEIRO e DANTAS, 2017). Houve, no entanto, uma discrepância entre os dados quantitativos e qualitativos em relação aos surdos. Os dados qualitativos sugeriram boa recepção pelos surdos, mas os quantitativos sugeriram um resultado mais consolidado.

Para chegar a conclusões mais exatas sobre a condição RBS, a pesquisa de Souza (2020) desenvolveu essa análise estatística sobre os dados coletados por Monteiro (2016), a fim de se obter maior conhecimento dos dados quantitativos que não tiveram a atenção necessária anteriormente. Para tal análise de dados, Souza (2020) fez uso do Modelo Linear Misto (LMM) e o pacote *lmer*, usado para análise no programa R (BATTES, 2020). O LMM descreve a relação de uma variável resposta e covariáveis sendo observadas em conjunto, a partir do *software* gratuito RStudio.

Souza (2020) teve que limpar os dados para obter maior precisão estatística. Para tal, ela estabeleceu uma rotina de trabalho, a qual consistia em assistir a cada vídeo com a gravação do olhar durante os testes, observando o número e a duração das fixações durante a leitura das legendas por parte dos participantes.

Sobre a duração das fixações, a autora excluiu da análise todas que tinham duração menor que 80ms e maiores que 500ms. Para Souza, as durações mais curtas seriam apenas o espectador vagando seu olho pela tela e não propriamente lendo. As mais longas, por sua vez, podem indicar algum tipo de distração. Foram consideradas como variáveis de estudo os movimentos de deflexão, regressão e fixação.

Depois de analisar estatisticamente as condições propostas, Souza (2020) concluiu que, independentemente da velocidade e do grupo, a má segmentação leva a maiores durações das fixações. Isso implica que os participantes passaram mais tempo na leitura das legendas do que percebendo a imagem. Esta implicação pode significar que a percepção da obra audiovisual se torna prejudicada pela má segmentação.

Souza (2020) também apontou que a condição RBS é a mais adequada para ambos os grupos. As fixações nessa condição foram mais curtas e isso teve um impacto estatístico maior que as demais situações para os dois grupos. A velocidade, portanto, torna-se um fator fundamental para a boa compreensão do gênero propaganda política, mas não da maneira como se acreditava previamente. Antes, quanto mais lenta a legenda, melhor; mas se provou que quanto mais rápida e melhor segmentada, mais benefício para a compreensão.

Quanto a regressões, os dois grupos apresentaram mais movimentos regressivos nas legendas lentas do que nas rápidas, indicando possível desconforto. A RBS, no entanto, se mostrou mais confortável novamente para surdos e ouvintes. A segmentação nessa variável não se mostrou relevante, sendo a velocidade o principal fator para se chegar a uma menor quantidade de regressões (SOUZA, 2020). Para surdos, tanto a RBS quanto a RMS obtiveram menos regressões. Esses resultados foram diferentes dos encontrados em Vieira (2016), que observou como a má segmentação influencia diretamente a quantidade de regressões, abrindo, assim, mais espaço para pesquisas experimentais futuras.

A conclusão dessas pesquisas é que o padrão de velocidade mais adequado para surdos e ouvintes é o rápido, 180ppm, e que a boa segmentação proporciona menos esforço e mais conforto para que ambos os públicos desfrutem a obra audiovisual da melhor maneira possível (VIEIRA, 2016; MONTEIRO e DANTAS, 2017; SOUZA, 2020). O presente estudo,

portanto, se utiliza da velocidade de 180ppm e esteve atento à segmentação das suas legendas, para evitar que problemas técnicos influenciem negativamente os resultados.

2.7 Teoria das Emoções Básicas e Teoria da Universalidade das Expressões Faciais

Para termos um meio de medirmos quantitativamente as emoções dos participantes, trabalharemos segundo algumas teorias da psicologia sobre emoções e suas expressões. As expressões faciais são apenas uma maneira de comunicar-se não verbalmente. Segundo Vieira (2015), podemos manifestar nossas emoções, ainda, por gestos, entonação e postura corporal. As expressões faciais, no entanto, são as mais importantes no contexto intersocial, as quais nos permitem trocar, inferir e transmitir informações sobre as emoções com maior facilidade de recepção.

Para Vieira (2015), os estudos sobre expressões faciais ganharam força a partir dos estudos de Ekman e Friesen (1978), nos quais avaliaram a expressão e o reconhecimento das emoções em diferentes populações. Seus resultados corroboraram a teoria evolucionista das emoções de Darwin (2009), que afirma que primatas conseguem exprimir suas emoções por meio de suas expressões faciais.

Para averiguar como se manifestam as expressões faciais, antes, procuramos definir o termo emoção. Vieira (2015, p. 19) o define como:

Um processo de nível básico, que não pode ser completamente dissociado de outros processos cognitivos, que é compartilhado por diferentes espécies, possuindo importância e características evolutivas e que tem seu significado compartilhado através de mecanismos verbais, bem como, através de sinais não verbais de comunicação.

Compreendemos a emoção como um *continuum* de intensidades que podem variar de acordo com o indivíduo e o contexto (EKMAN, 1970). As emoções, portanto, não são estados individualizados e isolados uns dos outros, elas são inter-relacionadas e manifestam-se em conjunto. Apesar de termos uma emoção dominante em determinadas situações, as outras fazem-se presentes em menor grau e em auxílio à que está sendo exprimida (VIEIRA, 2015). As emoções básicas seriam, então, respostas construídas evolutivamente e com caráter universal, podendo ser reconhecidas por toda uma espécie. Ou seja, todos os seres humanos são capazes de reconhecer as emoções básicas de outros, não importando o local de nascimento ou cultura (VIEIRA, 2015). Essas emoções básicas teriam sido essenciais para a sobrevivência e evolução da espécie.

Uma emoção básica, para ser classificada como tal, deve seguir alguns parâmetros comuns. Ela deve apresentar sinais universais com respostas fisiológicas diferentes. Essas emoções são compartilhadas por outras espécies, principalmente, os primatas. Elas ocorrem espontaneamente e apresentam um início de reação muito rápido, podendo ser de duração muito curta ou não. Essas emoções ainda apresentam pensamentos e memórias distintas associadas a cada reação (EKMAN e CORDARO, 2011). As emoções básicas não podem ser aprendidas por meio da cultura ou do contexto social do indivíduo, já que ela é uma habilidade proveniente do ciclo evolutivo da espécie (VIEIRA, 2015).

As pesquisas de Ekman e Friesen (1978) investigaram a função dos sinais não verbais dentre várias culturas e atribuíram às expressões faciais o cargo de transmissor universal das emoções. Esses sinais são importantes para o ser humano interagir entre si e reconhecer perigos ou segurança entre seus pares. A hipótese sobre a importância das expressões faciais já havia sido postulada anteriormente por Darwin, envolvendo, ainda, a entonação e a postura corporal como transmissor de informações sobre as emoções, mas até estudos mais recentes ela ainda não poderia ser confirmada (DARWIN, 2009; VIEIRA, 2015).

Dentre as emoções tidas como básicas, temos: alegria, medo, nojo, raiva, surpresa, tristeza e, posteriormente, o neutro (DUTTA e BARMAN, 2020). Elas são categorizadas da seguinte maneira, de acordo com Dutta e Barman (2020):

- Alegria – é tida como um estado de espírito que pode refletir satisfação, prazer ou felicidade. A alegria é uma das emoções mais comuns de se notar. Por vezes, a associamos com prazer. Ela é caracterizada por uma expressão facial que faz com que alguém levante os cantos da boca;
- Medo – provém de estímulos de perigo ou ameaça. É um mecanismo básico de sobrevivência que ocorre em resposta a uma presença traumática, como dor ou a ameaça iminente de dor. O medo faz a pessoa levantar as sobrancelhas, abrir ligeiramente a boca e arregalar os olhos de forma mais ampla do que o normal;
- Nojo – provém de coisas ofensivas, intragáveis ou infecciosas. Ele é relativo a algo repulsivo e é experimentado, principalmente, em retaliação ao sentido do paladar e, secundariamente, em retaliação a qualquer coisa que provoque uma sensação semelhante por meio dos sentidos do olfato e do tato. O nojo, no entanto, é uma das seis emoções que diminuem a

frequência cardíaca. Ele é caracterizado por uma expressão facial que faz com que a pessoa levante o lábio superior, enrugue a ponte do nariz e erga as bochechas;

- Raiva – está presente na gama que vai de uma pequena irritação à cólera intensa. Quando o indivíduo opta por agir contra qualquer ameaça que esteja causando trauma ou dor, a raiva é considerada a resposta, emoção dominante, e que é expressa de uma forma cognitiva e fisiológica. Fisicamente, a raiva faz com que alguém experimente um aumento da frequência cardíaca, pressão arterial e níveis anormais de adrenalina e noradrenalina. A raiva é caracterizada por uma expressão facial que faz com que alguém abaixe as sobrancelhas, pressione os lábios com firmeza e arregale os olhos;
- Surpresa – é incorporada a um breve estado de ser, sendo invocado por um evento relevante inesperado. No entanto, a surpresa nem sempre é uma experiência emocional traumática. A surpresa também foi categorizada no espectro de ser neutra, agradável e desagradável. Ela é caracterizada por uma expressão facial que faz com que alguém arqueie as sobrancelhas, abra bem os olhos e abaixe o queixo;
- Tristeza – faz-se presente nos sentimentos de desvantagem, perda e desamparo. O indivíduo reage à tristeza, frequentemente, calando-se, sentindo falta de energia e necessidade de se isolar. A tristeza é definida como o oposto direto da felicidade e pode ser caracterizada como dor, sofrimento, miséria e melancolia. A tristeza é caracterizada por uma expressão facial que faz com que alguém abaixe os cantos da boca e erga a parte interna das sobrancelhas.

Ainda há o neutro que, segundo Ekman (2003), seria a ausência de expressões faciais. É dela que, geralmente, a face parte para expressar alguma emoção (EKMAN, 2003).

Segundo Dutta e Barman (2020), as emoções são percebidas através das contrações e relaxamentos dos músculos da face humana. Ressaltamos que nem todo movimento facial pode ser caracterizado como decorrente de uma emoção sentida pelo indivíduo ou uma emoção básica. Assim sendo, é possível que uma máquina treinada pode se tornar capaz de analisar e medir as expressões faciais humanas. Vieira (2015) afirma que é de

relativa facilidade manipular o reconhecimento das expressões e as pesquisas testam isso cada vez mais de maneira automática.

Esta pesquisa utiliza-se de um programa treinado no reconhecimento e na mensuração de expressões faciais. O programa em questão é o *Visage*, que foi desenvolvido por este pesquisador, para que as expressões dos participantes sejam medidas quadro a quadro. Frisamos que, como nenhuma emoção é sentida de maneira integral, o programa mensura pequenas variações e todas as emoções ao mesmo tempo.

No próximo capítulo, falaremos sobre os passos metodológicos para o desenvolvimento deste estudo, desde o perfil dos participantes ao desenho experimental e à manipulação dos estímulos.

3 METODOLOGIA

Esta seção dedica-se à apresentação dos passos metodológicos deste estudo. A pesquisa tem por propósito central analisar o impacto da tradução dos efeitos sonoros na recepção fílmica pelo público surdo por meio da análise das emoções provocadas por essas traduções. A base teórico-metodológica vem da Tradução Audiovisual Acessível (TAVa), mais especificamente, da Legendagem para Surdos e Ensurdecidos (LSE). Nesta seção, também está presente a descrição do tipo de pesquisa no qual este estudo se enquadra, bem como o contexto no qual a pesquisa foi realizada, a elaboração e veiculação dos estímulos, bem como a manipulação das legendas desenvolvida pelos pesquisadores. Apresentaremos também o desenho experimental da pesquisa e os procedimentos tomados para se coletar dados em meio à pandemia de COVID-19.

3.1 Tipo de pesquisa

É considerada uma pesquisa experimental, uma vez que testamos a influência da tradução dos efeitos sonoros na emoção dos espectadores surdos, tendo como ferramenta um *software* que capta dados das emoções expressas pelos participantes em suas reações aos estímulos apresentados. Dessa forma, este estudo se debruça em dados quantitativos.

Os procedimentos quantitativos que se fazem presentes na aplicação do *software Visage* são considerados exploratórios. Inicialmente iríamos utilizar o poder de processamento da API⁶ de detecção facial da *Microsoft Azure*, mas, devido ao seu alto custo monetário para o processamento de dados, optamos por desenvolver nós mesmos uma aplicação semelhante ao provido pelo serviço da *Microsoft*, mas de maneira gratuita. Consideramos relevante mencionar que o *software*, intitulado *Visage*, desenvolvido por este pesquisador especificamente para esta pesquisa, tem a finalidade de captar dados das emoções expressas pelos participantes em suas reações aos estímulos apresentados. O *Visage*, bem como os demais elementos metodológicos deste estudo, será descrito em maiores detalhes nas subseções a seguir.

⁶ *Application Programming Interface*, em português, Interface de Programação de Aplicações, é um conjunto de padrões providos por uma empresa, que permite a criação de aplicações por desenvolvedores de maneira mais fácil e prática.

3.2 Contexto da pesquisa

Esta dissertação faz parte do programa de Pós-Graduação em Estudos da Tradução (POET), da Universidade Federal do Ceará (UFC). A pesquisa foi realizada no Laboratório de Psicolinguística e Ciências Cognitivas, atualmente coordenado pela Dra. Elisângela Nogueira Teixeira, da UFC, em parceria com o Laboratório de Tradução Audiovisual (LATAV), da Universidade Estadual do Ceará (UECE), coordenado pela Dra. Silvia Malena Modesto Monteiro.

3.3 Estímulos

Os estímulos para o estudo são formados por quatro trechos de um filme de ação estrangeiro: *Dunkirk* (2017). *Dunkirk* é uma obra ficcional baseada em fatos, que busca representar a Operação Dínamo, ocorrida em 1940. Essa operação ocorreu durante a 2ª Guerra Mundial e tinha por objetivo evacuar as tropas expedicionárias inglesas do território da França, visto que elas estavam encurraladas pelas forças armadas da Alemanha nazista na cidade de Dunquerque. O diretor do filme, Christopher Nolan, apresenta a operação por três pontos de vista diferentes: i) o soldado encurralado na cidade; ii) o cidadão britânico que se voluntaria no resgate das tropas por meio de seu barco; e iii) o avião que provê suporte aéreo para as tropas britânicas. Faremos uso, aqui, de dois desses arcos narrativos: o ponto de vista do soldado e o do avião.

Esse filme foi escolhido por apresentar diversas cenas de ação dramática, nas quais os personagens são construídos com a finalidade de vivenciar o drama da história (COMPARATO, 2018). A forte dependência do fator sonoro para a imersão e compreensão da ação pelo espectador, assim como o forte apelo emocional, pelo fato de o filme ter arcos dramáticos de grande intensidade, também justificam a escolha dessa obra como *corpus* do estudo. Um último fator que impacta diretamente na escolha é a baixa bilheteria e relativo desconhecimento do público geral, diminuindo, assim, as chances dos participantes já o terem assistido antes e isso influenciar os resultados. No entanto, para garantir resultados mais confiáveis sobre a influência, participaram apenas surdos que não assistiram ao filme.

O filme foi adquirido por meio da plataforma digital e *online* de *streaming* *Netflix*⁷. Por meio de uma assinatura do serviço, foi realizado seu *download*, a fim de se ter um arquivo disponível de maneira *offline*. As cenas do filme foram captadas por meio de um *software* de captura de tela chamado *OBS Studio*⁸ a fim de se possuir arquivos editáveis e manipuláveis. As cenas foram capturadas para passarem pelo processo de edição depois, visto que a plataforma não permite a manipulação do arquivo baixado. O programa *OBS Studio* é gratuito e permite fazer gravações de tela, sons emitidos pelo computador, bem como de um microfone. Apesar de seu objetivo ser o *streaming* da tela do usuário, ele funciona de maneira satisfatória para capturar a tela do computador.

Diante disso, e conforme mencionado anteriormente, os participantes surdos tiveram acesso a quatro trechos distintos do filme, com duração de 2 a 3 minutos cada, sendo o suficiente para conter uma cena completa do filme (começo, meio e fim) e manter a narrativa fílmica contextualizada. 16 participantes foram expostos a dois trechos, cada um com uma condição distinta. Todos os trechos foram legendados por este pesquisador, seguindo o padrão preconizado pelo grupo LEAD sobre a LSE.

Os trechos foram legendados em português de duas formas: com e sem a tradução dos efeitos sonoros, seguindo a orientação proposta por Nascimento (2018) sobre a tradução dos efeitos sonoros na LSE.

Os quatro trechos selecionados e editados têm a seguinte duração:

- Trecho 1 – 2 minutos e 20 segundos;
- Trecho 2 – 2 minutos;
- Trecho 3 – 2 minutos e 42 segundos;
- Trecho 4 – 2 minutos e 49 segundos;

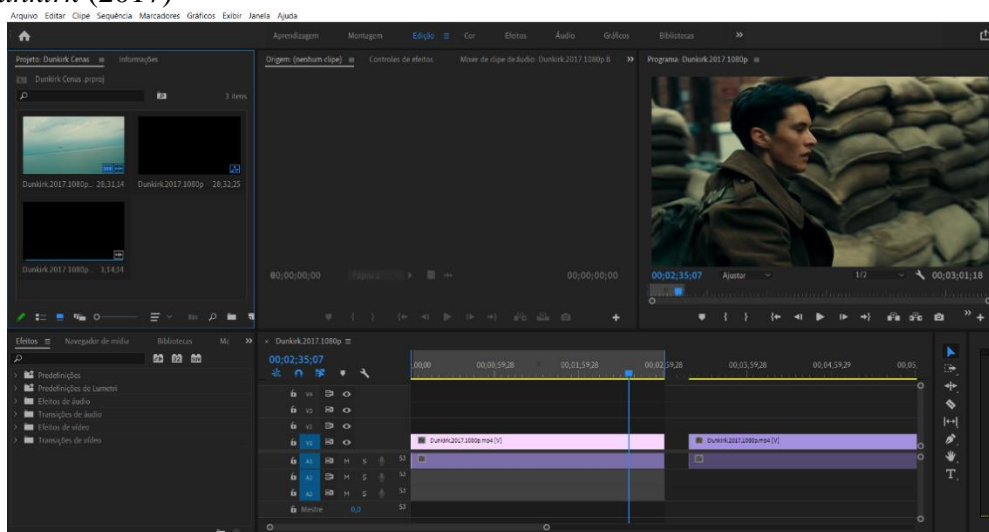
Os trechos foram cortados e editados do filme inteiro por meio do programa de edição *Adobe Premiere Pro*⁹ (FIGURA 2) a partir de uma licença pessoal dos *softwares* da plataforma Adobe. Esse programa permite demarcar, separar e extrair trechos da obra fílmica e renderizá-los em novos extratos separados, a partir de uma licença pessoal dos *softwares* da plataforma Adobe. É necessária, também, razoável experiência prévia em edição para usufruir das mais diversas funcionalidades do *Adobe Premiere Pro* em sua totalidade.

⁷ <https://www.netflix.com/br/title/80170278>

⁸ <https://obsproject.com/pt-br/download>

⁹ <https://www.adobe.com/br/products/premiere.html>

Figura 2 – Interface do programa de edição Adobe Premiere Pro com cena do filme *Dunkirk* (2017)



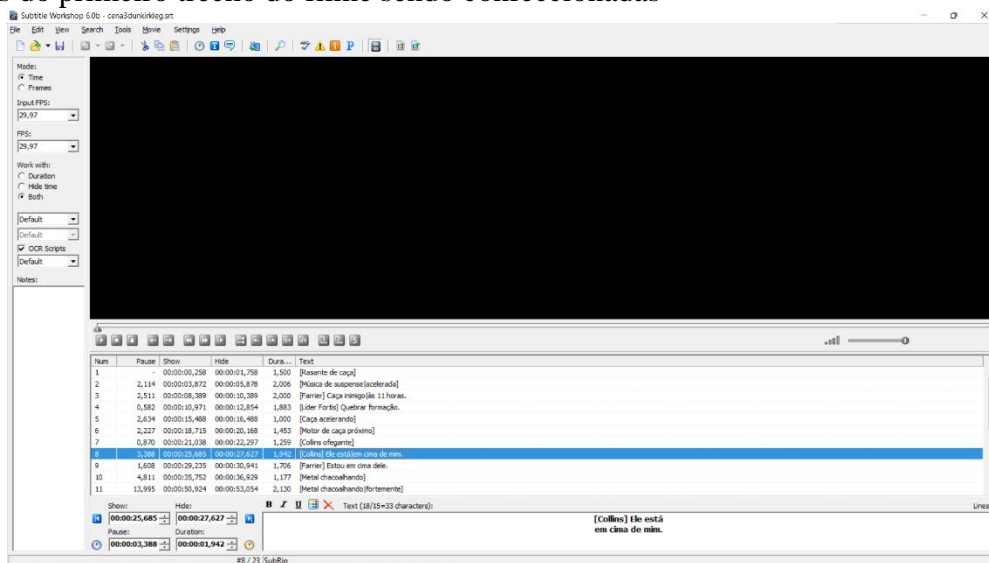
Fonte: Elaborado pelo autor.

As legendas para os dois grupos foram confeccionadas no *software Subtitle Workshop 6.0b*¹⁰ (FIGURA 3). Esse programa já se mostrou funcional em toda a trajetória do grupo LEAD, bem como nas pesquisas de Vieira (2016) e Assis (2016). O *software* permite, em suas funcionalidades, marcar o tempo de entrada e saída das legendas, sincronizá-las com o áudio, marcar o número de caracteres por segundo em cada legenda, além de possibilitar revisões sobre questões ortográficas e espaços vazios.

O *software* provê, também, as funcionalidades de criação e edição de legendas para qualquer arquivo de vídeo, visto que a extensão base de uma legenda é o .txt, um arquivo simples de texto. Outras funcionalidades que o *Subtitle* provê são recursos importantes para acelerar e facilitar o trabalho do legendistas, como: ajuste de *timing* em todas as legendas ao mesmo tempo, possibilidade de alteração da velocidade, posição das legendas do projeto de maneira prática, realce de erros na quantidade de caracteres e duração. Por ser um programa de fácil aprendizado e de livre acesso, sua popularização ajudou a difundir o trabalho de legendagem entre fãs de produções audiovisuais (ASSIS, 2016). O programa se mostra, portanto, adequado para uma boa execução da tarefa de legendagem desta pesquisa.

¹⁰ <http://subworkshop.sourceforge.net/>

Figura 3 – Interface do programa de legendagem *Subtitle Workshop 6.0b* com legendas do primeiro trecho do filme sendo confeccionadas



Fonte: Elaborado pelo autor.

O primeiro trecho selecionado tem 2 minutos e 20 segundos de duração e compreende três cenas não contínuas¹¹, caracterizando-se, assim, uma pequena montagem. Essa montagem foi executada para poupar tempo da pesquisa com os estímulos, além de retirar cenas que não fizessem sentido narrativo no extrato selecionado. Essa cena apresenta dois soldados se passando por médicos para salvar um ferido em uma maca, mas sendo o seu real objetivo embarcar em um navio-hospital para fugir dali. Ao entregar o ferido, eles são expulsos do navio, por não serem da brigada hospitalar. A cena como um todo é uma corrida contra o tempo para seus protagonistas.

Esse trecho foi escolhido por apresentar pouco diálogo, 11 legendas, muita ação em tela e 16 legendas para efeitos sonoros, sendo o som fator essencial para a compreensão fílmica. Por ser uma espécie de corrida contra o tempo, a angústia pode ser uma das emoções suscitadas com o trecho, podendo gerar algumas das emoções básicas universais utilizadas nesta pesquisa, como raiva e surpresa.

A trilha sonora *off-screen* desta cena é composta por um instrumento de cordas tocando notas curtas e agudas, cada vez mais rápidas e mais altas. A trilha acompanha o ritmo da ação em tela que se acelera e ruma em direção a um clímax dramático conforme os sons se

¹¹ Uma edição de cenas não contínuas ocorre quando se cria uma cena a partir de trechos em diferentes momentos cronológicos de um filme, tornando-se, assim, uma montagem.

intensificam e progridem. O som aqui, portanto, potencializa as diversas emoções que o espectador pode vir a sentir. O senso de urgência no núcleo narrativo se torna maior pelo som que o acompanha, podendo, portanto, provocar maiores e mais evidentes reações emocionais do espectador.

O primeiro trecho se torna, então, um elemento essencial para a avaliação do impacto sonoro e sua tradução para o público surdo, justamente por apresentar alta dependência de sua trilha para entregar o impacto dramático desejado pelo diretor.

O segundo extrato tem duração de 2 minutos e é uma montagem de três cenas não contínuas, assim como o primeiro. Ele apresenta uma luta, a chamada *dogfighting* pelos ingleses, entre dois caças britânicos e dois alemães. A cena acontece majoritariamente dentro da cabine de um piloto inglês caçando o inimigo no ar. A montagem e posição das câmeras causam um efeito de clausura e angústia, por não conseguir acertar o oponente. Ao final do extrato, o britânico abate um caça alemão, mas recebe o contra-ataque do remanescente.

Este trecho apresenta poucas linhas de diálogo, novamente 11 legendas e 9 efeitos sonoros. Sua trilha sonora é frenética e entrega variados efeitos sonoros, de tiros a voos em alta velocidade. Esperamos, com esse extrato, dados semelhantes aos do primeiro, visto que ambos investem sua carga dramática em sentimentos associados à angústia e ansiedade. Por ser mais curto, o efeito sentimental pode ser potencializado.

O terceiro trecho apresenta duração de 2 minutos e 42 segundos. Ele é novamente uma montagem de cinco trechos não contínuos. Nesta cena, um dos pilotos britânicos é abatido e cai no mar. Durante 2 minutos, ele tenta sair de sua aeronave enquanto ela afunda e vai matando-o afogado. Ele é salvo ao final da cena por uma família que se dirigia num barco à praia de Dunquerque para ajudar no esforço de salvação das tropas expedicionárias.

Este extrato apresenta mais um arco dramático do que propriamente a ação de um filme de guerra. Ainda assim, ele pode ser caracterizado, também, como uma corrida contra o tempo, entregando toda a ação necessária para evocar no espectador apreensão pela vida do personagem em cena. A carga sonora, apresentando 20 legendas para tradução de efeitos sonoros e somente 2 de diálogo, corrobora com o tom de urgência e emergência da situação. Ela é ressaltada pela trilha *off-screen*, marcada por batidas pesadas escalando a situação e os efeitos sonoros de água enchendo a cabine e a respiração ofegante do piloto.

O último extrato apresenta duração de 2 minutos e 49 segundos, sendo o único trecho que não apresenta montagem. Nesta cena, os soldados resgatados da praia estão em um navio de guerra que é atacado à noite, por um submarino com vários torpedos. Ao ser

atingida, a embarcação começa a afundar e os tripulantes se desesperam para não morrerem afogados. Novamente, temos uma corrida contra o tempo, mas com a angústia de acompanhar os soldados se desesperando com a morte por afogamento iminente. O trecho apresenta muitos sons produzidos por humanos, principalmente de gritos, sendo ao total 19 legendas destinadas a este fim, e somente uma para diálogos. Há, também, a presença de uma trilha sonora angustiante, que dá ritmo compassado à tentativa de escape dos tripulantes.

Todos os extratos escolhidos têm curta duração, menos de 3 minutos. Essa duração pequena e homogênea é proposital e busca ser uma métrica confiável, por causa da sua uniformidade temporal. Essa métrica tem a finalidade de não cansar o espectador, evitando, assim, que ele perca sua atenção ou se distraia do vídeo. A prevenção da distração nos proporcionará dados mais consistentes entre todos os participantes da pesquisa.

Quase todos os trechos também apresentam cenas de ação ininterrupta, forte apelo visual e grande dependência sonora para corroborar com o drama em tela. Por apresentarem arcos narrativos e emoções suscitadas semelhantes em todos os extratos, esperamos ter resultados homogêneos nas aplicações, sendo as discrepâncias causadas pelo efeito da tradução de efeitos sonoros.

3.3.1 Manipulação da LSE

Nascimento (2018) estabelece um padrão na tradução dos efeitos sonoros, para que esse parâmetro possibilite emoção a quem não ouve, da mesma forma que aos ouvintes ao ouvirem os sons nas produções audiovisuais. Dessa forma, utilizamos as convenções estabelecidas por ela em língua portuguesa.

A manipulação das legendas com a tradução dos efeitos sonoros foi realizada a partir da convencionalidade sobre os efeitos sonoros de Nascimento (2018), no que tange a essa tradução intersemiótica. A LSE seguiu também os padrões técnicos preconizados pelos pesquisadores do grupo LEAD (ARAÚJO, 2004, 2008). A velocidade das legendas foi de 180ppm, uma vez que os resultados dos estudos experimentais realizados com a técnica do rastreamento ocular por Vieira e Araújo (2017) e Souza (2020) apontam que é nessa velocidade que os surdos têm uma melhor recepção. A cor das legendas foi amarela com borda na cor preta (ARAÚJO, 2012). Também foram divididas, quando necessárias, seguindo os critérios de segmentação visual, corte de cena, retórica, fala dos personagens, linguística e sintaxe, seguindo o proposto pelas pesquisas de Araújo (2012). Todas as legendas seguem o

padrão de exposição mínima de 1 segundo e máxima de 4 segundos, para evitar releituras (VIEIRA e ARAÚJO, 2017; VIEIRA et al, 2017; MONTEIRO e DANTAS 2017; SOUZA 2020; VIEIRA et al, 2020).

Para os testes, os vídeos foram legendados em duas variáveis: LSE sem tradução dos efeitos sonoros e LSE com tradução dos efeitos sonoros. Para esquematizar e organizar a produção das legendas, separamos as mesmas em dois blocos de trabalho: LSE-CSOM e LSE-SSOM. O bloco LSE-CSOM contém legendas com tradução dos efeitos sonoros na LSE, uma para cada um dos quatro trechos selecionados. O bloco LSE-SSOM também contém legendas para os quatro trechos em LSE, mas, por sua vez, não apresenta tradução de efeitos sonoros. O bloco LSE-CSOM teve criação a partir das legendas pertencentes à Netflix, mas destacamos que não tivemos acesso direto às legendas da plataforma. Fizemos uso delas para compreendermos melhor alguns trechos de diálogo ininteligíveis, a fim de fazermos uma tradução própria.

O primeiro bloco, LSE-CSOM, destina-se à LSE com tradução de efeitos sonoros. Este bloco de legendas, contendo os quatro trechos selecionados, foi feito com base no áudio e no texto das legendas em inglês providos pelo serviço Netflix. Houve falas dos personagens no idioma de partida, o inglês, que não conseguimos compreender e identificar. Para sanar esses impasses na tradução, recorreremos ao texto-fonte da legenda em inglês. Uma vez compreendida a fala e seu contexto, mediante a legenda-fonte, fizemos a tradução e confeccionamos as legendas.

O segundo bloco, LSE-SSOM, apresentou legendas sem a tradução dos efeitos sonoros. Para esse estímulo, utilizamos a mesma LSE-CSOM, retirando apenas a tradução dos sons.

As duas manipulações, LSE-CSOM e LSE-SSOM, foram realizadas nos quatro trechos do filme. A LSE-CSOM foi a base da LSE-SSOM, no entanto, tivemos que fazer ajustes para manter a velocidade de 180 ppm em todas as legendas dos estímulos. O Quadro 2 apresenta as diferenças entre as legendas de um mesmo trecho. A diferença entre os blocos é a ausência completa das legendas que contém tradução dos sons diegéticos. As legendas 5 e 9 do bloco LSE-CSOM não existem no bloco LSE-SSOM, seguindo a manipulação dos vídeos.

Quadro 2 – Comparativo entre os blocos de legenda com e sem tradução de efeitos sonoros

COMPARATIVO ENTRE LSE-CSOM E LSE-SSOM	
LSE-CSOM – TRECHO 1	LSE-SSOM – TRECHO 1
5 00:00:10,823 --> 00:00:13,088 [música de tensão]	SEM LEGENDA
6 00:00:24,039 --> 00:00:25,235 [Capitão] É o último?	6 00:00:24,039 --> 00:00:25,235 [Capitão] É o último?
7 00:00:25,335 --> 00:00:26,572 [Médico] Sim, senhor.	7 00:00:25,335 --> 00:00:26,572 [Médico] Sim, senhor.
8 00:00:29,486 --> 00:00:30,486 Soltem as cordas.	8 00:00:29,486 --> 00:00:30,486 Soltem as cordas.
9 00:00:31,450 --> 00:00:32,450 [bomba]	SEM LEGENDA

Fonte: Elaborado pelo autor

Finalizado o processo de manipulação das legendas, os vídeos passaram pela análise de um consultor surdo do grupo LEAD, para verificar se não havia nenhum problema ou incompreensão que pudesse prejudicar os testes. O consultor indicou duas alterações. A primeira foi iniciar toda tradução de efeito sonoro com letras minúsculas, contudo, indicadores de falantes deveriam permanecer com inicial em maiúsculas, como pode ser visto nas legendas 6 e 7 da coluna LSE-CSOM do Quadro 2. Outra sugestão foi a eliminação de traduções dos efeitos sonoros que já fossem percebidas pela imagem, evitando leituras desnecessárias.

3.4 Instrumento qualitativo

Aplicamos dois questionários distintos, o primeiro com a finalidade de traçar o perfil demográfico dos participantes (APÊNDICE A) e o segundo que buscou questionar os sentimentos dos participantes ante os estímulos (QUADRO 3). Ambos os questionários foram

aplicados de maneira *online* e por meio da plataforma Google Forms. O segundo questionário foi aplicado, mas não foi analisado nesta pesquisa, por não termos tempo hábil para executar essa tarefa.

O primeiro questionário para as entrevistas tem como base o utilizado por Vieira (2016), apenas sendo modificado para se adequar ao consumo de mídia atual – serviços de *streaming* –, além de adicionar a escala Likert, quando necessário, para oferecer informações menos subjetivas aos participantes.

O questionário demográfico compreende 20 questões, tendo perguntas sobre dados mais gerais, como idade, gênero e grau de instrução; e mais específicos, como frequência em que o entrevistado assiste a filmes e afinidade com a LSE (APÊNDICE A). Nenhum nome foi utilizado, para não comprometer a identidade dos participantes. Para identificá-los, entretanto, usaremos a letra “S”, seguida de um número cardinal referente à ordem de participação.

O segundo questionário (QUADRO 3) teve por finalidade levantar dados sobre os sentimentos experimentados pelos participantes durante a fase de estímulos. Para tal, eles responderam a um questionário de autopercepção de suas emoções. O questionário foi aplicado duas vezes para cada participante. Uma vez para cada trecho assistido e logo após o estímulo. A finalidade dele, no âmbito desta pesquisa, é somente colher dados para um futuro trabalho de doutorado, não sendo analisados na presente pesquisa.

A base para a confecção desse questionário é a Escala Diferencial das Emoções (DES - *Differential Emotions Scale*) de Izard (1977). Essa gradação, elaborada por Izard, traz 10 emoções básicas que são comuns a todos os humanos, as quais são chamadas de fatores. Para a aplicação com as populações, cada um dos fatores apresenta três adjetivos correspondentes à referida emoção, chamado pela pesquisadora de item, para guiar e facilitar o reconhecimento pelo pesquisado do que ele estava sentindo (IZARD, 1977). Como esta pesquisa faz uso das seis emoções básicas universais (alegria, tristeza, raiva, medo, nojo e surpresa), mais o neutro, quatro das emoções propostas no DES foram retiradas da escala que será aplicada nesta pesquisa.

Quadro 3 – Questionário com Escala Diferencial das Emoções de Izard (1977) adaptado para a pesquisa

Pergunta: Como você se autoavalia para cada uma das emoções listadas abaixo ao assistir ao primeiro trecho do filme?					
Emoção/Escala	1 Não senti	2 Senti pouco	3 Senti razoável	4 Senti muito	5 Senti a todo momento
Alegria (encantado, feliz, contente)					
Tristeza (desanimado, triste, desencorajado)					
Raiva (enfurecido, zangado, furioso)					
Medo (assustado, temeroso, apreensivo)					
Nojo (aversão, enojado, repulsa)					
Surpresa (surpreso, maravilhado, admirado)					

Fonte: Izard, Carroll E. (1977). Adaptado pelo autor.

Conforme informamos anteriormente, os dados do questionário DES não serão utilizados na análise desta dissertação, incluímos neste capítulo apenas como registro de uma das etapas que os participantes vivenciaram neste estudo.

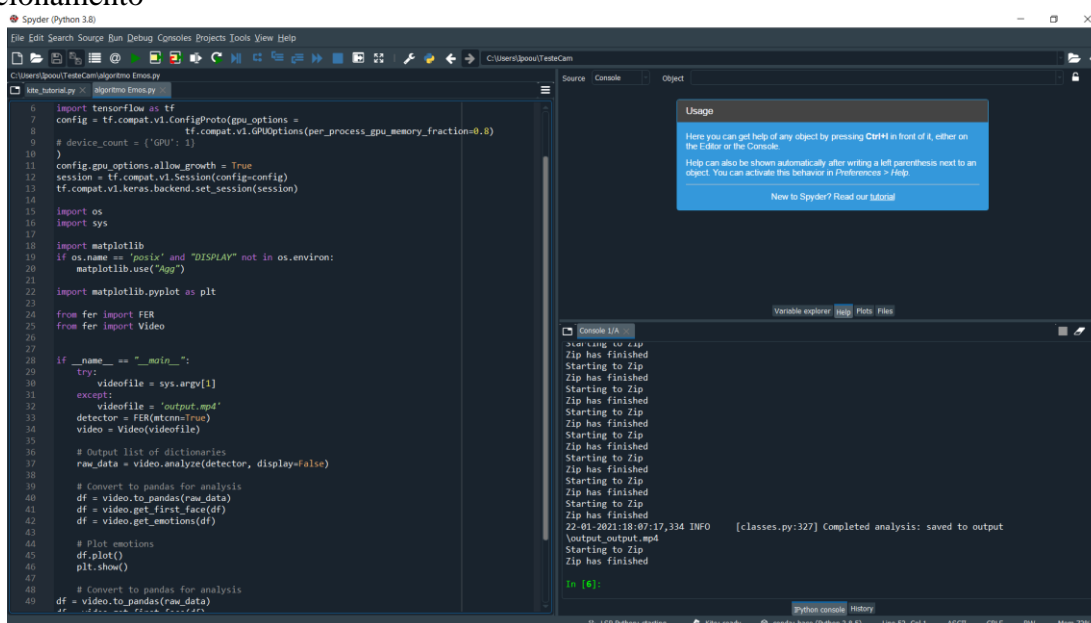
3.5 Instrumento quantitativo

3.5.1 O programa desenvolvido - Visage

A aplicação, nomeada pelo pesquisador de *Visage*, foi desenvolvida em *Python* em sua versão 3.8. *Python* é uma linguagem de programação de alto nível, desenvolvida na década de 1990, e que tem se tornado uma das linguagens mais utilizadas na atualidade no desenvolvimento de aplicações diversas (GITHUB, 2020). Uma linguagem é considerada de alto nível quando ela se aproxima mais da escrita e compreensão do ser humano do que da sintaxe da máquina (DEITEL e DEITEL, 2019). *Python* também é uma linguagem multiparadigma, o que confere a ela grande versatilidade de possibilidades de desenvolvimento em diversos contextos (DEITEL e DEITEL, 2019).

O algoritmo desenvolvido foi escrito e executado na IDE *Spyder* (FIGURA 4). Uma IDE (*Integrated Development Environment*) é, em português, conhecida por Ambiente Integrado de Desenvolvimento. Na criação de algoritmos e, por consequência, aplicações e programas, ela é uma facilitadora para a programação. Por conter diversos recursos e suporte integrado para a atividade de codificação, como correção automática de sintaxe e compilador incorporado em sua interface, ela se torna fundamental para o bom desenvolvimento das aplicações.

Figura 4 – Interface da IDE *Spyder* com parte do algoritmo escrito e em funcionamento



```

Spyder (Python 3.8)
File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\bonni\TextCami\algoritmo Emes.py
line_editor.py x algoritmo Emes.py

6 import tensorflow as tf
7 config = tf.compat.v1.ConfigProto(gpu_options =
8     tf.compat.v1.GPUOptions(per_process_gpu_memory_fraction=0.8)
9 )
10 # device_count = ('GPU': 1)
11 config.gpu_options.allow_growth = True
12 session = tf.compat.v1.Session(config=config)
13 tf.compat.v1.keras.backend.set_session(session)
14
15 import os
16 import sys
17
18 import matplotlib
19 if os.name == 'posix' and "DISPLAY" not in os.environ:
20     matplotlib.use("Agg")
21
22 import matplotlib.pyplot as plt
23
24 from fer import FER
25 from fer import Video
26
27
28 if __name__ == "__main__":
29     try:
30         videofile = sys.argv[1]
31     except:
32         videofile = 'output.mp4'
33     detector = FER(mtcnn=True)
34     video = Video(videofile)
35
36 # Output list of dictionaries
37 raw_data = video.analyze(detector, display=False)
38
39 # Convert to pandas for analysis
40 df = video.to_pandas(raw_data)
41 df = video.get_first_faces(df)
42 df = video.get_emotions(df)
43
44 # Plot emotions
45 df.plot()
46 plt.show()
47
48 # Convert to pandas for analysis
49 df = video.to_pandas(raw_data)
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

Fonte: Elaborado pelo autor

A fim de criar o programa com as características desejadas – reconhecimento facial e análise de suas expressões – e que possa ser executado em qualquer máquina, ou seja, gerar um executável, o programa fez uso de algumas bibliotecas encontradas no repositório *PyPi*¹². Foram utilizadas cinco bibliotecas diferentes, para prover funções básicas e avançadas do programa.

Bibliotecas, na programação, são conjuntos de diretrizes que executam uma determinada tarefa. Elas são fundamentais para desenvolver um programa aplicado e acelerar o processo de desenvolvimento, já que ela provê partes mais fundamentais já prontas,

¹² <https://pypi.org/>

restando ao programador desenvolver seu código em cima desses conjuntos pré-programados. A analogia com uma biblioteca é válida para compreender seu funcionamento. Quando alguém adentra o recinto, há vários livros à sua disposição para ler e estudar. A partir do conhecimento ali contido, ela pode construir novos saberes, baseado no que foi aprendido.

*Tensorflow*¹³ é uma biblioteca voltada para cientistas de dados, que reúne diversas outras bibliotecas específicas, para que o pesquisador tenha menos trabalho ao montar seu programa. No nosso programa, dentre funções mais gerais de *machine learning* e inteligência artificial utilizados no *Visage*, a biblioteca *Tensorflow* permite fazer uso da placa gráfica e acelerar o processamento dos dados imagéticos. *Os* e *sys* são bibliotecas voltadas para, respectivamente, a conversação com o sistema operacional e o processamento das linhas de código, de entrada e saída, sendo ambas vitais para a interação com a máquina e para o funcionamento da aplicação como um todo. *Matplotlib* é a biblioteca que permite que sejam plotados gráficos com os dados numéricos gerados a partir da interpretação das emoções. Fer (GOODFELLOW, *et al.*, 2013) é o centro nevrálgico do *Visage*, já que contém os códigos necessários para o programa executar o reconhecimento facial, a geração de uma área delimitadora das faces e, a posteriori, executar a interpretação das expressões em sentimentos. Dentre eles, a biblioteca *Fer* faz uso da arquitetura MTCNN (ZHANG, *et al.*, 2016) para o rastreio facial, tendo sua acurácia e funcionalidade validadas em estudos da área de computação e *machine learning* (ZHANG, *et al.*, 2016). Há outras arquiteturas com a mesma finalidade, como a *DeepFace*, mas que não foram utilizadas, pois a sua implementação só foi testada depois que começamos a utilizar o *Visage* em escala.

O funcionamento do programa segue alguns passos que serão aqui descritos. Antes de se iniciar o programa, precisa-se de um *input*, ou seja, uma entrada, que pode ser tanto uma imagem estática como um vídeo. Para esta pesquisa, em específico, as entradas foram constituídas exclusivamente de vídeos. Então, grava-se um outro vídeo, com a face do indivíduo reagindo ao estímulo, devendo ele estar devidamente enquadrado e iluminado. O vídeo com a expressão facial é utilizado para alimentar o programa e ele inicia o processamento dos dados. O processamento de dados apresenta algumas subetapas:

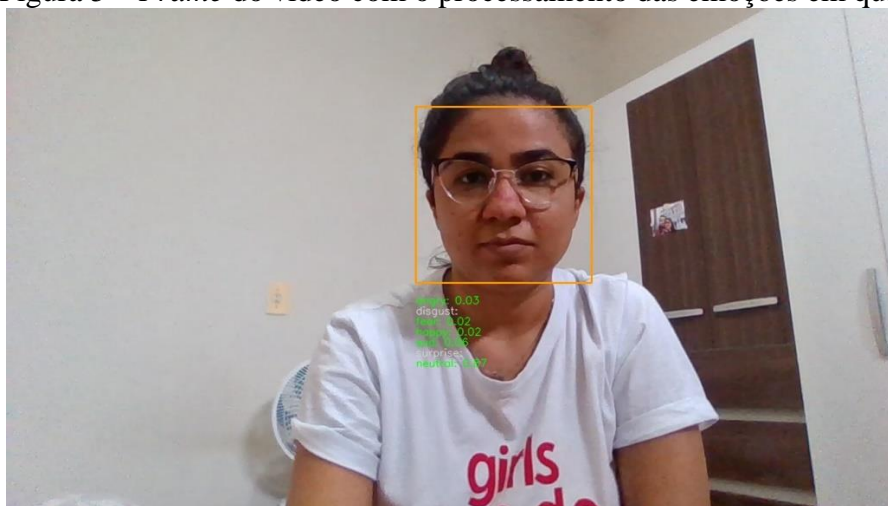
1. Quebra do vídeo em *frames*, quadros;
2. Pesquisa da face no quadro;

¹³ <https://www.tensorflow.org/>

3. Delimitação e *tracking* da face por meio de uma representação visual de um quadrado de cor laranja;
4. Interpretação da expressão facial e escrita na tela de cada uma das seis emoções básicas universais, com valores de 0 a 1.0 no Índice de Intensidade da Emoção (FIGURA 5);
5. Finalização da análise de todos os quadros;
6. Plotagem de gráfico com todas as emoções, em inglês, numa escala de intensidade por tempo (FIGURA 6);
7. Exportação de uma planilha com os valores adquiridos em cada um dos quadros (FIGURA 7);
8. Finalização do programa.

Este processo todo pode ser demorado, baseado na duração do clipe que foi alimentado no programa. Ele ocorre de forma automática, sem o usuário ver o que está sendo feito pelo programa. Ao final do processamento, é aberta uma pasta com o vídeo decupado em imagens separadas, quadros do vídeo com a face do participante devidamente marcada (FIGURA 5), além do gráfico com a frequência das emoções (FIGURA 6) e uma planilha com os dados em forma numérica (FIGURA 7).

Figura 5 – *Frame* do vídeo com o processamento das emoções em quadro¹⁴

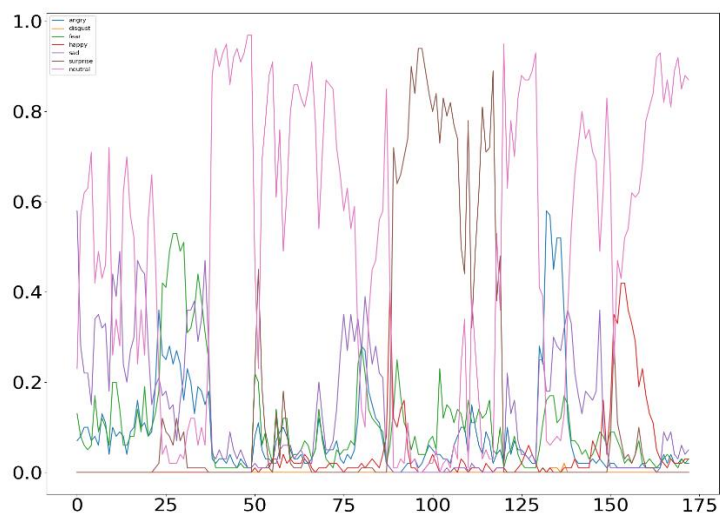


Fonte: Elaborado pelo autor.

¹⁴ A pessoa nesta imagem não faz parte do conjunto de participantes desta pesquisa. A voluntária é o cônjuge deste pesquisador

A Figura 5 mostra como o programa *Visage* apresenta a função de rastreamento facial e a análise de expressões faciais que classificam as emoções básicas humanas, de acordo com o proposto pela teoria das emoções universais.

Figura 6 – Gráfico gerado pelo programa com a prevalência de emoções pelo tempo



Fonte: Elaborado pelo autor.

A Figura 6 apresenta o gráfico gerado pelo *Visage*, que relaciona a intensidade da expressão facial e sua emoção correspondente pelo tempo decorrido do vídeo. O Índice de Intensidade da Emoção (IIE) vai de 0 (ausente) a 1.0 (completamente presente). Ressaltamos que este índice tem que apresentar a soma das partes igual a 1. Sendo assim, ele também funciona como um medidor de predominância de uma emoção pelo tempo. O tempo é medido em quadros, onde 1 segundo equivale a 24 quadros.

Figura 7 – Tabela gerada pelo *Visage*, com os dados numéricos de um teste

	A	B	C	D	E	F
1	angry0,box0,disgust0,fear0,happy0,neutral0,sad0,surprise0					
2	0.07,"(481, 254, 349, 349)",0.0,0.13,0.0,0.23,0.58,0.0					
3	0.08,"(477, 250, 352, 352)",0.0,0.08,0.0,0.57,0.28,0.0					
4	0.1,"(475, 250, 356, 356)",0.0,0.06,0.0,0.62,0.22,0.0					
5	0.1,"(477, 252, 354, 354)",0.0,0.05,0.0,0.63,0.22,0.0					
6	0.07,"(479, 254, 350, 350)",0.0,0.06,0.0,0.71,0.15,0.0					
7	0.08,"(479, 255, 347, 347)",0.0,0.17,0.0,0.42,0.34,0.0					
8	0.06,"(483, 252, 348, 348)",0.0,0.09,0.0,0.49,0.35,0.0					
9	0.13,"(480, 254, 350, 350)",0.0,0.12,0.0,0.43,0.32,0.0					
10	0.11,"(480, 255, 349, 349)",0.0,0.1,0.0,0.46,0.33,0.0					
11	0.04,"(475, 241, 358, 358)",0.0,0.06,0.0,0.72,0.18,0.0					
12	0.1,"(479, 252, 349, 349)",0.0,0.2,0.0,0.26,0.44,0.0					
13	0.08,"(480, 254, 347, 347)",0.0,0.2,0.0,0.34,0.39,0.0					
14	0.09,"(478, 252, 349, 349)",0.0,0.14,0.0,0.28,0.49,0.0					
15	0.08,"(476, 242, 356, 356)",0.0,0.06,0.0,0.62,0.24,0.0					
16	0.04,"(475, 242, 355, 355)",0.0,0.06,0.0,0.7,0.2,0.0					
17	0.09,"(475, 243, 355, 355)",0.0,0.08,0.0,0.57,0.27,0.0					
18	0.1,"(476, 243, 355, 355)",0.0,0.08,0.0,0.52,0.3,0.0					
19	0.16,"(483, 252, 347, 347)",0.0,0.14,0.0,0.24,0.47,0.0					
20	0.1,"(482, 251, 350, 350)",0.0,0.09,0.0,0.36,0.45,0.0					
21	0.11,"(481, 253, 348, 348)",0.0,0.19,0.0,0.26,0.44,0.0					
22	0.08,"(483, 250, 349, 349)",0.0,0.08,0.0,0.56,0.28,0.0					

Fonte: Elaborado pelo autor

A Figura 7 apresenta a tabela com os dados numéricos gerados pelo *Visage*. Cada linha apresenta os valores de todas as emoções captadas, bem como as coordenadas do posicionamento da face no plano do vídeo. As linhas equivalem aos quadros do vídeo, sendo a primeira uma descrição da ordem das emoções na tabela. As colunas equivalem as emoções e a caixa que é gerada pela face do participante. Os nomes das emoções estão ordenadas alfabeticamente em inglês.

3.6 Seleção dos participantes

Foram formados dois grupos de 8 participantes surdos cada. O grupo 1 foi formado por 8 surdos e o grupo 2 também por 8, totalizando, assim, 16 pessoas. Tivemos 8 participantes do sexo masculino e 8 do feminino. A faixa de idade é entre 20 e 41 anos, sendo a idade mínima de 20 anos e a máxima de 41, média de 28,75 anos, mediana de 24,5 e desvio-padrão de 7,27. Esperávamos chamar participantes com idade mais homogênea, mas, por causa da pandemia do COVID-19, tivemos dificuldades em ter participantes surdos. Dentre eles, 15 surdos são naturais do estado do Ceará e 1 do estado de Pernambuco.

Todos os surdos têm formação básica, sendo 1 com ensino técnico, 4 com superior incompleto, 4 com pós-graduação e 7 com superior completo. Os surdos são fluentes em Libras, tendo-a como L1, tendo aprendido na escola; apresentam surdez pré-linguística, se

consideram leitores regulares da língua portuguesa e são todos usuários assíduos de filmes legendados por meio de *streaming* e cinema.

Foi pedido a todos os surdos seus testes audiométricos. Todos os participantes assinaram um Termo de Consentimento Livre e Esclarecido para participar da pesquisa. Para executarmos a pesquisa com seres humanos submetemos nosso trabalho à avaliação do Comitê de Ética, sendo aprovado no CAAE 51396921.2.0000.5054.

3.7 Desenho experimental

Com o propósito de averiguar os impactos que a tradução de efeitos sonoros na LSE gera no público surdo, a pesquisa tem como norte o seguinte questionamento:

- A tradução dos efeitos sonoros na LSE de um filme de ação influencia na emoção dos espectadores surdos durante a recepção?

O objetivo geral deste estudo é investigar se a tradução dos efeitos sonoros na LSE de um filme de ação intensifica a emoção dos espectadores surdos. Ademais, nossos objetivos específicos são:

- 1) Investigar a intensidade das emoções na recepção da LSE por meio do modelo de tradução dos efeitos sonoros proposto por Nascimento (2018);
- 2) Mensurar a diferença sobre as emoções mediante o uso e não-uso da tradução de efeitos sonoros na LSE.

Para ilustrar a execução das atividades com os agrupamentos de surdos que proveram os dados essenciais a este trabalho, é apresentado, a seguir, o quadro metodológico da pesquisa, seguindo o Quadrado Latino (QUADRO 4).

Quadro 4 – Quadrado Latino com a disposição dos participantes por trechos a serem expostos

VÍDEO	CONDIÇÃO	PAR 1	PAR 2
1	LSE-CSOM	S1 / S9	S3 / S11
	LSE-SSOM	S2 / S10	S4 / S12
2	LSE-CSOM	S5 / S13	S7 / S15
	LSE-SSOM	S6 / S14	S8 / S16
3	LSE-CSOM	S2 / S10	S4 / S12
	LSE-SSOM	S1 / S9	S3 / S11
4	LSE-CSOM	S6 / S14	S8 / S16

	LSE-SSOM	S5 / S13	S7 / S15
--	----------	----------	----------

Fonte: Elaborado pelo autor.

O Quadro 4 ilustra como os participantes foram expostos aos trechos. Cada participante foi exposto a somente dois vídeos, cada um em uma condição distinta: LSE-CSOM e LSE-SSOM. Os 16 participantes, portanto, foram separados em dois grupos: grupo 1 e 2, para efeito de comparação dos estímulos. Este estudo apresenta 8 estímulos, sendo cada vídeo de cada condição assistido quatro vezes. Um estímulo de um participante não foi gravado por problemas com o programa de captura da *webcam*.

Todas essas legendas foram manipuladas em português formal e adaptadas para a LSE. A única diferença entre elas é a presença ou ausência de traduções de efeitos sonoros, seguindo o modelo de Nascimento (2018) em LSE.

Os voluntários que participaram dos testes foram alocados presencialmente no laboratório de psicolinguística da UFC. Os participantes tiveram um intérprete de Libras, que é bolsista pesquisador do Programa Institucional de Bolsas de Iniciação Científica – PIBIC/UFC e é aluno do 7º semestre do curso Letras – Libras, da UFC. Todos assistiram aos vídeos em um computador com o sistema operacional Windows 10, com tela de 21 polegadas, para garantir uma boa visibilidade tanto dos vídeos quanto da LSE. Tivemos disponível para gravação, uma câmera do tipo *webcam* embutida no monitor, com resolução de 720p, filmando-os a 60 quadros por segundo.

O computador emitiu os sons dos extratos fílmicos, para servir de marcador temporal para a edição dos vídeos com as faces dos participantes. Isso se deu porque há um atraso entre começar a gravação manualmente e a veiculação do trecho, sendo necessária a edição, para evitar a computação de dados referentes aos momentos em que o participante não estava sendo exposto ao filme.

Os dois agrupamentos, de 8 membros cada, tiveram seus dados coletados individualmente. Antes de ser exposto a cada trecho, os participantes preencheram o questionário demográfico em português, com o auxílio do intérprete de Libras. Após a finalização do questionário, foram expostos aos dois trechos do filme correspondentes ao seu código.

Ao finalizar cada vídeo, os participantes preencheram o questionário *Differential Emotion Scale* (DES) para cada um dos trechos assistidos, totalizando dois por participante. Todos os questionários aqui descritos foram disponibilizados pela ferramenta Google Forms no mesmo computador utilizado para coleta, a fim de que os participantes pudessem

preenchê-los com maior rapidez e sem rasuras. A plataforma de questionários Google Forms também foi escolhida por proporcionar uma análise de dados de maneira mais prática e rápida, sendo possível agrupar e gerar gráficos, por exemplo, com pouco esforço, se comparado à aplicação por meio de papel. Uma dificuldade que surgiu foi a ausência de internet no laboratório, inviabilizando a aplicação do questionário no computador da coleta. Esta barreira foi transposta mediante uso de um celular, com os questionários em sua tela, entregue ao participante para execução da tarefa.

A sequência de coleta foi: 1. Preenchimento dos questionários; 2. Assinatura do Termo de Consentimento Livre e Esclarecido; 3. Exibição dos trechos, tendo a face monitorada pelo *Visage*; 4. Salvamento dos dados em uma pasta no Google Drive.

Todos os vídeos com a gravação das faces dos participantes tiveram que ser editados, para apresentar a duração exata de exposição ao filme, eliminando tempos desnecessários. Todas as edições foram feitas no mesmo programa utilizado para cortar os trechos do filme. Os arquivos foram convertidos para a taxa de atualização de 30 quadros por segundo, para evitar longos tempos de trabalho no *Visage*. Ressaltamos que executamos testes prévios com ambas as taxas de quadros por segundo e os resultados foram os mesmos. Uma vez editados, o programa analisava os vídeos, rastreando quadro a quadro as faces, em busca das expressões que indicassem as emoções expostas por cada participante durante a fase de recepção dos estímulos e gerando planilhas e gráficos sobre as emoções. A análise de cada trecho de dois minutos e meio de gravação, em torno de 4500 quadros, pelo *Visage* levou, em média, duas horas para ser executada. Foram utilizados dois computadores, com placas de vídeo dedicadas, simultaneamente, para cortar o tempo total de processo de todos os vídeos ao meio.

Devido à pandemia do COVID-19, iniciamos a coleta somente no final de outubro de 2021, após aprovação no Comitê de Ética. O avanço da vacinação no âmbito estadual e municipal proveu mais segurança na aplicação deste experimento.

Para aplicação da pesquisa de maneira presencial, seguimos os protocolos de segurança estabelecidos pelas autoridades sanitárias. As medidas tomadas foram:

- Presença de, no máximo, três pessoas dentro do laboratório;
- Uso de máscara por participantes e pesquisadores;
- Higienização do computador utilizado a cada participante;
- Manutenção da distância mínima de dois metros entre os indivíduos dentro da sala;

- Álcool em gel 70% para higienização das mãos.

Ao final de todas as coletas, iniciamos a análise dos dados quantitativos gerados pela captura facial provenientes do *software Visage*.

3.8 Variáveis e medidas

A pesquisa apresenta dois tipos de variáveis: dependentes e independentes. As variáveis independentes são as LSEs com a presença ou não da tradução de efeitos sonoros. Já as dependentes são os dados relativos às seis emoções da Teoria Universal das Emoções mais o neutro.

3.8.1 Variáveis independentes

De acordo com Saldanha e O'Brien (2014), uma variável de pesquisa é dita independente quando ela pode ser manipulada pelo pesquisador para ver qual será o resultado sobre as variáveis dependentes.

Nesta pesquisa, a LSE-CSOM e a LSE-SSOM são necessárias para podermos investigar a influência da tradução de efeitos sonoros nos sentimentos dos surdos. Essa variação nos permite perceber se há mesmo alguma diferença para os surdos quando os sons são traduzidos.

3.8.2 Variáveis dependentes

São consideradas variáveis dependentes tudo que for influenciável pelas variáveis independentes (SALDANHA e O'BRIEN, 2014). Nesta pesquisa, as variáveis dependentes, portanto, são os dados das emoções que podem ser lidas pelas expressões faciais, de acordo com a Teoria das Emoções Universais (VIEIRA, 2015), ao todo seis emoções e o neutro, que é a ausência de emoção:

- Alegria (*Happy*);
- Tristeza (*Sad*);
- Raiva (*Angry*);
- Medo (*Fear*);
- Nojo (*Disgust*);

- Neutro (*Neutral*);
- Surpresa (*Surprise*).

Todas essas emoções são lidas a partir das expressões faciais dos participantes e, a posteriori, são interpretadas pelo programa *Visage*, de reconhecimento e expressões faciais e de interpretação de emoções.

3.9 Análise dos dados

Esta pesquisa obteve dados dos dois questionários e da análise das emoções pelo programa *Visage*. Os dados do primeiro questionário (QUADRO 3) serviram para traçar o perfil demográfico dos participantes. Esse tipo de dado é útil para traçar perfil de nossos participantes.

O segundo questionário, DES, forneceu dados quali-quantitativos, por meio de uma escala numérica que representa a autopercepção do participante quanto a suas emoções quando assiste aos trechos do filme. No entanto, esses dados não foram analisados nesta pesquisa, ficando para trabalhos a posteriori, uma vez que não tivemos tempo hábil para fazer tanto a análise estatística como seu cruzamento com os dados provenientes do *Visage*.

Esses dados foram analisados por um teste e um modelo estatísticos: Análise de Variância (ANOVA) e o Modelo Linear Misto, apresentados na próxima seção. O teste ANOVA foi utilizado por demonstrar a existência de significância entre variáveis. Fizemos uso, também, do LMM, por ser um modelo aplicável a variáveis categóricas – CSOM e SSOM –, ideal para nossa pesquisa.

Os valores numéricos da escala proveniente do programa são os mesmos que permitem gerar gráficos da prevalência das emoções pelo tempo de exposição. Os gráficos nos permitem identificar se as emoções em cada trecho são comuns para cada participante.

O programa utilizado para se executar a análise estatística foi o RStudio, programa que faz uso da linguagem R, própria para a área de estatística. Esta linguagem permite diversas abordagens sobre dados, como tarefas de cálculo de estatística descritiva e inferencial. Assim como a linguagem Python, com bibliotecas certas, o R consegue exportar gráficos e tabelas para a melhor visualização dos dados estatísticos, além de executar diversas análises estatísticas.

A comparação das informações matemáticas provenientes do *Visage* nos proporcionaram maior poder de inferência sobre o comportamento emocional dos participantes durante a recepção da LSE com ou sem a tradução dos efeitos sonoros.

No próximo capítulo veremos os dados obtidos e como trabalhamos estatisticamente com eles para podermos fazer inferências a respeito da funcionalidade da tradução de efeitos sonoros.

4 RESULTADOS

Este capítulo está dividido em duas subseções. A primeira se dedica a apresentar os dados brutos coletados pelo *Visage* e a sua análise descritiva. A segunda subseção trata dos dados retirados a partir da Análise de Variância (ANOVA) e do Modelo Linear Misto. A partir da apresentação desses dados, faremos inferências quanto aos resultados obtidos.

4.1 Análise descritiva dos resultados quantitativos

Nossa pesquisa contou com o total de 16 participantes ($n=16$), sendo cada um estimulado por dois trechos distintos do filme *Dunkirk* (2017). Um trecho era veiculado na condição CSOM, com tradução de efeitos sonoros; e o outro na condição SSOM, sem tradução dos efeitos sonoros. O total de vídeos com as faces dos participantes foi de 31, gerando 135.790 linhas de dados quantitativos sobre as emoções dos participantes, com cada linha representando um quadro captado pela filmagem da face pela *webcam*. Os dados referentes à condição CSOM tiveram um total de 70.143 linhas; e os da condição SSOM, 65.647.

A análise desses dados foi feita separando e cruzando as variáveis preditoras (CSOM e SSOM) com as respostas (*angry, disgust, fear, happy, neutral, sad, surprise*), com o objetivo de perceber diferenças significativas entre as condições preditoras. Ressaltamos que quando formos nos referir às variáveis, faremos uso do seu nome em inglês, o mesmo compreendido pelo programa. Para nos referirmos à emoção, no entanto, utilizaremos o correspondente em português.

Em algumas das variáveis respostas, retiramos seus *outliers*, para não confundir a análise estatística dos dados. *Outlier*, valor atípico, é todo valor que se distancia muito do restante, podendo indicar, assim, um erro de coleta ou pontos díspares que não representem os valores como um todo.

Adiante, apresentamos os dados obtidos seccionados pelas variáveis resposta. Estas variáveis, bem como alguns termos estatísticos, têm seus nomes apresentados em inglês, visto que é dessa maneira que o programa RStudio e o *Visage* geram seus dados. Todos os valores têm algum tipo de arredondamento, para não apresentarmos números excessivamente longos. As médias têm arredondamento de três casas decimais, enquanto as demais

apresentam, pelo menos, duas casas decimais, com exceção do erro-padrão, que por ser um número muito pequeno, requereu quatro dígitos.

4.1.1 Angry

A variável resposta *angry*, raiva, apresentou valores baixos no IIE, que vai de 0 – para não apresentou – a 1 – para apresentou completamente. As médias em ambas as variáveis preditoras, CSOM e SSOM, foi de 0,106, indicando que não foi uma emoção muito sentida em ambos os grupos (TABELA 4). Os valores baixos do erro-padrão indicam alta confiabilidade destas médias. A repetição das médias nas duas condições nos sugere que não houve diferença significativa entre elas, conforme a tabela 4.

Tabela 4 – Dados descritivos da variável resposta *angry*

MANIPULAÇÃO	MEAN	MEDIAN	SD	SE	MIN	MAX
CSOM	0.106	0.08	0.09	0.0003	0	0.83
SSOM	0.106	0.08	0.10	0.0004	0	0.79

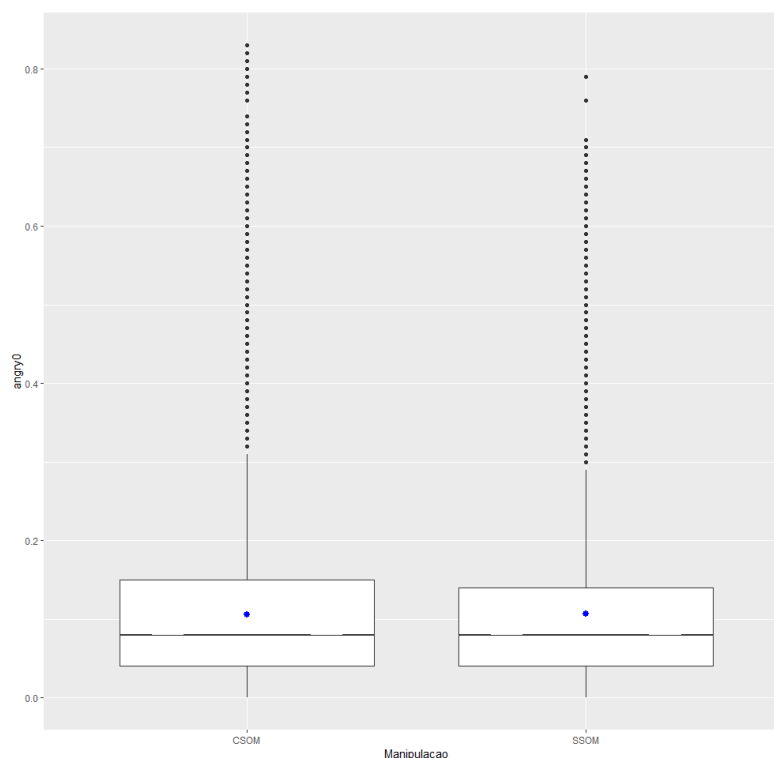
Fonte: Elaborada pelo autor

Mean = Média; SD = Desvio-padrão; SE = Erro-padrão; Min = Mínimo; Max = Máximo.

Outros dados que também podemos observar pela Tabela 4 são as medianas. A mediana de 0,08 em ambas as condições evidencia que foi uma emoção pouco sentida pelos participantes. A mediana permite sabermos o valor central de uma sequência. Neste caso, ela permite inferirmos que, apesar de uma média mais alta, o sentimento de raiva foi muito pouco sentido e o que possivelmente eleva a média é a presença de muitos *outliers* (GRÁFICO 1).

O desvio-padrão entre as medidas apresenta variação de 0,01, indicando que a condição CSOM tem sua distribuição marginalmente mais homogênea que a outra condição. As duas condições, entretanto, apresentam um valor próximo de zero (0,09 e 0,10), o que indica haver homogeneidade dos valores.

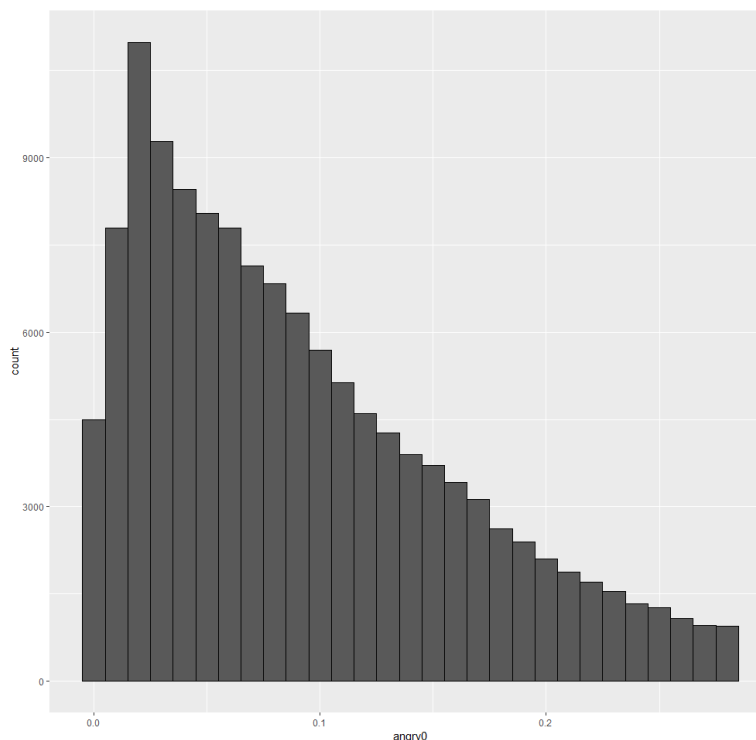
O gráfico 1 é um *boxplot* dos valores de *angry* e suas médias nas duas condições, CSOM e SSOM.

Gráfico 1 – *Boxplot* da variável *angry0* em ambas as variáveis predictoras

Fonte: Elaborado pelo autor

Quando geramos um gráfico *boxplot* (GRÁFICO 1) com os valores da emoção raiva, percebemos que há grande presença de valores atípicos nesta medida, ou seja, os *outliers*. A presença dos *outliers* poderia, de certa forma, influenciar negativamente a confiabilidade das análises estatísticas que executamos. Então, estabelecemos uma linha de corte, a fim de retirar todos em ambas as condições. A linha de corte estabelecida para esta finalidade foi de 0,28.

Uma vez eliminados os valores atípicos, podemos observar que a distribuição dos dados se torna assimétrica, neste caso, uma assimetria positiva, ou seja, os dados geram uma cauda para a direita (GRÁFICO 2). Provavelmente, essa assimetria acontece por raiva ser uma emoção, mediante os estímulos, pouco observada nos participantes.

Gráfico 2 – Histograma da variável *angry* totalizada e sem *outliers*

Fonte: Elaborado pelo autor

Não esperávamos dados de valor elevado, como 0,5 ou superior, na emoção raiva, já que os trechos selecionados do filme não visam suscitar tais emoções. Estes valores estão de acordo com nossas expectativas.

Fazendo uso somente de uma análise descritiva, não podemos inferir, portanto, que a legendagem de efeitos sonoros em um filme de ação cause diferença significativa na emoção raiva.

4.1.2 *Disgust*

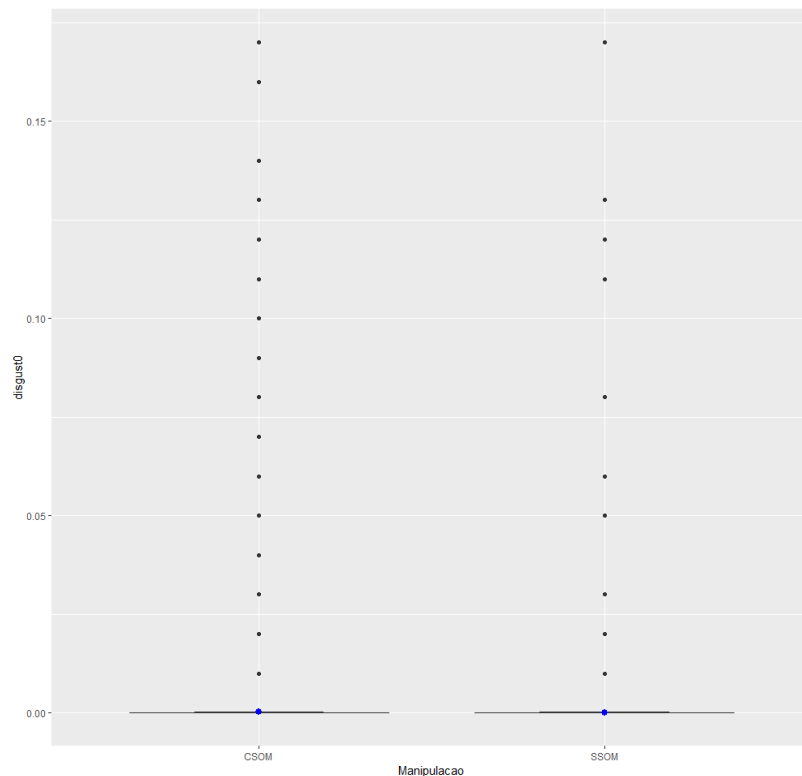
A variável resposta *disgust*, nojo, apresentou valores próximos de zero no IIE. A média na variável preditora CSOM foi de um número que se aproxima de zero: 0,00018; e na variável SSOM, seu valor é de $8.96e-05$, também aproximando-se de zero. Podemos, portanto, inferir que não foi uma emoção sequer sentida pelos participantes surdos (TABELA 5), tendo valores gerados por erro de captação ou pequenos espasmos musculares captados pelo programa. Os valores do erro-padrão também são muito baixos, indicando alta confiabilidade das médias que tendem a zero.

Tabela 5 - Dados descritivos da variável resposta *disgust*

MANIPULAÇÃO	MEAN	MEDIAN	SD	SE	MIN	MAX
CSOM	0.00018	0	0.003	0.0000	0	0.17
SSOM	8.96e-05	0	0.0015	0.0000	0	0.17

Fonte: Elaborado pelo autor

Outros dados que também podemos observar na Tabela 5 são as medianas, que é zero em ambos os casos. A mediana zerada indica que não foi uma emoção expressa pelos participantes. O desvio-padrão entre as medidas apresenta variação de 0,03, para CSOM; e 0,0015, para SSOM, indicando que a condição SSOM tem sua distribuição marginalmente mais homogênea que a outra condição. Ambas, entretanto, apresentam um valor próximo de zero (0,03 e 0,0015), o que indica haver homogeneidade dos valores. O mínimo de 0 e o máximo atingindo o valor de 0,17 permite-nos inferir que nojo não foi sequer significativamente expressa. O gráfico 3 traz os valores para a variável *disgust* em um *boxplot*.

Gráfico 3 – *Boxplot* da variável *disgust* em ambas as variáveis predictoras

Fonte: Elaborado pelo autor

O gráfico 3 permite-nos observar a dispersão de poucos *outliers* na emoção nojo. Não conseguimos gerar gráficos de distribuição, visto que os valores se situavam em zero. A média e mediana se aproximando de 0 nos permite inferir que esta emoção é irrelevante para os trechos selecionados do filme de ação *Dunkirk* (2017). O esperado era um valor baixo para esta emoção, mas que não chegasse a zero.

Uma hipótese que se levanta para este resultado é de que as cenas não continham imagens gráficas de lesão corporal ou afins, sendo, assim, mais focada na ação do que em causar repulsa nos indivíduos. Apesar de rodarmos mais testes estatísticos em relação a esta emoção, acreditamos que o sentimento de nojo não contribuiu, nesta pesquisa, para a confirmação de nossas hipóteses.

4.1.3 Fear

A variável resposta *fear*, medo, tem alta relevância para este trabalho, visto que é uma emoção que esperamos que nossos estímulos a suscitem. Ela, entretanto, apresentou valores tendendo a zero no IIE. A média na variável preditora CSOM foi de um número próximo de zero, 0,058; e na variável SSOM, seu valor é de 0,054, também aproximando-se de zero, havendo uma diferença marginal entre as duas. A partir das médias, inferimos que o medo foi uma emoção pouco sentida em ambos os grupos (TABELA 6). Os baixos valores do erro-padrão também indicam alta confiabilidade destas médias.

Tabela 6 - Dados descritivos da variável resposta *fear*

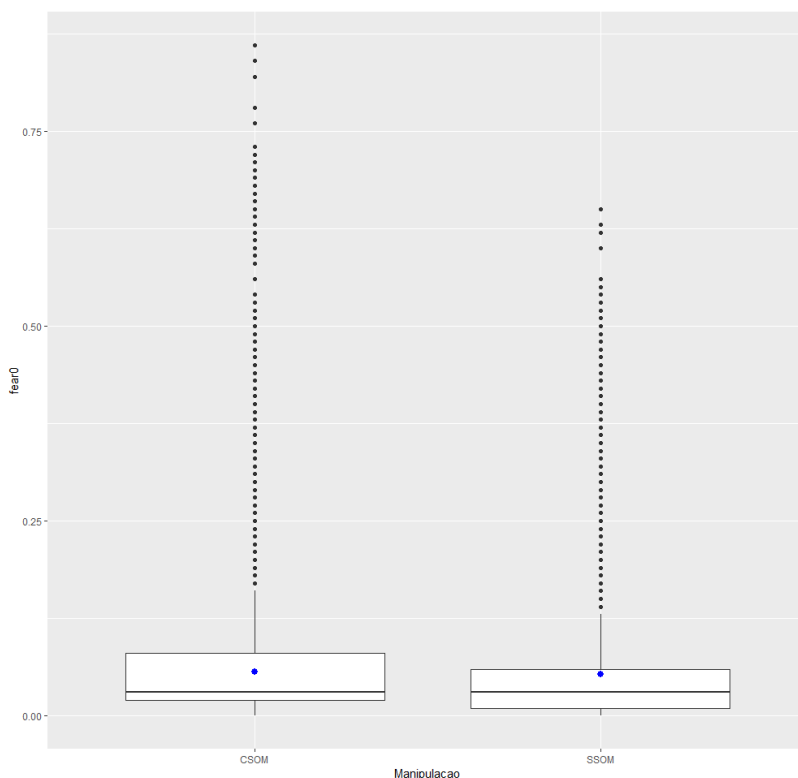
MANIPULAÇÃO	MEAN	MEDIAN	SD	SE	MIN	MAX
CSOM	0.058	0.03	0.062	0.0002	0	0.86
SSOM	0.054	0.03	0.076	0.0003	0	0.65

Fonte: Elaborado pelo autor

A partir da Tabela 6, podemos observar o valor de 0,03 de ambas as medianas. A mediana muito baixa, mas não zerada, indica que medo foi uma emoção marginalmente expressa pelos participantes em ambas as condições. O desvio-padrão entre as medidas apresenta variação de 0,062, para a condição CSOM; e 0,076, para a SSOM, indicando que a condição CSOM tem sua distribuição marginalmente mais homogênea que a outra condição. O mínimo de 0 e o máximo atingindo o valor de 0,86 e 0,75, permite-nos imaginar que *fear*

teve breves picos de expressão, podendo significar presença de *outliers*. Para tal, geramos um diagrama de caixa para melhor visualizar estes dados (GRÁFICO 4).

Gráfico 4 – *Boxplot* da variável *fear* em ambas as variáveis predictoras

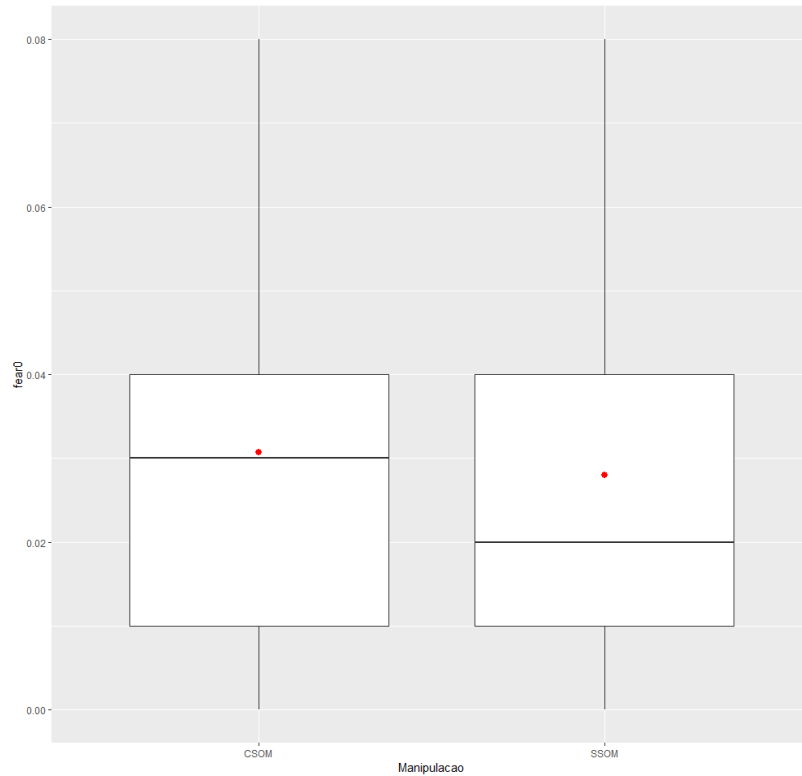


Fonte: Elaborado pelo autor

Observando o diagrama de caixa com os valores da emoção medo (GRÁFICO 4), percebemos que há grande presença de *outliers* nessa medida, principalmente na condição CSOM. A fim de limparmos os dados, estabelecemos a linha de corte de valores atípicos em 0,08, ponto onde não há mais a sua presença nos dados (GRÁFICO 5).

A ausência dos *outliers*, apresentada no gráfico 5, permite que visualizemos melhor a diferença entre as condições (GRÁFICO 5). Visualmente, a média e mediana da condição CSOM são marginalmente superiores as da condição SSOM. Para inferirmos algo, no entanto, são necessárias abordagens estatísticas mais aprimoradas que a observação da média e mediana, o que veremos na próxima seção deste capítulo.

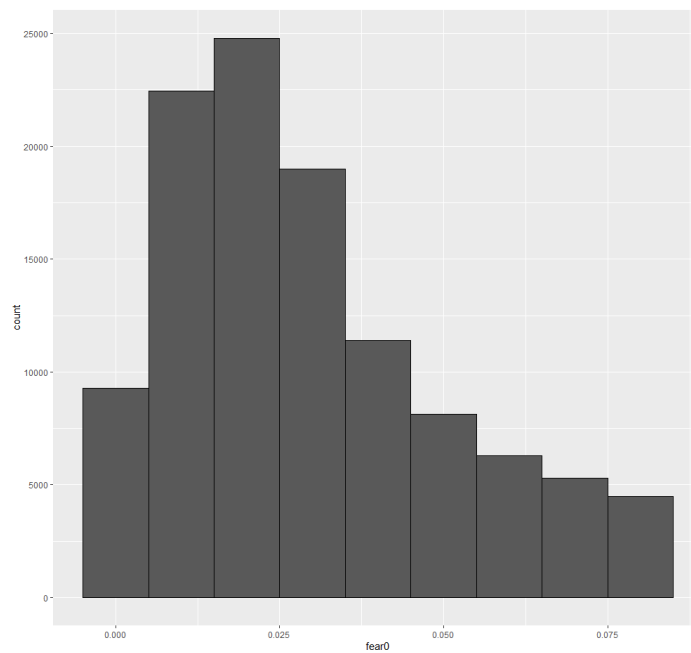
Gráfico 5 – *Boxplot* da variável *fear* em ambas as variáveis preditoras sem *outliers*



Fonte: Elaborado pelo autor

Gerando um gráfico de histograma (GRÁFICO 6), podemos ver novamente uma assimetria positiva nos dados da emoção medo.

Gráfico 6 – Histograma da variável *fear* totalizada



Fonte: Elaborado pelo autor

Esperávamos que a emoção medo fosse apresentar diferença mais significativa na condição CSOM, visto que os trechos do filme geram bastante angústia e medo, entretanto, não foi isso que se demonstrou. Fazendo uso somente de uma análise descritiva, poderíamos inferir que a legendagem de efeitos sonoros em um filme de ação não causa diferença significativa nesta emoção. Entretanto, ainda veicularemos nossas conclusões até que os resultados dos demais testes estatísticos em relação a esta emoção sejam apresentados.

4.1.4 Happy

A variável resposta *happy*, alegria, também apresentou valores tendendo a zero no IIE, mas com maior dispersão. A média na variável preditora CSOM foi próximo a zero: 0,055; e na variável SSOM, seu valor é de 0,049, aproximando-se ainda mais de zero, havendo uma diferença pouco considerável entre as duas. A partir das médias, inferimos que a alegria foi uma emoção pouco sentida em ambos os grupos (TABELA 7). Os baixos valores do erro-padrão também indicam alta confiabilidade das médias.

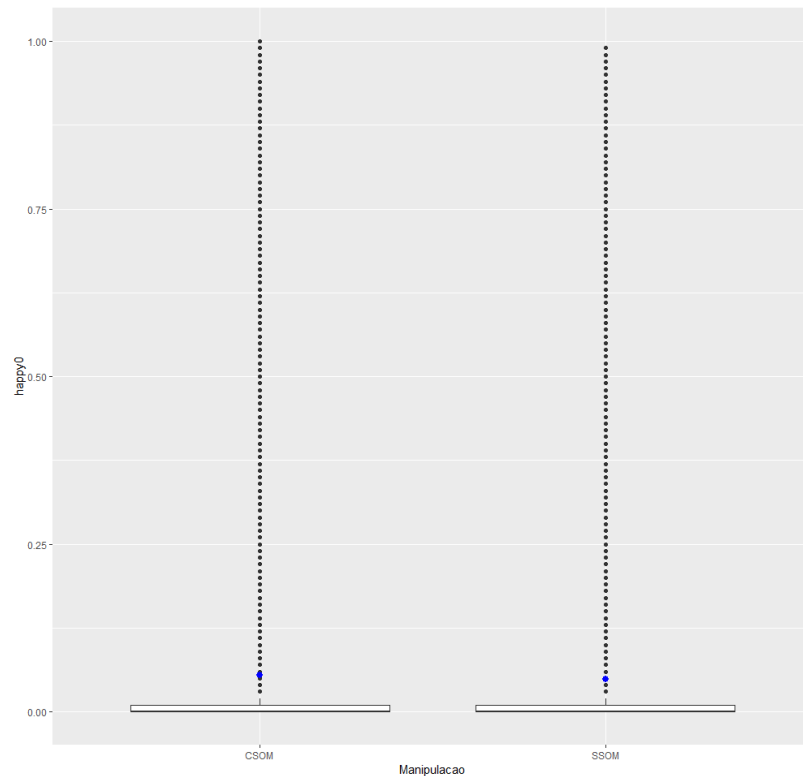
Tabela 7 - Dados descritivos da variável resposta *happy*

MANIPULAÇÃO	MEAN	MEDIAN	SD	SE	MIN	MAX
CSOM	0.055	0	0.196	0.0007	0	1
SSOM	0.049	0	0.175	0.0006	0	0.99

Fonte: Elaborado pelo autor

A partir da Tabela 7, podemos observar o valor de 0 em ambas as medianas. A mediana zerada indica que a alegria foi uma emoção que em boa parte do tempo sequer foi expressa pelos participantes, em ambas as condições.

O desvio-padrão entre as medidas apresenta variação de 0,196, para a condição CSOM; e 0,175, para a SSOM, indicando que a condição SSOM tem sua distribuição mais homogênea que a outra condição. Na CSOM, o valor mínimo foi 0 e o máximo é 1; em SSOM, o valor mínimo foi 0 e o máximo foi 0,99, permitindo-nos inferir que a alegria foi a emoção que teve picos de sua expressão. No entanto, esses picos foram breves, já que as médias e medianas são bem próximas a zero, podendo significar, então, presença de valores atípicos. Geramos um diagrama de caixa para melhor visualizar estes dados (GRÁFICO 7).

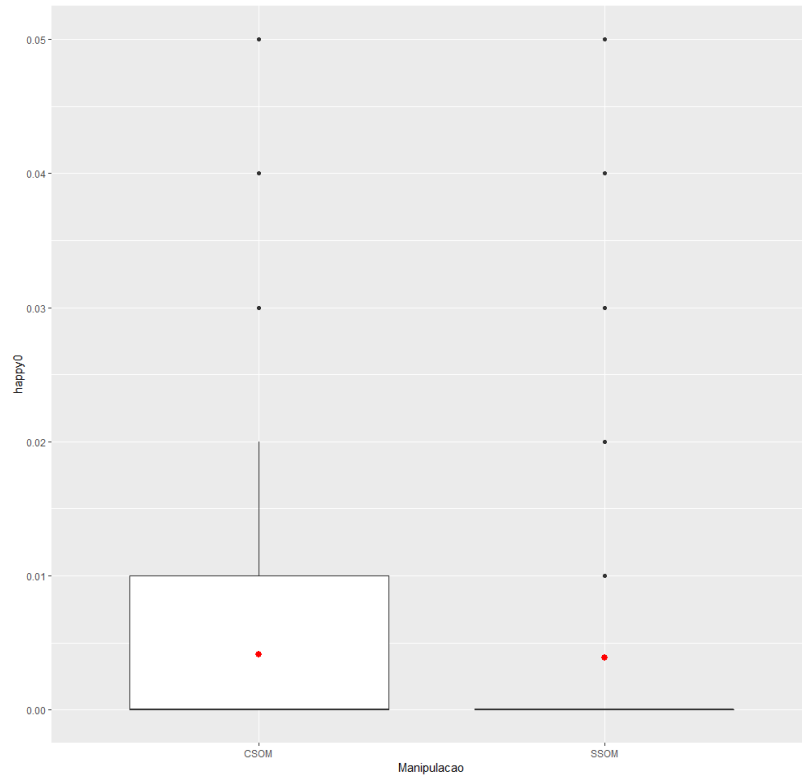
Gráfico 7 – *Boxplot* da variável *happy* em ambas as variáveis preditoras

Fonte: Elaborado pelo autor

Observando o diagrama de caixa com os valores da emoção alegria (GRÁFICO 7), percebemos que há grande presença de *outliers* nessa medida, em ambas as condições. Para a limpeza deles, estabelecemos a linha de corte de valores atípicos em 0,05 (GRÁFICO 7). Mesmo assim, ainda é possível observar a presença de *outliers* no gráfico. Decidimos, portanto, mantê-los, por perceber que, sem a sua existência, os valores zeravam completamente, não permitindo a nós fazer nenhuma análise estatística com nossos testes e modelos.

A diminuição da presença dos valores atípicos permite que visualizemos melhor a diferença entre as condições (GRÁFICO 8). Visualmente, a média e mediana da condição CSOM são marginalmente superiores as da condição SSOM. A distribuição da condição SSOM é zerada, enquanto a condição CSOM ainda apresenta distribuição de valores no terceiro quartil, correspondente ao valor 0,01, e uma pequena variabilidade superior, atingindo 0,02. Esses dados nos permitem perceber uma diferença mínima entre os valores das duas condições. O gráfico 8 mostra a distribuição dos dados com filtragem dos *outliers*.

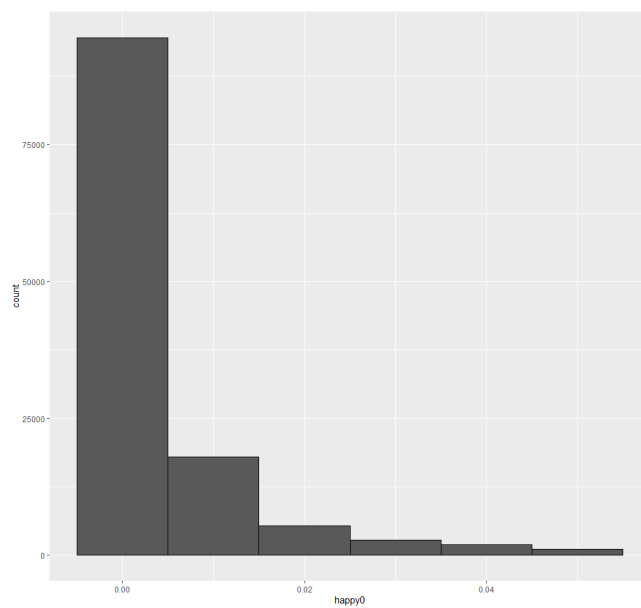
Gráfico 8 – *Boxplot* da variável *happy* em ambas as variáveis preditoras sem muitos *outliers*



Fonte: Elaborado pelo autor

No gráfico 9, com a redução dos *outliers*, podemos visualizar uma assimetria positiva nos dados da emoção alegria. A maioria dos valores, entretanto, é zero.

Gráfico 9 – Histograma da variável *happy* totalizada



Fonte: Elaborado pelo autor

Esperávamos que a emoção alegria não fosse apresentar grandes valores nas duas condições, visto que os trechos do filme não buscam suscitar essa emoção. Partindo somente do uso da análise descritiva, inferimos que a legendagem de efeitos sonoros em um filme de ação não causa diferença significativa na emoção alegria. Os dados ainda serão apresentados devidamente e analisados por nossos modelos estatísticos para que possamos chegar a uma melhor conclusão.

4.1.5 *Neutral*

A variável resposta *neutral*, neutro, para alguns estudiosos da área não é considerada uma expressão de emoção, mas sim, sua ausência (DUTTA e BARMAN, 2020). No entanto, sua relevância é grande para esta pesquisa, já que ela demonstra se houve alguma reação, ou não, aos estímulos. As médias foram as segundas maiores de todas as variáveis preditoras, 0,361, na CSOM; e 0,385, na SSOM, apresentando uma diferença de 0,024. A partir das médias, inferimos que o neutro foi muito expresso pelos participantes (TABELA 8). Os baixos valores do erro-padrão também indicam alta confiabilidade das médias.

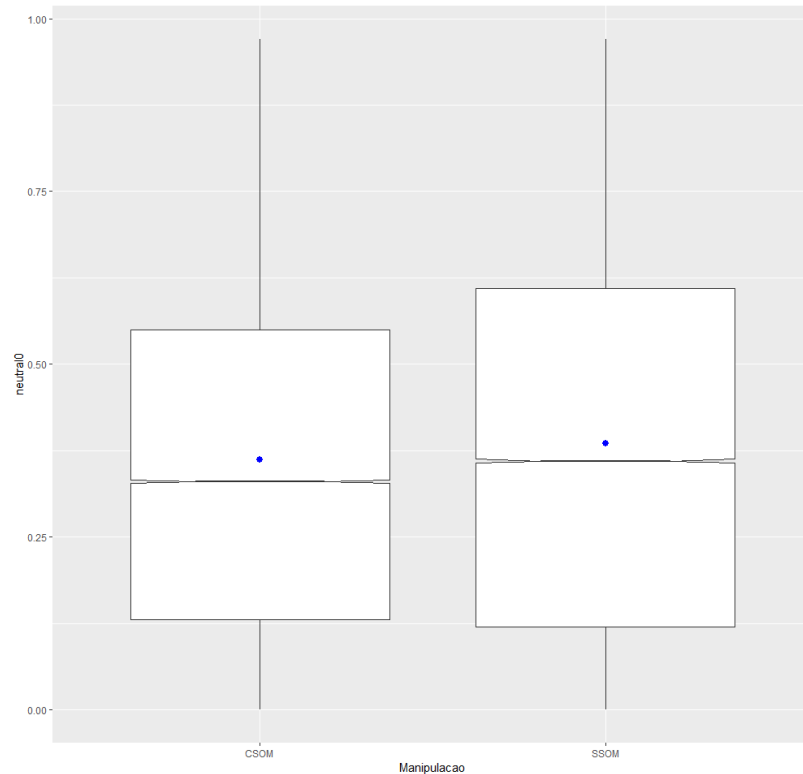
Tabela 8 - Dados descritivos da variável resposta *neutral*

MANIPULAÇÃO	MEAN	MEDIAN	SD	SE	MIN	MAX
CSOM	0.361	0.33	0.252	0.001	0	0.97
SSOM	0.385	0.36	0.275	0.001	0	0.97

Fonte: Elaborado pelo autor

A Tabela 8 apresenta valores distintos em ambas as medianas: 0,33 e 0,36. A mediana nessa tabela indica que a expressão neutra foi predominante para as duas condições durante os estímulos, sendo pouco superior na condição SSOM. Valores mais altos da mediana indicam permanência maior de tempo neste estado.

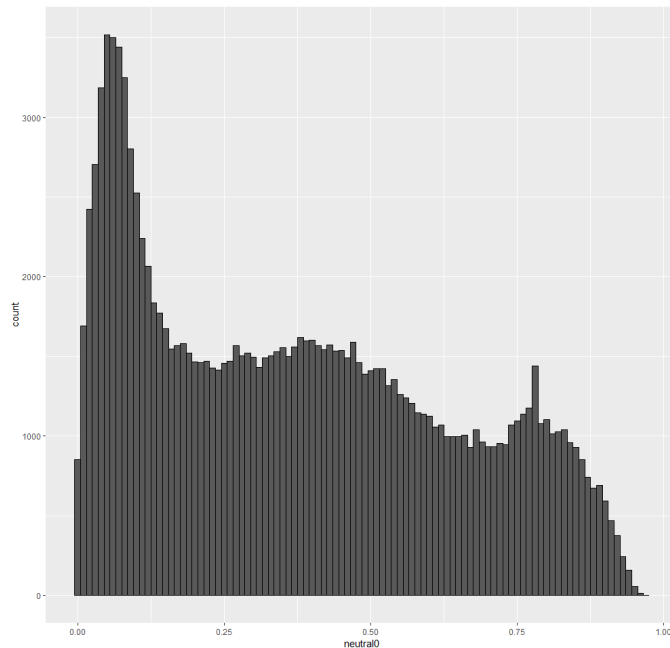
O desvio-padrão entre as medidas apresenta variação de 0,252, para a condição CSOM; e 0,275, para a SSOM, indicando que a condição CSOM tem sua distribuição mais homogênea que a outra condição. O mínimo de 0 e o máximo atingindo o valor de 0,97 nas duas condições. Esses valores não permitem inferir algo relevante, visto que as médias são também altas. Geramos um *boxplot* para melhor visualizar a possível existência de *outliers*, que não ocorreu (GRÁFICO 10).

Gráfico 10 – *Boxplot* da variável *neutral* em ambas as variáveis preditoras

Fonte: Elaborado pelo autor

Pelo gráfico 10, percebemos que não há a presença de valores discrepantes nesta medida, em ambas as condições. A análise fica facilitada com a ausência dos *outliers*. Podemos também notar que a média e mediana da condição SSOM são superiores às da condição CSOM. A distribuição da condição SSOM é maior, enquanto que a da condição CSOM ainda apresenta distribuição menor e o terceiro quartil é inferior ao seu correspondente em SSOM. Esses dados nos permitem perceber que há diferenças significativas entre os valores nas duas condições.

No gráfico 11, podemos perceber uma assimetria positiva nos dados de *neutral*, novamente não nos permitindo ver uma distribuição normal. Apesar de haver valores maiores entre 0 e 0,13, a maior concentração de valores, como apontado pela média e mediana, está entre 0,25 e 0,5.

Gráfico 11 – Histograma da variável *neutral* totalizada

Fonte: Elaborado pelo autor

Esperávamos que o neutro fosse apresentar grandes valores nas duas condições, visto que as emoções partem do neutro e, na maior parte do tempo, estamos nesse estado (EKMAN e FRIESEN, 1978; DUTTA e BARMAN, 2020). A análise dos dados do neutro nos permitiu inferir que a condição CSOM provocou mais emoções em geral que a SSOM. A variável preditora SSOM, portanto, deve manter o espectador numa falta de emoções, enquanto a outra condição aparenta suscitar mais emoções, sejam elas quais forem, por passarem menos tempo no neutro. Novamente, teremos conclusões mais afirmativas quando apresentarmos toda a nossa análise estatística.

4.1.6 Sad

A variável *sad*, tristeza, apresentou a maior média de valores de todas as variáveis resposta. As médias foram 0,414, na CSOM; e 0,401, na SSOM (TABELA 9). A diferença numérica entre as médias é a segunda maior registrada: 0,13. A partir das médias, inferimos que a tristeza foi a emoção mais presente em ambos os grupos. Há alta confiabilidade das médias, visto que o erro-padrão é baixo.

Tabela 9 - Dados descritivos da variável resposta *sad*

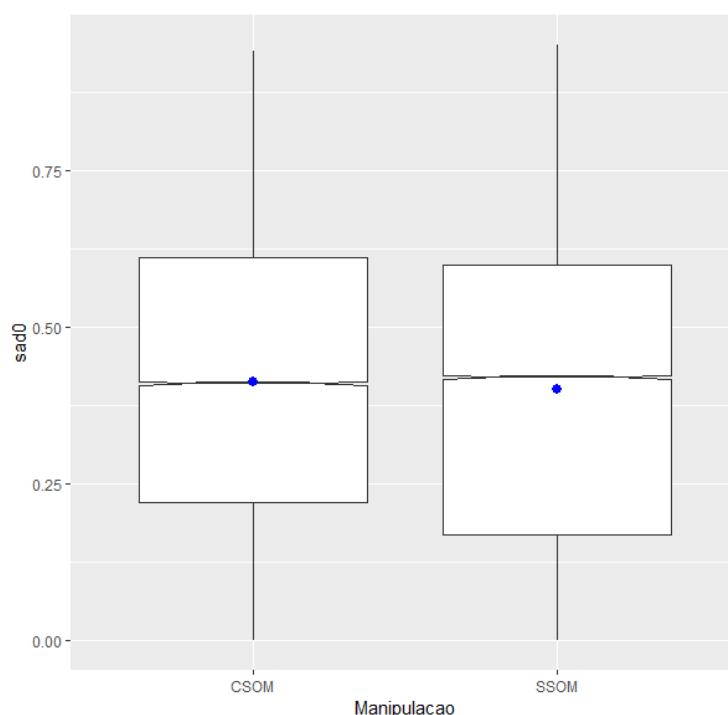
MANIPULAÇÃO	MEAN	MEDIAN	SD	SE	MIN	MAX
CSOM	0.414	0.41	0.239	0.0009	0	0.94
SSOM	0.401	0.42	0.253	0.0009	0	0.95

Fonte: Elaborado pelo autor

A Tabela 9 traz medianas marginalmente diferentes: 0,41 e 0,42, nas condições CSOM e SSOM, respectivamente. A mediana indica que a tristeza foi predominante nas duas condições durante os estímulos, sendo pouco superior na condição SSOM. As medianas, nesse valor, permitem-nos inferir que, durante a maior parte do tempo de coleta, os participantes aparentavam estar tristes.

O desvio-padrão entre as medidas apresenta variação de 0,239, para a condição CSOM; e 0,253, para a SSOM, sugerindo que a condição CSOM tem sua distribuição mais homogênea que a outra condição. O mínimo de 0 e o máximo atingindo o valor de 0,94, na condição CSOM, e 0,95, na SSOM, não indica uma diferença significativa, a não ser quando aliadas com as médias e medianas. Sendo assim, podemos inferir a ausência total de valores atípicos. Geramos um *boxplot*, para melhor visualizar a distribuição das médias e valores obtidos na emoção tristeza (GRÁFICO 12).

Gráfico 12 – *Boxplot* da variável *sad* em ambas as variáveis preditoras



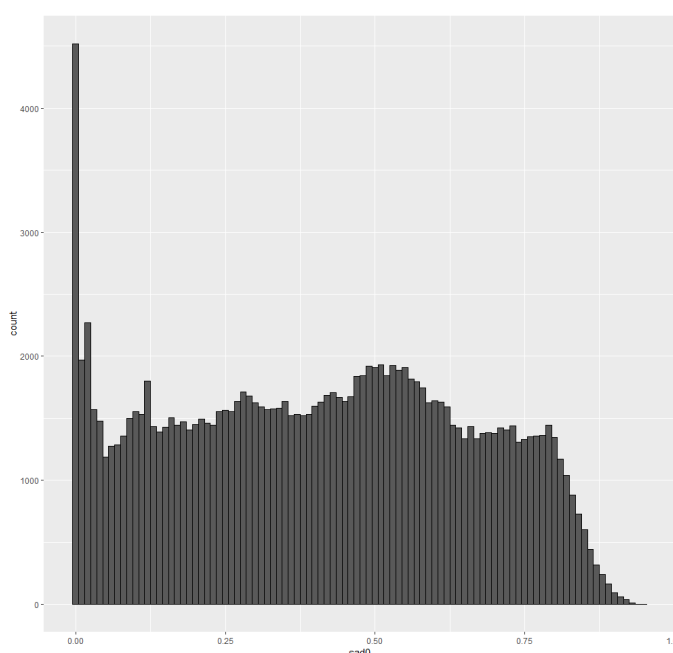
Fonte: Elaborado pelo autor

No gráfico 12, percebemos que realmente não há a presença de *outliers* nesta medida, em ambas as condições. A mediana da condição SSOM é marginalmente superior ao da condição CSOM, como anteriormente citado. A distribuição mais dispersa da condição SSOM força seu primeiro quartil mais para baixo, enquanto a da condição CSOM ainda

apresenta distribuição menos dispersa e terceiro quartil superior ao seu correspondente em SSOM. Esses dados nos fazem perceber que há diferenças significativas entre os valores das duas condições.

Pelo gráfico de histograma (Gráfico 13), podemos ver um grande platô nos valores centrais e um pico no valor de 0,0. Esse pico isolado pode significar que houve algum problema na captação dos dados ou que simplesmente ele foi o mais presente. Ainda assim, consideramos os dados válidos.

Gráfico 13 – Histograma da variável *sad* totalizada



Fonte: Elaborado pelo autor

Esperávamos que a tristeza fosse apresentar valores relevantes, mas não tanto quanto os dados revelaram. *Sad* nos permite inferir que a condição CSOM provocou mais tristeza que a SSOM. Seria, então, um ponto favorável à nossa hipótese. Contudo, conclusões mais categóricas, novamente, só poderão ser feitas quando apresentarmos toda a nossa análise estatística.

4.1.7 Surprise

A variável resposta *surprise*, surpresa, também apresentou valores tendendo a zero no IIE, mas com maior dispersão. Esperávamos valores mais relevantes desta emoção, visto que os trechos apresentam momentos que suscitam a surpresa, entretanto, não foi o que

se mostrou quando olhamos para os dados brutos. A média na variável preditora CSOM foi de 0,006; e em SSOM, 0,004, havendo, assim, diferença pouco considerável entre as duas. A partir das médias, inferimos que a surpresa foi uma emoção pouco sentida em ambos os grupos (Tabela 10). Os baixos valores do erro-padrão também indicam alta confiabilidade das médias.

Tabela 10 - Dados descritivos da variável resposta *surprise*.

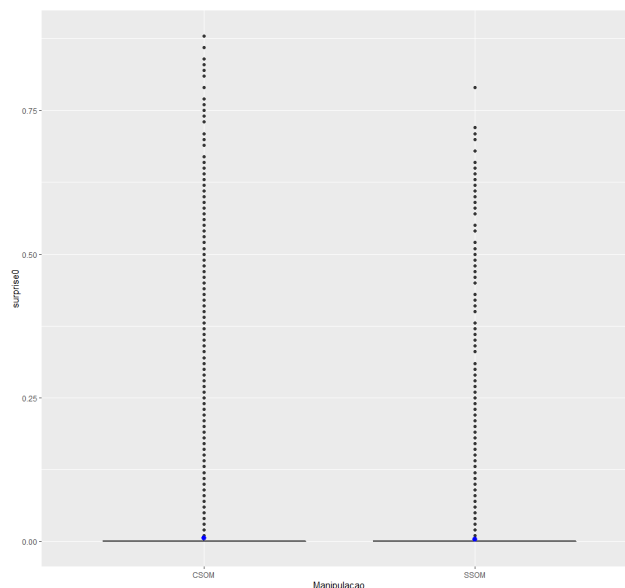
MANIPULAÇÃO	MEAN	MEDIAN	SD	SE	MIN	MAX
CSOM	0.006	0	0.035	0.0001	0	0.88
SSOM	0.004	0	0.023	0.0001	0	0.79

Fonte: Elaborado pelo autor

A Tabela 10 nos mostra valor de 0 em ambas as medianas. A mediana zerada indica que essa emoção não foi expressa pelos participantes.

O desvio-padrão entre as medidas apresenta variação de 0,035, para a condição CSOM; e 0,023, para a SSOM, indicando que a condição SSOM tem sua distribuição mais homogênea. O mínimo de 0 e o máximo atingindo o valor de 0,88 e 0,79 em CSOM e SSOM, respectivamente, permite-nos imaginar que surpresa teve picos da sua expressão. Esses picos podem ter sido bem breves, já que médias e medianas são bem próximas de zero, podendo significar, então, presença de muitos *outliers*. O gráfico 14 apresenta o gráfico *boxplot* com a dispersão dos valores e os *outliers*.

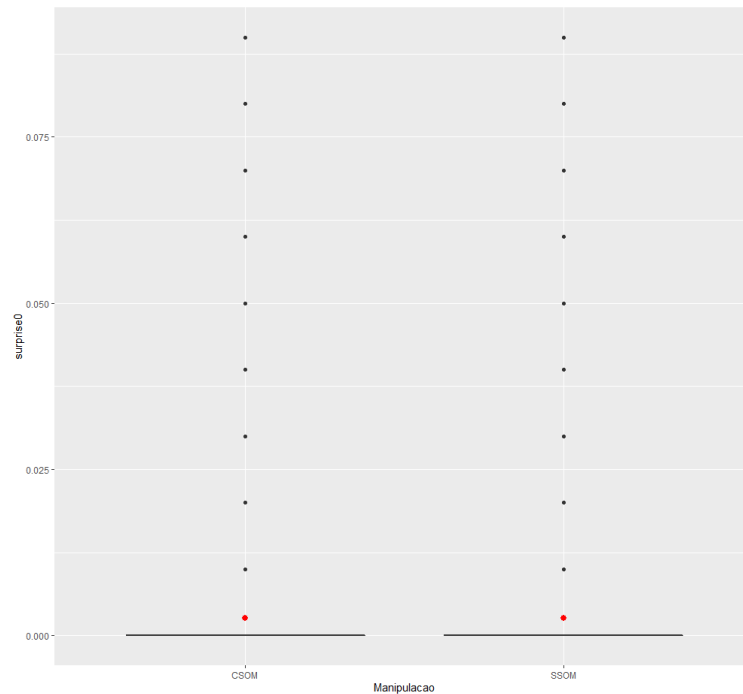
Gráfico 14 – *Boxplot* da variável *surprise* em ambas as variáveis predictoras



Fonte: Elaborado pelo autor

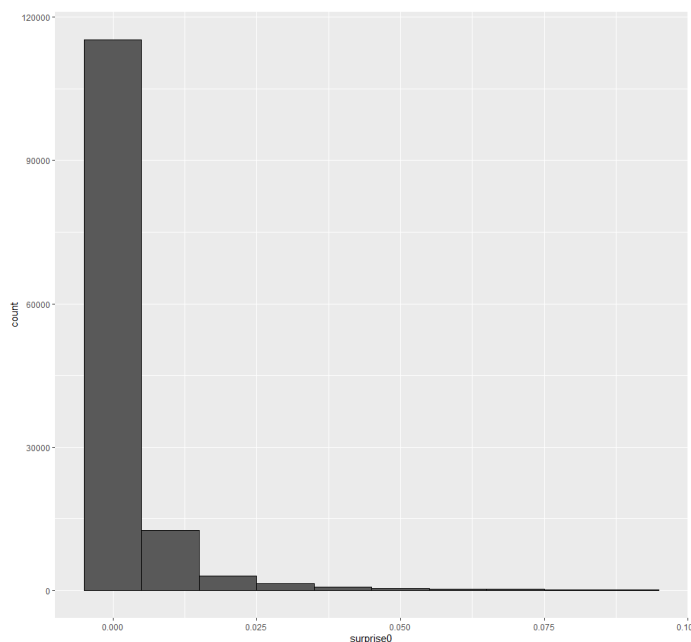
O *boxplot* com os valores da emoção surpresa (GRÁFICO 14) mostra grande presença de *outliers* para ambas as condições. Para a filtragem deles, estabelecemos a linha de corte de valores atípicos em 0,09 (GRÁFICO 15). Mesmo assim, ainda é possível observar a presença desses dados. Esses *outliers* foram mantidos a fim de não zerar os dados da variável *surprise*, caso filtrássemos completamente os valores atípicos. Visualmente, a média na condição CSOM é marginalmente superior à da condição SSOM. Ambas as distribuições das caixas se situam em zero, sendo a média superior a esse valor por causa dos *outliers*.

Gráfico 15 – *Boxplot* da variável *surprise* em ambas as variáveis preditoras sem muitos *outliers*



Fonte: Elaborado pelo autor

Pelo histograma (GRÁFICO 16), podemos visualizar uma assimetria positiva nos dados da emoção surpresa. A maioria dos valores, quase 120 mil de um total de 135 mil, entretanto, é de zero.

Gráfico 16 – Histograma da variável *surprise* totalizada

Fonte: Elaborado pelo autor

Esperávamos que a emoção surpresa fosse apresentar valores de maior relevância, já que os trechos de ação buscam provocar surpresas ao acontecer um fato inesperado. Os dados são tão baixos que inferimos que, apesar de haver alguma diferença entre as médias das condições, surpresa não é uma emoção presente nos trechos selecionados. Na próxima seção, verificaremos se há alguma relevância estatística entre os valores apresentados sobre as emoções em ambas as variáveis preditoras.

4.2 Resultados da Análise de Variância (ANOVA) e o Modelo Linear Misto (LMM)

O teste de análise de variância (ANOVA) busca verificar se há diferença entre as hipóteses. Para tal, ela compara médias das amostras e procura significância, respeitando o valor de confiança, que, em nosso caso, é de 5%, ou $p \leq 0,05$. As duas hipóteses que são comparadas são: hipótese nula, H_0 , em que não há relação entre as amostras (μ) e elas são encaradas como iguais ($H_0: \mu_1 = \mu_2$ se $p \geq 0,05$); hipótese alternativa, H_1 , em que há diferença significativa entre as amostras ($H_1: \mu_1 \neq \mu_2$ se $p \leq 0,05$).

Em nosso caso, a H_0 afirmaria que os valores da variável resposta *angry* não se 0, diferem em ambas as variáveis preditoras – CSOM e SSOM –, ou seja, $\mu_1 = \mu_2$ se $p \leq 0,05$. A H_1 seria o contrário: há diferença dos valores de *angry* entre as variáveis preditoras, sendo, então, $\mu_1 \neq \mu_2$ se $p \geq 0,05$.

No RStudio, rodamos o seguinte código para executarmos a ANOVA: “*anovaangry = aov(angry0 ~ Manipulacao, data = dados)*”. Para relatar os dados gerados pela ANOVA, utilizamos a mesma notação que é o padrão para esses casos: f (graus de liberdade, tamanho da amostra) = valor estatístico do teste, p = valor de p . No nosso mesmo exemplo da emoção raiva, obteríamos o seguinte resultado: $f(1, 128.770) = 408,7, p < 0,0001$.

Executamos, também, um teste Eta ao quadrado, para sabermos o tamanho do efeito. O código para execução no RStudio foi o seguinte: “*eta_squared(anovaangry)*”. O teste de Eta ao quadrado serve para medirmos o tamanho do efeito observado no teste ANOVA. Ele é importante para complementar os dados gerados pela análise de variância, ao indicar o efeito da diferença das variáveis. A medida do Eta ao quadrado estabelece alguns intervalos de tamanho do efeito relatado (η^2). Se $\eta^2 = 0,01$, o tamanho do efeito é pequeno; se $\eta^2 = 0,06$, é médio; e se $\eta^2 = 0,14$, é grande. Fazendo uso do mesmo exemplo, reportamos os dados da seguinte maneira: $f(1, 128.770) = 408,7, p < 0,0001, \eta^2 < 0.003$.

Para complementarmos a análise decorrente do teste ANOVA, fizemos uso do LMM, pois ele nos possibilita uma previsão de valores para populações mais amplas, já que ele executa diversas estimativas numéricas para as variáveis resposta em relação às preditoras (BATTES, 2020). O LMM é, ainda, um modelo baseado em contrastes, sobre o que se quer comparar. Em nossa pesquisa o contraste é um só: a condição CSOM e SSOM. Temos, ainda, dois efeitos aleatórios presentes em nossa pesquisa: “Participante” e “Trecho”. Efeitos aleatórios são todos aqueles que são particulares para cada indivíduo, e sua demarcação é fundamental para executarmos o modelo estatístico.

Para rodarmos o LMM, fizemos uso do seguinte código-exemplo: “*lmm_angry <- lmer(data = dados, angry0 ~ Manipulacao + (1|Participante) + (1|Trecho))*”. O modelo retorna diversos dados, inclusive testes-t, que iremos apresentar em tabelas. Os resultados que interessam para nossa pesquisa são as chamadas variáveis fixas, mais precisamente, a estimativa dos coeficientes *intercept* (β_0), linear, e *slope* (β_1), angular. O coeficiente linear indica a estimativa de valor quando a variável CSOM for zero. O coeficiente angular, por sua vez, demonstra o crescimento, ou diminuição, dos resultados na condição CSOM em relação ao ponto de partida da SSOM, ou seja, o tamanho do efeito.

A seguir, apresentaremos os resultados das variáveis resposta, divididos por suas respectivas subseções. Apresentaremos, primeiro, o resultado do teste ANOVA e Eta ao quadrado e, depois, passaremos para os resultados do LMM.

4.2.1 Angry

Para a variável resposta *angry*, obtivemos $f(1, 128.770) = 408,7$, $p < 0,0001$, $\eta^2 = 0,003$ (TABELA 11). O baixo valor-p descarta nossa hipótese nula – não há diferença entre as medidas – e nos indica que há diferença entre as condições CSOM e SSOM (hipótese alternativa). Mesmo obtendo um valor muito baixo de η^2 , acreditamos que o tamanho do efeito é risório.

Tabela 11 – ANOVA da variável resposta *angry*

	DF	SUM SQ	MEAN SQ	F VALUE	PR (>F)	N ²
MANIPULACAO	1	1.9	1.9190	408.7	<2e-16	0.00316
RESIDUALS	128770	604.7	0.0047			

Fonte: Elaborado pelo autor

O LMM apresentou dados convergentes aos resultados da ANOVA. A tabela 12 mostra os efeitos randômicos da variável resposta *angry*, enquanto que a tabela 13 mostra os efeitos fixos. Os dados da tabela 12 são de pouca relevância para nossa análise, mas veiculamo-los para termos transparência e para que possamos observar todos os dados gerados pelo nosso modelo.

Tabela 12 – Estimativas de efeito randômico da variável resposta *angry*

RANDOM EFFECTS			
GROUPS	NAME	VARIANCE	STD. DEV.
PARTICIPANTE	(Intercept)	0.0026687	0.05166
TRECHO	(Intercept)	0.0001424	0.01193
RESIDUAL		0.0024162	0.04915
NUMBER OF OBS: 128.772; GROUPS: - ; PARTICIPANTE: 16; TRECHO: 4			

Fonte: Elaborado pelo autor

A partir dos dados da Tabela 13, temos $\beta_0 = 0,0087$ e $\beta_1 = 0,0008$. Esses valores baixos nos indicam que o tamanho do efeito da condição CSOM é positivo. O valor-p do teste-t abaixo de 0,0001 nos indica que o modelo é funcional.

Tabela 13 – Estimativas de efeito fixo da variável resposta *angry*

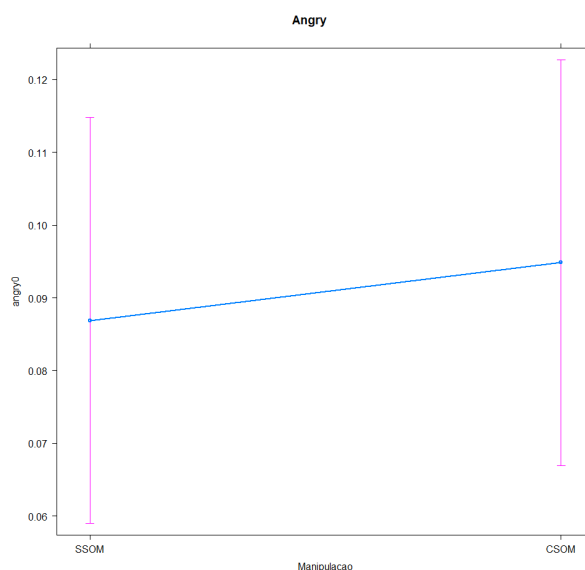
FIXED EFFECTS					
	Estimate	Std. Error	df	t value	Pr(> t)
(INTERCEPT)	8.689e-02	1.423e-02	1.660e+01	6.107	1.29e-05

MANIPULACAOC SOM	7.923e-03	2.837e-04	1.288e+05	27.922	< 2e-16
-------------------------	-----------	-----------	-----------	--------	---------

Fonte: Elaborado pelo autor

A fim de observar o efeito, geramos um gráfico que relaciona as estimativas da condição SSOM para CSOM (GRÁFICO 17).

Gráfico 17 – Plot da variável *angry* de SSOM para CSOM



Fonte: Elaborado pelo autor

O gráfico apresenta, visualmente, variação considerável na emoção raiva, quando se traduz os efeitos sonoros. Esse resultado confirma nossas hipóteses de que o uso da tradução sonora impacta positivamente nas emoções do público surdo e que o não uso impacta negativamente. Entretanto, os dados numéricos do teste ANOVA e do LMM apresentam um efeito marginal, mas considerado por nós relevante.

4.2.2 Disgust

Para a variável resposta *disgust*, obtivemos $f(1, 135.539) = 12,6$, $p < 0,001$, $\eta^2 < 0,0001$ (TABELA 14). O p-valor irrelevante descarta a H_0 e abre espaço para analisarmos nossa H_1 . O baixo valor-p nos indica que há diferença entre as condições CSOM e SSOM. Mesmo obtendo um valor muito baixo de η^2 , acreditamos novamente que o tamanho do efeito é risório.

Tabela 14 – ANOVA da variável resposta *disgust*

	DF	SUM SQ	MEAN SQ	F VALUE	PR (>F)	N ²
MANIPULACAO	1	0.00001	7.546e-06	12.6	0.000387	9.292276e-05
RESIDUALS	135539	0.08120	5.990e-07			

Fonte: Elaborado pelo autor

Há, mais uma vez, a corroboração dos dados do teste ANOVA pelo LMM. A tabela 15 mostra os efeitos randômicos da variável resposta *disgust*, enquanto que a tabela 16 mostra os efeitos fixos.

Tabela 15 – Estimativas de efeito randômico da variável resposta *disgust*

RANDOM EFFECTS			
GROUPS	NAME	VARIANCE	STD. DEV.
PARTICIPANTE	(Intercept)	6.216e-09	7.884e-05
TRECHO	(Intercept)	1.522e-09	3.902e-05
RESIDUAL		5.922e-07	7.695e-04
NUMBER OF OBS: 135.541; GROUPS: - ; PARTICIPANTE: 16; TRECHO: 4			

Fonte: Elaborado pelo autor

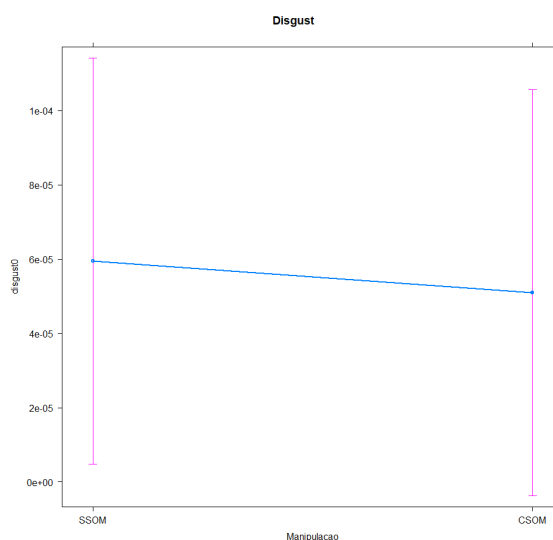
Na tabela 16, temos $\beta_0 = 5.945e-05$ e $\beta_1 = -8.521e-06$. No caso da variável *disgust*, a estimativa do efeito foi negativo, ou seja, a condição CSOM apresenta queda na presença desta emoção em relação a não legendagem dos efeitos sonoros.

Tabela 16 – Estimativas de efeito fixo da variável resposta *disgust*

FIXED EFFECTS					
	Estimate	Std. Error	df	t value	Pr(> t)
(INTERCEPT)	5.945e-05	2.791e-05	8.725e+00	2.130	0.0629
MANIPULACAOCSOM	-8.521e-06	4.315e-06	1.340e+05	-1.975	0.0483

Fonte: Elaborado pelo autor

Plotamos novamente um gráfico de relação entre as condições SSOM e CSOM (GRÁFICO 18).

Gráfico 18 – *Plot* da variável *disgust* de SSOM para CSOM

Fonte: Elaborado pelo autor

Pelo gráfico 18, percebemos a variação negativa na emoção nojo, quando se legenda os efeitos sonoros. Esse resultado vai ao encontro de nossa hipótese; o uso da tradução sonora impacta positivamente nas emoções do público surdo, consideramos, portanto, ela relevante.

4.2.3 Fear

Para a variável resposta *fear*, obtivemos $f(1, 111.104) = 534,4$, $p < 0,001$, $\eta^2 = 0,004$ (TABELA 17). Com o baixo valor de p , podemos descartar nossa H_0 e analisarmos nossa H_1 . O teste ANOVA aponta para diferença estatística entre as condições CSOM e SSOM. Novamente, η^2 apresenta um valor muito baixo (0,004), mas que ainda assim nos faz acreditar que o tamanho do efeito é relevante.

Tabela 17 – ANOVA da variável resposta *fear*

	DF	SUM SQ	MEAN SQ	F VALUE	PR (>F)	N ²
MANIPULACAO	1	0.20	0.20112	453.4	<2e-16	0.004063938
RESIDUALS	111104	49.29	0.00044			

Fonte: Elaborado pelo autor

O LMM novamente confirma os resultados obtidos pelo teste ANOVA, incluindo o tamanho do efeito muito pequeno. A tabela 18 mostra os efeitos randômicos da variável resposta *fear*, enquanto que a tabela 19 mostra os efeitos fixos.

Tabela 18 – Estimativas de efeito randômico da variável resposta *fear*

RANDOM EFFECTS			
GROUPS	NAME	VARIANCE	STD. DEV.
PARTICIPANTE	(Intercept)	3.086e-04	0.017568
TRECHO	(Intercept)	2.571e-06	0.001603
RESIDUAL		2.299e-04	0.015163
NUMBER OF OBS: 111.106; GROUPS: - ; PARTICIPANTE: 16; TRECHO: 4			

Fonte: Elaborado pelo autor

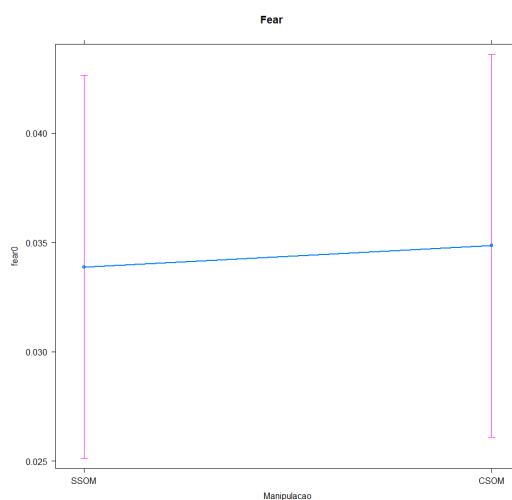
Em relação a *fear*, LMM retornou que $\beta_0 = 3,387e-02$ e que $\beta_1 = 9,756e-04$ (TABELA 19). A variável *fear* apresentou resultado positivo em seu coeficiente *slope*, indicando que nosso modelo aponta para diferença real entre as condições CSOM e SSOM. Apesar disso, o valor do coeficiente foi pequeno ($\beta_1 < 0,0001$). O tamanho do efeito, então, foi considerado relevante, mesmo apresentando pequena diferença entre as condições.

Tabela 19 – Estimativas de efeito fixo da variável resposta *fear*

FIXED EFFECTS					
	Estimate	Std. Error	Df	t value	Pr(> t)
(INTERCEPT)	3.387e-02	4.465e-03	1.581e+01	7.586	1.18e-06
MANIPULACAOCOM	9.756e-04	9.692e-05	1.108e+05	10.067	< 2e-16

Fonte: Elaborado pelo autor

A relação entre as condições SSOM e CSOM teve novamente um gráfico plotado, para sua melhor visualização (GRÁFICO 19).

Gráfico 19 – Plot da variável *fear* de SSOM para CSOM

Fonte: Elaborado pelo autor

É possível perceber que há uma variação na emoção medo, quando se legenda os efeitos sonoros. Esse resultado reforça nossas hipóteses, apesar de ser um efeito estatístico reduzido.

4.2.4 Happy

Para a variável *happy*, obtivemos $f(1, 123.427) = 20,27$, $p < 0,001$, $\eta^2 = 0,0001$ (TABELA 20). O valor-p baixo descarta nossa H_0 e permite analisarmos nossa H_1 . O teste ANOVA aponta mais uma vez para diferença estatística entre as condições CSOM e SSOM. O η^2 apresenta um valor muito baixo (0,0001), indicando que o tamanho do efeito é mínimo.

Tabela 20 – ANOVA da variável resposta *happy*

	DF	SUM SQ	MEAN SQ	F VALUE	PR (>F)	N ²
MANIPULACAO	1	0.002	0.0016375	20.27	6.72e-06	0.000164215
RESIDUALS	123427	9.970	0.0000808			

Fonte: Elaborado pelo autor

Os resultados obtidos pela ANOVA, incluindo o tamanho do efeito, são corroborados pelo LMM. A tabela 21 mostra os efeitos randômicos da variável resposta *happy*, enquanto que a tabela 22 mostra os efeitos fixos.

Tabela 21 – Estimativas de efeito randômico da variável resposta *happy*

RANDOM EFFECTS			
GROUPS	NAME	VARIANCE	STD. DEV.
Participante	(Intercept)	6.386e-05	0.0079912
Trecho	(Intercept)	5.180e-07	0.0007197
Residual		5.175e-05	0.0071939
NUMBER OF OBS: 123.429; GROUPS: - ; PARTICIPANTE: 16; TRECHO: 4			

Fonte: Elaborado pelo autor

Em relação a *happy*, LMM retornou que $\beta_0 = 5,403e-03$ e que $\beta_1 = 2,941e-04$ (TABELA 22). A variável *happy* apresentou resultado positivo em seu β_1 , indicando diferença estatística entre as condições CSOM e SSOM. O valor desse coeficiente, entretanto, foi muito

pequeno novamente ($\beta_1 < 0,0001$), mas ainda assim, acreditamos que o tamanho do efeito é relevante.

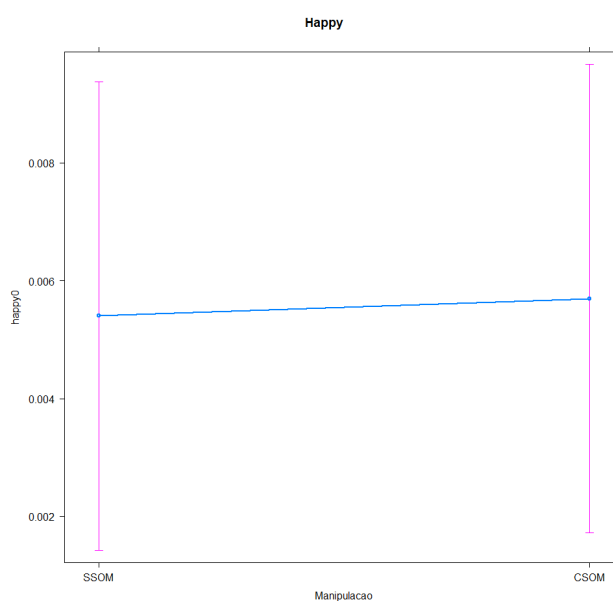
Tabela 22 – Estimativas de efeito fixo da variável resposta *happy*

FIXED EFFECTS					
	Estimate	Std. Error	df	t value	Pr(> t)
(INTERCEPT)	5.403e-03	2.030e-03	1.583e+01	2.661	0.0172
MANIPULACAOCSOM	2.941e-04	4.213e-05	1.234e+05	6.979	2.98e-12

Fonte: Elaborado pelo autor

Plotamos novamente um gráfico para melhor visualizar a relação entre SSOM e CSOM na variável *happy* (GRÁFICO 20).

Gráfico 20 – *Plot* da variável *happy* de SSOM para CSOM



Fonte: Elaborado pelo autor

A variação entre as condições é existente, mas é mínima (GRÁFICO 20). Esse resultado reforça nossas hipóteses, apesar de não ser esperada diferenças significativas para o gênero do filme por nós utilizado.

4.2.5 Neutral

Para a variável resposta *neutral*, obtivemos $f(1, 135.788) = 269,5$, $p < 0,001$, $\eta^2 = 0,002$ (TABELA 23). O valor de p mínimo descarta outra vez nossa H_0 e abre margem

novamente para analisarmos nossa H₁. O teste ANOVA aponta para a existência de diferença estatística entre as condições CSOM e SSOM. Assim, como todas as demais variáveis resposta, η^2 apresenta um valor muito baixo (0,002), indicando que o tamanho do efeito é mínimo.

Tabela 23 – ANOVA da variável resposta *neutral*

	DF	SUM SQ	MEAN SQ	F VALUE	PR (>F)	N ²
MANIPULACAO	1	19	18.656	269.5	<2e-16	0.001980591
RESIDUALS	135788	9401	0.069			

Fonte: Elaborado pelo autor

O LMM confirma mais uma vez os resultados da ANOVA, incluindo o tamanho do efeito. A tabela 24 mostra os efeitos randômicos da variável resposta *neutral*, enquanto que a tabela 25 mostra os efeitos fixos.

Tabela 24 – Estimativas de efeito randômico da variável resposta *neutral*

RANDOM EFFECTS			
GROUPS	NAME	VARIANCE	STD. DEV.
PARTICIPANTE	(Intercept)	0.045712	0.21380
TRECHO	(Intercept)	0.002089	0.04571
RESIDUAL		0.024016	0.15497
NUMBER OF OBS: 135.790; GROUPS: - ; PARTICIPANTE: 16; TRECHO: 4			

Fonte: Elaborado pelo autor

Em relação a *neutral*, LMM retornou que $\beta_0 = 0,039$ e que $\beta_1 = - 0,002$ (TABELA 25). A variável *neutral* apresentou resultado negativo em seu coeficiente *slope*, indicando que nosso modelo aponta para uma redução de valores, quando comparamos as condições SSOM e CSOM. O efeito negativo sugere que os participantes na condição CSOM passaram menos tempo em estado neutro do que sua contraparte da condição SSOM. Apesar disso, o valor do coeficiente foi mínimo novamente ($\beta_1 = - 0,002$). O tamanho do efeito é, então, considerado por nós relevante, mesmo apresentando diferença significativa.

Tabela 25 – Estimativas de efeito fixo da variável resposta *neutral*

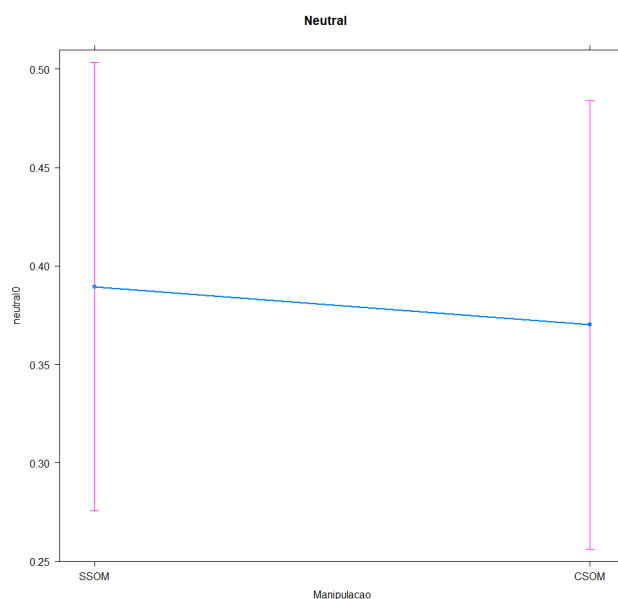
FIXED EFFECTS					
	Estimate	Std. Error	df	t value	Pr(> t)
(INTERCEPT)	3.895e-01	5.814e-02	1.690e+01	6.70	3.85e-06

MANIPULACAOC SOM	-1.948e-02	8.687e-04	1.358e+05	-22.43	< 2e-16
-------------------------	------------	-----------	-----------	--------	---------

Fonte: Elaborado pelo autor

A relação entre as condições SSOM e CSOM teve novamente um gráfico plotado, para sua melhor visualização (GRÁFICO 21).

Gráfico 21 – *Plot* da variável *neutral* de SSOM para CSOM



Fonte: Elaborado pelo autor

Ao se observar o gráfico 21, percebemos que há uma variação negativa no estado neutro, quando se legenda os efeitos sonoros. Esse resultado corrobora nossas hipóteses, visto que imaginávamos que o participante, quando não é apresentado a legendas com a tradução dos efeitos sonoros, tende a passar mais tempo em estado neutro do que sua contraparte. Mesmo o tamanho do efeito sendo muito pequeno, este resultado contribuiu a favor da tradução dos efeitos sonoros.

4.2.6 *Sad*

Para a variável resposta *sad*, obtivemos $f(1, 135.788) = 94,68$, $p < 0,001$, $\eta^2 = 0,0007$ (TABELA 26). O baixo valor de p descarta nossa H_0 e abre margem para analisarmos nossa H_1 . O teste ANOVA apresenta novamente a existência de diferenças estatísticas entre as condições SSOM e CSOM. Novamente, η^2 apresenta um valor muito baixo (0,0007), indicando que o tamanho do efeito é muito pequeno.

Tabela 26 – ANOVA da variável resposta *sad*

	DF	SUM SQ	MEAN SQ	F VALUE	PR (>F)	N ²
Manipulacao	1	6	5.721	94.68	<2e-16	0.000696782
Residuals	135788	8205	0.060			

Fonte: Elaborado pelo autor

Os resultados da ANOVA são novamente corroborados pelo LMM, incluindo o tamanho do efeito. A tabela 27 mostra os efeitos randômicos da variável resposta *sad*, enquanto que a tabela 28 mostra os efeitos fixos.

Tabela 27 – Estimativas de efeito randômico da variável resposta *sad*

RANDOM EFFECTS			
GROUPS	NAME	VARIANCE	STD. DEV.
Participante	(Intercept)	0.042010	0.20496
Trecho	(Intercept)	0.001432	0.03784
Residual		0.018886	0.13743
NUMBER OF OBS: 135.790; GROUPS: - ; PARTICIPANTE: 16; TRECHO: 4			

Fonte: Elaborado pelo autor

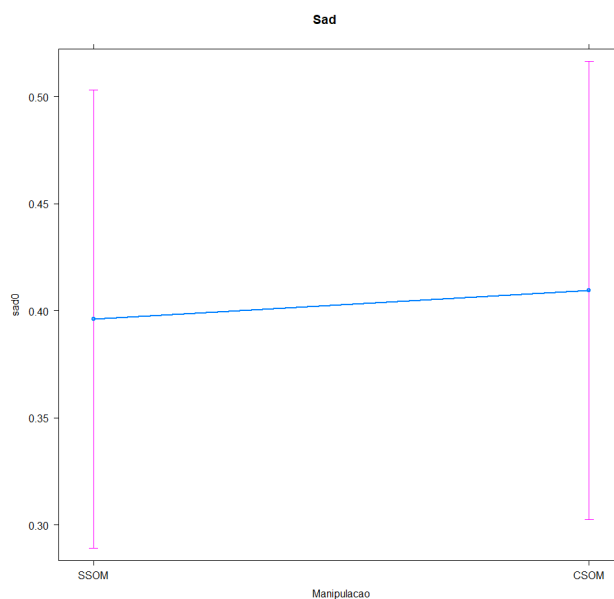
Em relação a *sad*, LMM retornou $\beta_0 = 0,04$ e que $\beta_1 = 0,001$ (TABELA 28). A variável *sad* apresentou resultado positivo em β_1 , indicando que nosso modelo aponta para o aumento de valores quando se parte da condição SSOM em direção a CSOM. O efeito positivo indica que os participantes na condição CSOM sentiram mais a emoção tristeza do que sua contraparte. Contudo, o valor do coeficiente foi muito pequeno ($\beta_1 = 0,001$). Consideramos o tamanho do efeito, portanto, relevante e apresentando diferença significativa.

Tabela 28 – Estimativas de efeito fixo da variável resposta *sad*

FIXED EFFECTS					
	Estimate	Std. Error	df	t value	Pr(> t)
(Intercept)	3.962e-01	5.463e-02	1.701e+01	7.252	1.35e-06
ManipulacaoCSOM	1.334e-02	7.704e-04	1.358e+05	17.311	< 2e-16

Fonte: Elaborado pelo autor

A relação entre as condições SSOM e CSOM teve novamente um gráfico plotado para sua melhor visualização (GRÁFICO 22).

Gráfico 22 – Plot da variável *sad* de SSOM para CSOM

Fonte: Elaborado pelo autor

Observando o gráfico 22, podemos verificar variação positiva na emoção tristeza, quando se legenda os efeitos sonoros. Esse resultado corrobora nossas hipóteses, visto que imaginávamos que a emoção tristeza fosse marginalmente superior na condição CSOM, mesmo o tamanho do efeito sendo muito pequeno.

4.2.7 Surprise

Para a variável resposta *surprise*, obtivemos $f(1, 125.477) = 7,256$, $p < 0,01$, $\eta^2 = 5,78e-05$ (TABELA 29). O baixo valor de p descarta H_0 e nos permite analisar H_1 . O teste ANOVA apresenta novamente a existência de diferenças estatísticas entre as condições SSOM e CSOM. O η^2 apresenta um valor muito pequeno ($5,78e-05$), indicando que o tamanho do efeito é mínimo.

Tabela 29 – ANOVA da variável resposta *surprise*

	DF	SUM SQ	MEAN SQ	F VALUE	PR (>F)	N ²
Manipulacao	1	0.002	0.0015824	7.256	0.00707	5.782487e-05
Residuals	125477	27.364	0.0002181			

Fonte: Elaborado pelo autor

Novamente, os testes ANOVA e LMM estão em consonância, validando os resultados. A tabela 30 mostra os efeitos randômicos da variável resposta *surprise*, enquanto que a tabela 31 mostra os efeitos fixos.

Tabela 30 – Estimativas de efeito randômico da variável resposta *surprise*

RANDOM EFFECTS			
GROUPS	NAME	VARIANCE	STD. DEV.
Participante	(Intercept)	1.944e-05	0.004409
Trecho	(Intercept)	1.731e-06	0.001316
Residual		5.548e-05	0.007449
NUMBER OF OBS: 134.594; GROUPS: - ; PARTICIPANTE: 16; TRECHO: 4			

Fonte: Elaborado pelo autor

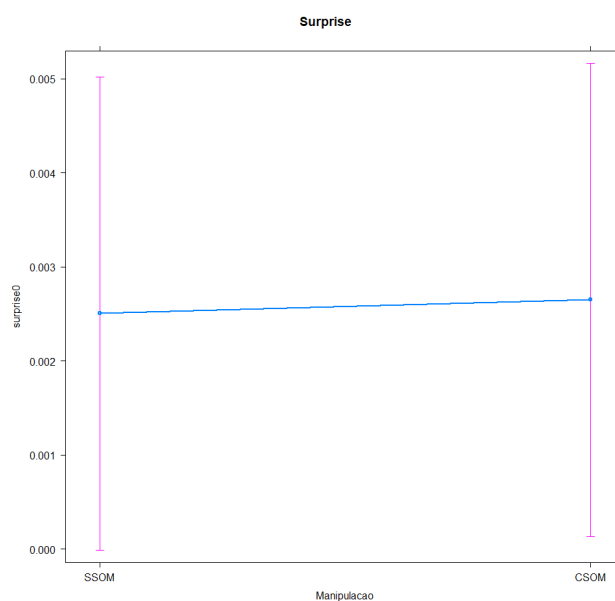
Em relação a *sad*, LMM retornou $\beta_0 = 0,0003$ e que $\beta_1 = 1.454e-04$ (TABELA 31). A variável *surprise* apresentou resultado positivo em β_1 , indicando que nosso modelo aponta para o aumento de valores quando se parte da condição SSOM para CSOM. O efeito, mínimo, indica que os participantes na condição CSOM sentiram marginalmente mais a emoção surpresa do que sua contraparte, tendo coeficiente, como já dito, mínimo ($\beta_1 < 0,0001$). Consideramos, portanto, o tamanho do efeito relevante, além do fato do modelo apresentar diferença significativa entre as condições.

Tabela 31 – Estimativas de efeito fixo da variável resposta *surprise*

FIXED EFFECTS					
	Estimate	Std. Error	df	t value	Pr(> t)
(Intercept)	2.503e-03	1.284e-03	1.488e+01	1.949	0.070380
ManipulacaoCSOM	1.454e-04	4.189e-05	1.346e+05	3.470	0.000521

Fonte: Elaborado pelo autor

Plotamos um gráfico de relação entre SSOM e CSOM (GRÁFICO 23), para observarmos melhor sua variação.

Gráfico 23 – Plot da variável *surprise* de SSOM para CSOM

Fonte: Elaborado pelo autor

Ao analisarmos o gráfico 23, verificamos que há uma pequena variação positiva na emoção surpresa, quando se compara SSOM com CSOM. Esse resultado contribuiu para a confirmação de nossas hipóteses, visto que imaginávamos que a emoção tristeza saísse pouco superior na condição CSOM que na SSOM. O tamanho do efeito, no entanto, é marginal, mas relevante.

Na próxima seção, apresentaremos a interpretação dos dados, bem como os resultados obtidos pela análise estatística. Apresentaremos, também, nossas considerações sobre o que pode ter influenciado os resultados deste estudo.

4.3 Discussão dos resultados

Para coletar os dados em nosso estudo, manipulamos quatro trechos com menos de 3 minutos do filme *Dunkirk* (2017) e os expusemos a um total de 16 participantes surdos. Para reconhecer e quantificar as emoções de cada participante, utilizamos como ferramenta o *software* de reconhecimento de expressões faciais *Visage*. Este programa gerou dados quantitativos de cada uma das seis emoções básicas universais postuladas por Ekman (1970), incluindo, também, o estado neutro.

Nossa coleta gerou em torno de 135 mil linhas de dados, que foram analisadas primeiramente descritivamente e depois estatisticamente pelos testes ANOVA, Eta ao quadrado e o LMM.

A análise descritiva retornou dados muito semelhantes entre nossas condições, com e sem a tradução de efeitos sonoros. As médias da variável *angry*, por exemplo, foram as mesmas: 0,106, fato que se repetiu com suas medianas: 0,08. Esses resultados foram proporcionalmente semelhantes para as outras variáveis, com exceção de *neutral* e *sad*, que apresentaram discrepâncias maiores.

Por se tratar de um filme de ação, esperávamos que as seguintes emoções tivessem seu efeito demonstrado na maior parte do tempo nas emoções dos espectadores: raiva, medo, tristeza e surpresa. Entretanto, não foi o constatado pelos dados das médias e nem das análises estatísticas.

O que observamos é que, na condição CSOM, nojo teve a menor média, tendendo a zero (0,0002); seguida por surpresa (0,006); alegria (0,055); medo (0,058); raiva (0,11); neutro (0,36) seguido, finalmente, pela tristeza (0,41). Já na condição SSOM, as médias foram, da menor para a maior: nojo (8,96e-05); surpresa (0,004); alegria (0,049); medo (0,054); raiva (0,106); neutro (0,385) e, por fim, a tristeza (0,401). Percebemos que, em ambas as condições, a sequência das variáveis por ordem de grandeza é a mesma: nojo, surpresa, alegria, medo, raiva, neutro e tristeza.

Para complementar o estudo descritivo, executamos análises estatísticas e, a partir delas, buscamos compreender se a tradução dos efeitos sonoros influenciou ou não a manifestação das emoções. Os resultados do teste ANOVA, Eta ao quadrado e o LMM indicam haver diferença estatística entre as condições em todas as variáveis, não sendo, entretanto, significativa, como apresentaremos a seguir.

Angry apresentou os seguintes resultados: $f(1, 128.770) = 408,7$; $p < 0,0001$, $\eta^2 < 0,003$; $\beta_1 = 0,0008$. O baixo valor de p indica haver diferença entre as condições, corroborando nossas hipóteses. Mesmo os valores de η^2 e β_1 sendo muito baixos, a diferença entre condições é estatisticamente relevante. *Disgust* apresentou os valores: $f(1, 135.539) = 12,6$; $p < 0,001$; $\eta^2 < 0,0001$; $\beta_1 = -8.521e-06$. O valor de p aponta para diferença entre as condições, no entanto, o baixo valor de η^2 e o valor negativo de β_1 apontam para um decréscimo de intensidade na condição CSOM e, novamente, diferença estatisticamente risória. Os valores de *fear* foram: $f(1, 111.104) = 534,4$; $p < 0,001$; $\eta^2 = 0,004$; $\beta_1 = 9,756e-04$. Temos novamente um valor- p indicando diferença entre as condições e valores relevantes para η^2 e β_1 . *Happy* teve o seguinte resultado: $f(1, 123.427) = 20,27$; $p < 0,001$; $\eta^2 = 0,0001$; $\beta_1 = 2,941e-04$. Novamente, valor- p , η^2 e β_1 baixos, indicando haver diferença entre as condições. Para *neutral* tivemos: $f(1, 135.788) = 269,5$; $p < 0,001$; $\eta^2 = 0,002$; $\beta_1 = -0,002$. Os

valores de p , η^2 e β_1 foram baixos novamente, indicando diferença. *Sad* apresentou os seguintes resultados: $f(1, 135.788) = 94,68$; $p < 0,001$; $\eta^2 = 0,0007$; $\beta_1 = 0,001$. Os valores de p , η^2 e β_1 apontaram novamente para diferença entre condições e estatisticamente risória. *Surprise* retornou os valores: $f(1, 125.477) = 7,256$; $p < 0,01$; $\eta^2 = 5,78e-05$; $\beta_1 = 1.454e-04$. Mais outra vez o valor de p , de η^2 e β_1 apontam para diferenças entre condições outra vez, e sendo estatisticamente relevante.

Nossos dados, apesar de evidenciarem a diferença entre as condições e corroborarem a adoção de modelos de tradução de efeitos sonoros, como o de Nascimento (2018), mostraram tamanhos de efeito marginais. Os β_1 e η^2 foram inferiores, em todas as variáveis, a 0,01 ou -0,01. Apesar de valores baixos, acreditamos que há diferença significativa entre as condições, visto que a variação nos músculos faciais para expressar uma emoção é realmente muito sutil. Os gráficos e as análises estatísticas, portanto, apresentaram uma tendência a uma evolução dos valores, quando a tradução dos efeitos sonoros está presente. Portanto, consideramos ter indícios do impacto positivo da tradução dos efeitos sonoros no público surdo.

Antes de explorarmos os motivos pelos quais acreditamos que as diferenças estatísticas foram relevantes, temos que discutir brevemente a aplicação do questionário de autopercepção das emoções (DES), que não entrou nesta análise. A aplicação do DES nos possibilita novas análises dos seus dados e o cruzamento deles com as informações geradas pelo *Visage*. Este novo *set* de informações nos permitirá identificar semelhanças e incongruências entre os instrumentos de coletas, e se os valores aqui apresentados são corroborados pela percepção do próprio participante. O cruzamento dos resultados de ambos os instrumentos pode nos ajudar a compreender melhor o impacto da tradução dos efeitos sonoros no público surdo. Não foi possível realizar a análise desse questionário, devido ao fator do tempo que dedicamos à análise dos dados do *software Visage*. Assim, nosso próximo passo será fazer um estudo detalhado sobre os dados do questionário DES e compará-lo aos resultados dos dados gerados pelo *Visage*.

Sobre os dados do *Visage*, para explicar melhor as diferenças entre as condições existentes (CSOM e SSOM) elaboramos algumas hipóteses que buscam compreender essas questões. Em relação ao alto valor de *sad*, variáveis externas podem ter influenciado esse alto valor, como: mau humor prévio dos participantes, cansaço, falta de vontade de participar da pesquisa e a leitura dos rostos em estado neutro pelo programa como se estivessem exprimindo tristeza. Acreditamos que este último problema seja o mais provável, visto que

alguns participantes apresentaram semblante abatido, mesmo reportando estarem bem, sendo esse o seu natural, ou neutro. Para contrapor essas variáveis externas, precisaríamos, portanto, avaliar também o questionário de autopercepção (DES), a fim de eliminar possíveis influências externas aos dados.

O decréscimo da emoção nojo pode ter acontecido devido a expressões de não entendimento por parte dos espectadores enquanto estavam realizando as tarefas. Na condição SSOM, por exemplo, a falta de legendas sobre os efeitos sonoros pode ter causado dúvida e, por consequente, expressões faciais que o programa interpretou como nojo. Esse acontecimento pode ter ocorrido, também, em valores atribuídos a outras emoções, como raiva e tristeza.

As variáveis dependentes medo e surpresa também apresentaram valores abaixo do esperado em ambas as condições, o que pode ter sido devido à pouca duração dos trechos. Os trechos, com duração abaixo dos 3 minutos, mesmo que editados, não permitem que o espectador conheça toda a trama dramática do filme e possam se sentir devidamente impactadas por suas consequências em tela. Essa falta de envolvimento com o texto dramático pode influenciar negativamente quando analisamos somente excertos de um filme e não ele por completo.

Em relação à detecção de efeitos diminutos, acreditamos que nossa ferramenta e nossa abordagem têm impacto direto nesses resultados. É possível que a ferramenta usada, *Visage*, não tenha tanto aprimoramento para perceber todas as sutilezas dos músculos da face quando expressa determinada emoção, gerando, portanto, dados diminutos da emoção expressa. Apesar de haver outras bibliotecas de detecção facial, utilizamos uma comprovadamente eficiente, MTCNN, mas há outras que reportam valores mais acurados, como a *DeepFace*, mas que não conseguimos implementar a tempo para a coleta de dados.

Há, também, a questão da controvérsia a respeito da eficácia do modelo de expressão das emoções pela face de Ekman (1970, 2003), de que ela não seria eficaz para representar todas as emoções que o ser humano é capaz de sentir (BARRETT, ADOLPHS, *et al.*, 2019). Segundo Barret et al. (2019), a face apresenta expressa emoções experienciadas pelo ser humano, mas nem todas são necessariamente expressas na face, apesar sentidas. Sendo assim, os modelos que se baseiam em análise facial para descobrir a manifestação de emoções podem não reportar com detalhes todas as emoções.

Apesar de termos algumas observações críticas sobre os dados coletados pelo *Visage*, podemos observar que os resultados sobre as variáveis preditoras intensificam a

importância da tradução dos efeitos sonoros para o público surdo. Uma vez que, mesmo tendo diferenças bem pequenas, os dados convergem para a influência dessa tradução intersemiótica na emoção dos espectadores. Outros métodos para a coleta de dados a respeito das emoções se fazem necessárias para validarem esses dados. O uso de entrevistas, questionários e até mesmo um eletroencefalograma também podem ser consideradas ferramentas úteis para complementar os dados analisados.

5 CONSIDERAÇÕES FINAIS

Nossa pesquisa nasceu no contexto do grupo Legendagem e Audiodescrição da UECE, que busca estudar a acessibilidade nos meios audiovisuais. Ela foi concebida a partir de questionamentos da funcionalidade do modelo técnico de tradução de efeitos sonoros proposto por Nascimento (2018). A autora elaborou parâmetros de tradução para que os legendistas pudessem padronizar suas legendas, a fim de proporcionar uma recepção que atenda às necessidades dos espectadores surdos e ensurdecidos. Nosso intento foi, portanto, prosseguir com os estudos do grupo, verificando se a tradução dos efeitos sonoros provoca impacto em relação às emoções sobre o gênero ação.

Nosso principal questionamento de pesquisa foi: “A tradução dos efeitos sonoros na LSE de um filme de ação influencia na emoção dos espectadores surdos durante a recepção?”. Logo, nossos objetivos foram: 1) Investigar a intensidade das emoções na recepção da LSE por meio do modelo de tradução dos efeitos sonoros proposto por Nascimento (2018); 2) Mensurar a diferença sobre as emoções mediante o uso e não-uso da tradução de efeitos sonoros na LSE.

Para cumprir os objetivos deste estudo, desenhamos uma pesquisa experimental que avaliasse a influência tradução dos efeitos sonoros durante a recepção de um filme de ação. Para executar tal tarefa, selecionamos quatro trechos de ação do filme *Dunkirk* (2017) e manipulamos sua LSE, seguindo o modelo de tradução de efeitos sonoros proposto por Nascimento (2018).

Após a exposição dos participantes aos trechos em ambas as condições, alimentamos os arquivos de vídeo com suas faces no programa *Visage* e obtivemos os dados quantitativos, que nos permitiram averiguar a intensidade das emoções expressas por suas faces.

Os dados nos mostraram que a emoção mais sentida foi a tristeza, além do estado neutro da face, dominando, as duas somadas, 77% do tempo das emoções expressas pelos participantes na condição CSOM. As menos sentidas foram nojo e surpresa, aproximadamente, 2% do total. As demais emoções, raiva, alegria e medo, apresentaram valores medianos, sendo expressas durante 21% do tempo total de exposição.

Para mensurarmos a influência da tradução dos efeitos sonoros sobre a emoção dos participantes, desenvolvemos um estudo estatístico dos dados, fazendo uso de três técnicas distintas: análise descritiva, uso da Análise de Variância e do Modelo Linear Misto.

As três técnicas estatísticas apresentaram os mesmos resultados: há diferença na expressão das emoções entre surdos que assistem a trechos de um filme de ação com legendas com e sem tradução de efeitos sonoros. Há indícios estatísticos de que a intensidade das emoções é aumentada quando os surdos são expostos às legendas com tradução dos efeitos sonoros e diminuída quando assistem sem essa tradução. Os dados estatísticos, entretanto, demonstram que essa variação de intensidade é muito pequena, mas existente, sendo um indicativo de variação encontrado em nossa pesquisa. Esse resultado existente, mas marginal, portanto, rejeita a nossa hipótese nula e ainda nos estimula a buscar métodos mais sofisticados, para melhor compreender o impacto da tradução dos efeitos sonoros sobre as emoções.

Os resultados deste estudo sobre as reações da face revelaram que há indícios de que a tradução dos efeitos sonoros na LSE impacte positivamente nas emoções de espectadores surdos, quando estes assistem a um filme de ação. No entanto, esta pesquisa focou apenas no gênero ação, especificamente, o subgênero guerra, em que há forte carga informativa no canal sonoro, devido à inclusão de efeitos sonoros. Consideramos, portanto, que são necessários estudos sobre outros gêneros audiovisuais, como comédia, documentário ou drama. Também é importante que pesquisas derivativas se utilizem de participantes ouvintes para se estabelecer uma medida de grupo controle.

Nosso resultado não corrobora o de Aleksandrowicz (2019), que propõe que provavelmente a música não tenha tanto impacto no público quanto se imagina. Apesar dos instrumentos de pesquisa diferentes do nosso, os resultados de certa forma apontam para um efeito positivo sobre as emoções dos espectadores surdos, por meio da tradução de efeitos sonoros.

Esperamos que este estudo favoreça a área de formação de legendistas da LSE, principalmente, sobre a tradução de efeitos sonoros. Há ainda lacunas que precisamos compreender melhor, para contribuirmos com dados mais consolidados, como: aumento no número de participantes; estímulos maiores, possibilitando maior tempo de monitoramento da face; e maior quantidade de estímulos que gerem mais possibilidades de comparação e desenvolvimento de outros instrumentos de observação, como a análise de questionários de autopercepção e testes com eletroencefalograma.

Este estudo experimental buscou verificar a relevância da tradução dos efeitos sonoros na LSE pela recepção de espectadores surdos. Logo, acreditamos que o investimento em pesquisas experimentais, com instrumentos que possam medir as emoções, pode contribuir

para que possamos formar legendistas que forneçam acessibilidade ao público surdo. Assim, nossa principal contribuição, neste estudo, foi complementar as pesquisas desenvolvidas pela TAVa e possibilitar uma LSE que atenda às necessidades de recepção dos espectadores surdos e ensurdecidos.

REFERÊNCIAS

- ADERALDO, M. F. **Proposta de parâmetros descritivos para audiodescrição à luz da interface revisitada entre tradução audiovisual acessível e semiótica social – multimodalidade**. Universidade Federal de Minas Gerais. Belo Horizonte. 2014.
- ALEKSANDROWICZ, P. Can subtitles for the deaf and hard-of-hearing convey the emotions of film music? A reception study. **Perspectives**, 21 Junho 2019. 58-72.
- ALLEYNE, M. Sound Technology: Sounds Reel: Tracking the Cultural History of Film Sound Technology. In: HARPER, G.; DOUGHTY, R.; EISENTRAUT, J. **Sound and Music in Film and Visual Media: A Critical Overview**. 1. ed. [S.l.]: Bloomsbury Academics, 2009. Cap. 1, p. 15-41.
- ARAÚJO, V. L. S. Closed subtitling in Brazil. In: ORERO, P. **Topics in Audiovisual Translation**. 1. ed. Amsterdam: John Benjamins Publishing Company, v. 1, 2004. Cap. 20, p. 199-212.
- ARAÚJO, V. L. S. A Legendagem para Surdos no Brasil. In: LIMA, P. L. C.; ARAÚJO, A. D. **Questões de Linguística Aplicada**. Fortaleza: EdUECE, 2005. p. 163-188.
- ARAÚJO, V. L. S. Subtitling for the deaf and hard-of-hearing in Brazil In: Media for All. In: JORGE DÍAZ CINTAS, P. O. A. R. **Media for All: Subtitling for the Deaf, Audio Description and Sign Language**. Nova Jersey: Brill | Rodopi, v. 30, 2007. Cap. 5, p. 99-107.
- ARAÚJO, V. L. S. Por um modelo de legendagem para Surdos no Brasil. **Tradução & Comunicação - Revista Brasileira de Tradutores**, São Paulo, n. 17, p. 59-76, 2008.
- ARAÚJO, V. L. S. **Legendagem para surdos: em busca de um modelo para o Brasil**. CNPq. Fortaleza, p. 44. 2012. (306948/2008-7).
- ARAÚJO, V. L. S. O processo de legendagem no Brasil. **Revista do GELNE**, v. 4, n. 1, p. 1-6, 2016.
- ARAÚJO, V. L. S.; ASSIS, I. A. P. A segmentação linguística na legendagem para surdos e ensurdecidos (LSE) de ‘Amor Eterno Amor’: uma análise baseada em corpus. **Letras & Letras**, p. 156-184, dez. 2014.
- ARAÚJO, V. L. S.; NASCIMENTO, A. K. P. Investigando parâmetros de legendas para Surdos e Ensurdecidos no Brasil. **Tradução em Revista**, Rio de Janeiro, v. 2, p. 1-18, 2011.
- ARAÚJO, V. L. S.; VIEIRA, P. A.; MONTEIRO, S. M. M. Legendagem para surdos e ensurdecidos (LSE): Um estudo de recepção com surdos da região Sudeste. **Tradterm**, São Paulo, v. 22, n. 1, p. 283-302, dez. 2013.

ASSIS, Í. A. P. D. **Legendagem para surdos e ensurdecidos (LSE): análise baseada em corpus da segmentação linguística em ‘amor eterno amor’**. Fortaleza: Universidade Estadual do Ceará. p. 108. 2016.

BARRETT, L. F. *et al.* Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements. **Psychological Science in the Public Interest**, p. 68, 2019.

BATTES, D. M. **Lme4: Mixed-effects modeling with R**. Wisconsin: Springer, 2020. 145 p.

BORDWELL, D. **Sobre a História do Estilo Cinematográfico**. São Paulo: Ed. UNICAMP, 2013.

BORDWELL, D.; THOMPSON, K. **A arte do cinema: uma introdução**. 10. ed. São Paulo: EdUSP, v. 1, 2014. ISBN 9788531414541.

COMPARATO, D. **Da Criação ao Roteiro: Teoria e Prática**. 10. ed. São Paulo: Summus Editorial, 2018.

COUSINS, M. **The Story of Film**. [S.l.]: Pavilion, 2013.

D'YDEWALLE, G. *et al.* Watching subtitled television: automated reading behavior. **Communication research**, p. 650-666, 1991.

DARWIN, C. **The Expression of the Emotions in Man and Animals**. [S.l.]: Cambridge University Press, 2009.

DÍAZ CINTAS, J. Audiovisual translation in the third millennium. *In*: ANDERMAN, G.; ROGERS, M. **Translation Today: Trends and Perspectives**. [S.l.]: Multilingual Matters, 2003. Cap. 14, p. 192-204.

DÍAZ CINTAS, J.; REMAEL, A. **Audiovisual Translation: Subtitling**. Manchester: St. Jerome Pub., v. 11, 2007.

DÍAZ CINTAS, J.; REMAEL, A. **Subtitling: Concepts and Practices**. [S.l.]: ROUTLEDGE, 2020.

DUTTA, P.; BARMAN, A. **Human emotion recognition from face images**. [S.l.]: SPRINGER, 2020.

D'YDEWALLE, G. Reading a message when the same message is available auditorily. **Eye Movements: From Physiology to Cognition**, Amsterdam, 1987. 313-321.

D'YDEWALLE, G.; VAN RENSBERGEN, J.; POLLE, J. Reading a message when the same message is available auditorily in another language: the case of subtitling. *In*: J.K. O'REGAN,

A. L.-S. **Eye Movements from Physiology to Cognition**. Amsterdã e Nova Iorque: Elsevier, 1987. p. 313-321. ISBN 9780444701138.

EKMAN, P. Universal Facial Expressions of Emotion. **California Mental Health Research Digest**, 1970. p. 151-158.

EKMAN, P. **Emotions Revealed**. Nova Iorque: Henry Holt and Company, 2003.

EKMAN, P.; CORDARO, D. What is Meant by Calling Emotions Basic. **Emotion Review**, 2011. 364-370.

EKMAN, P.; FRIESEN, W. **Facial Action Coding System: A Technique for the Measurement of Facial Movement**. Palo Alto: Consulting Psychologists Press, 1978.

FRANCO, E. P. C.; ARAÚJO, V. L. S. Reading Television Checking Deaf People's Reactions to Closed Subtitling in Fortaleza, Brazil. **The Translator**, 2003. 249-267.

GAMBIER, Y. The position of audiovisual translation studies. *In*: MILLÁN, C.; BARTRINA, F. **The routledge handbook of translation studies**. LONDON: ROUTLEDGE, 2012. p. 45-59.

GITHUB. The State of the Octoverse. **GitHub**, 2020. Disponível em: <https://octoverse.github.com>. Acesso em: 21 jan. 2021.

GOODFELLOW, I. J. *et al.* Challenges in Representation in Learning: A report on three machine learning contests. **arXiv**, 2013.

HARPER, G.; DOUGHTLY, R.; EISENTRAUT, J. **Sound and Music in Film and Visual Media: A Critical Overview**. [S.l.]: Bloomsbury Academic, 2009.

IZARD, C. E. **HUMAN EMOTIONS**. New York: Springer Science+Business, 1977.

JAKOBSON, R. **Linguística e Comunicação**. São Paulo: Cultrix, 1995. 63-86 p.

JIMÉNEZ HURTADO, C.; SEIBEL, C. **Traducción Accesible: Narratología y Semántica de la Audiodescripción**. El español, lengua de traducción para la cooperación y el diálogo. Madrid: ESLEtRA. 2007. p. 451-468.

KARAMITROGLOU, F. A Proposed Set of Subtitling Standards in Europe. **Translation Journal**, 1998. 1-15. Disponível em: http://www.sub2learn.ie/downloads/karamitroglou_fotiosa_proposed_set_of_subtitling_standards_in_europe.pdf. Acesso em: 11 ago. 2020.

LASTRA, J. **Sound Technology and the American Cinema: Perception, Representation, Modernity**. Nova Iorque: Columbia University Press, 2000.

LEHMAN, F. **Hollywood Harmony: Musical Wonder and the Sound of Cinema**. Oxford: Oxford University Press, 2018.

MONTEIRO, S. M. M.; DANTAS, J. F. Tradução audiovisual acessível (TAVa): a segmentação linguística na Legendagem para Surdos e Ensurdecidos (LSE) da campanha política na televisão em Fortaleza. **Trabalhos em Linguística Aplicada**, Campinas, 2017. 527-560.

NASCIMENTO, A. K. P. **Linguística de corpus e legendagem para surdos e ensurdecidos (LSE): uma análise baseada em corpus da tradução de efeitos sonoros na legendagem de filmes brasileiros em DVD**. Universidade do Estado do Ceará. Fortaleza, p. 109. 2013.

NASCIMENTO, A. K. P. **Convencionalidade nas legendas de efeitos sonoros na legendagem para surdos e ensurdecidos (LSE)**. Universidade de São Paulo. São Paulo, p. 241. 2018.

PEREGO, E. What would we read best? Hypotheses and suggestions for the location of line. **The Sign Language Translator and Interpreter**, Manchester, 2008. 35-63.

REID, H. Literature on the screen: subtitle translation for public broadcasting. In: WESTERWEEL, B. **Something understood: studies in Anglo-Dutch literary translation**. Amsterdã: Rodopi, 1990. p. 97-107.

ROMERO-FRESCO, P. Standing on quicksand: hearing viewers' comprehension and reading. In: DÍAZ-CINTAS, J.; MATAMALA, A.; NEVES, J. **New Insights into Audiovisual Translation and Media Accessibility: Media for all 2**. Amsterdam: Rodopi, 2010. p. 175-195.

SALDANHA, G.; O'BRIEN, S. **Research Methodologies in Translation Studies**. Nova Iorque: Routledge, 2014.

SELVATICI, C. **Closed Caption: Conquistas e Questões**. Pontifícia Universidade Católica do Rio De Janeiro - PUC-RIO. Rio de Janeiro, p. 87. 2010.

SOUZA, Gabrielle Tomaz de. **O efeito da segmentação e da velocidade de legendas no processamento linguístico de surdos e ouvintes em vídeos de campanhas políticas: um estudo de rastreamento ocular**. 2020. 82f. – Dissertação (Mestrado) – Universidade Federal do Ceará, Programa de Pós-graduação em Estudos da Tradução, Fortaleza (CE), 2020. Disponível em: <http://www.repositorio.ufc.br/handle/riufc/53279>. Acesso em: 25 mar. 2021.

VIEIRA, Héliida Arrais Costa. O padrão de trajetória visual para o reconhecimento de expressões faciais. 2015. 127 f., il. Tese (Doutorado em Ciências do Comportamento) - Universidade de Brasília, Brasília, 2015. Disponível em: <https://repositorio.unb.br/handle/10482/19707>. Acesso em: 20 mar. 2021

VIEIRA, PATRÍCIA ARAÚJO. A influência da segmentação e da velocidade na recepção de legendas para surdos e ensurdecidos (LSE). 2016. 247 f. Tese (Doutorado) - Universidade

Estadual do Ceará, 2016. Disponível em: <http://siduece.uece.br/siduece/trabalhoAcademicoPublico.jsf?id=82208>. Acesso em: 15 mar. 2021.

VIEIRA, P. A.; SANTIAGO ARAUJO, V. L. A influência da segmentação linguística na recepção de legendas para surdos e ensurdecidos (LSE) em documentários televisivos. **Domínios de Linguagem**, [S. l.], v. 11, n. 5, p. 1797-1824, 2017. DOI: 10.14393/DL32-v11n5a2017-22. Disponível em: <https://seer.ufu.br/index.php/dominiosdelinguagem/article/view/36913>. Acesso em: 15 mar. 2021.

VIEIRA, P. A.; ASSIS, I. A. P.; ARAÚJO, V. L. S. Tradução audiovisual: estudos sobre a leitura de legendas para surdos e ensurdecidos. **Cadernos de Tradução**, Florianópolis, v. 40, p. 97-124, 2020.

ZHANG, K. *et al.* Joint face detection and alignment using multitask cascaded convolutional networks. **IEEE Signal Processing Letters**, p. 1499-1503, 2016.

APÊNDICE

**APÊNDICE A – QUESTIONÁRIO DEMOGRÁFICO PARA TRAÇAR O PERFIL
DOS PARTICIPANTES**

QUESTIONÁRIO DEMOGRÁFICO	
Dados Pessoais	
Nome:	_____
E-mail:	_____
Gênero:	<input type="checkbox"/> feminino <input type="checkbox"/> masculino <input type="checkbox"/> outro
Idade:	____ anos.
Data de nascimento:	____/____/____:
Estado de nascimento:	Ceará <input type="checkbox"/> Outro <input type="checkbox"/> _____
Se marcou outro, desde quando (que ano) mora no Ceará?	_____.
Formação:	<input type="checkbox"/> Pós-graduado. <input type="checkbox"/> Ensino superior completo. <input type="checkbox"/> Ensino superior incompleto. <input type="checkbox"/> Ensino Médio <input type="checkbox"/> Ensino Fundamental <input type="checkbox"/> Outro _____
Renda mensal Familiar:	<input type="checkbox"/> Até 1 Salário-Mínimo <input type="checkbox"/> De 1 a 2 Salários-Mínimos <input type="checkbox"/> De 2 a 3 Salários Mínimos <input type="checkbox"/> De 3 a 5 Salários Mínimos <input type="checkbox"/> De 5 a 10 Salários Mínimos <input type="checkbox"/> Mais de 10 Salários Mínimos
Dados Demográficos	
1. Quando você ficou surdo?	<input type="checkbox"/> Eu nasci surdo. <input type="checkbox"/> Com ____ meses. <input type="checkbox"/> Com ____ anos.
2. Qual seu grau de surdez?	<input type="checkbox"/> Surdez leve <input type="checkbox"/> Surdez moderada <input type="checkbox"/> Surdez profunda <input type="checkbox"/> _____
3. Onde você aprendeu Libras pela primeira vez?	<input type="checkbox"/> Na família <input type="checkbox"/> Na escola <input type="checkbox"/> Na igreja <input type="checkbox"/> Na internet <input type="checkbox"/> _____

4. Com quantos anos você aprendeu Libras?

- Entre 0 e 5 anos
- Entre 5 e 11 anos
- Depois dos 11 anos
- Com _____ anos

5. Que nota você daria para seu conhecimento em Libras? Sendo 0 para nenhum e 5 para muito alto:

- 0 – Nenhum
- 1 – Muito Baixo
- 2 – Baixo
- 3 – Regular
- 4 – Alto
- 5 – Muito Alto

6. Que nota você daria para sua leitura em Língua Portuguesa? Sendo 0 para nenhum e 5 para muito alto:

- 0 – Nenhum
- 1 – Muito Baixo
- 2 – Baixo
- 3 – Regular
- 4 – Alto
- 5 – Muito Alto

7. Você tem alguma dificuldade com a visão?

- Sim Não Não sei

8. Se você marcou sim, qual?

- Miopia – Grau: OD (olho direito)_____ OE (olho esquerdo)_____
- Astigmatismo – Grau: OD_____ OE _____
- Hipermetropia – Grau: OD_____ OE _____
- Catarata
- Daltonismo
- Estrabismo
- Glaucoma
- Presbiopia

Outra(s) _____

9. Você usa óculos ou lentes de contato para a dificuldade com a visão?

Sim Não

10. Você é uma pessoa:

Destra (escreve com a mão direita)

Canhota (escreve com a mão esquerda)

11. Quando você assiste a TV, você costuma ver filmes ou programas legendados?

Sim Não

12. Em caso de resposta afirmativa à questão 11, com que frequência?

Todos os dias

3 vezes por semana

1 vez por semana

2 vezes por mês

Raramente

Não tenho uma frequência fixa

Não sei

13. Você costumava ir ao cinema antes da pandemia?

Sim Não

14. Em caso de resposta afirmativa à questão 13, com que frequência?

Mais de uma vez por semana

Uma vez por semana

Uma vez por mês

Uma vez a cada seis meses

Raramente

Não tenho uma frequência fixa

Não sei

15. Em caso de resposta afirmativa à questão 13, que gênero de filme você prefere ver no cinema? Marque até 3 opções.

Ação

Aventura

Comédia

- Documentário Drama Fantasia
 Ficção Científica Musical Policial
 Romance Suspense Terror

16. Você costuma assistir filmes e/ou séries por meio de serviços de *streaming*? Por exemplo: *Netflix, Amazon Prime, Globoplay*.

- Sim Não

17. Em caso de resposta afirmativa à questão 16, com que frequência?

- Todos os dias
 Uma a três vezes por semana
 Uma vez a cada 15 dias
 Uma vez por mês
 Uma vez a cada seis meses
 Raramente
 Não tenho uma frequência fixa
 Não sei

18. Em caso de resposta afirmativa à questão 17, que gênero de filme você prefere ver em casa?

- Ação Aventura Comédia
 Documentário Drama Fantasia
 Ficção Científica Musical Policial
 Romance Suspense Terror

19. Você costuma consumir entretenimento de vídeo com legendas para surdos? Por exemplo: legenda com tradução de efeitos sonoros e indicação de falante.

- Sim Não

20. Você prefere legendas para surdos ou legendas para ouvintes?

- Legendas para Surdos Legendas para Ouvintes