



**UNIVERSIDADE ESTADUAL DO CEARÁ**  
**CENTRO DE HUMANIDADES**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA APLICADA**  
**DOUTORADO EM LINGUÍSTICA APLICADA**

**ÍTALO ALVES PINTO DE ASSIS**

**A INFLUÊNCIA DO NÚMERO DE LINHAS E DA VELOCIDADE NO  
PROCESSAMENTO DE LEGENDAS POR SURDOS E OUVINTES: UM ESTUDO  
EXPERIMENTAL COM RASTREADOR OCULAR**

**FORTALEZA – CEARÁ**

**2021**

ÍTALO ALVES PINTO DE ASSIS

A INFLUÊNCIA DO NÚMERO DE LINHAS E DA VELOCIDADE NO  
PROCESSAMENTO DE LEGENDAS POR SURDOS E OUVINTES: UM ESTUDO  
EXPERIMENTAL COM RASTREADOR OCULAR

Tese apresentada ao Curso de Doutorado em Linguística Aplicada do Programa de Pós-Graduação em Linguística Aplicada do Centro de Humanidades da Universidade Estadual do Ceará, como requisito parcial à obtenção do título de doutor em Linguística Aplicada.  
Área de concentração: Linguagem e Interação.

Orientadora: Profa. Dra. Vera Lúcia Santiago Araújo.  
Coorientadora: Profa. Dra. Patrícia Araújo Vieira.

FORTALEZA – CEARÁ

2021

**Dados Internacionais de Catalogação na Publicação**

**Universidade Estadual do Ceará**

**Sistema de Bibliotecas**

Assis, Ítalo Alves Pinto de.

A influência do número de linhas e da velocidade no processamento de legendas por surdos e ouvintes: um estudo experimental com rastreador ocular [recurso eletrônico] / Ítalo Alves Pinto de Assis. - 2021

Um arquivo no formato PDF do trabalho acadêmico com 255 folhas.

Tese (doutorado) - Universidade Estadual do Ceará, Centro de Humanidades, Programa de Pós-Graduação em Linguística Aplicada, Fortaleza, 2021.

Área de concentração: Linguagem e Interação.

Orientação: Prof.<sup>a</sup> Ph.D. Vera Lúcia Santiago Araújo.

Coorientação: Prof.<sup>a</sup> Dra. Patrícia Araújo Vieira.

1. Tradução Audiovisual Acessível. 2. Legendagem para Surdos e Ensurdidos. 3. Rastreamento Ocular. 4. Velocidade da Legenda. 5. Número de Linhas. I. Título.

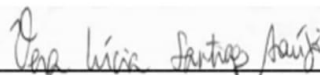
ÍTALO ALVES PINTO DE ASSIS

A INFLUÊNCIA DO NÚMERO DE LINHAS E DA VELOCIDADE NO  
PROCESSAMENTO DE LEGENDAS POR SURDOS E OUVINTES: UM ESTUDO  
EXPERIMENTAL COM RASTREADOR OCULAR

Tese apresentada ao Curso de Doutorado Acadêmico em Linguística Aplicada do Programa de Pós-Graduação em Linguística Aplicada da Universidade Estadual do Ceará, como requisito parcial para a obtenção do título de Doutor em Linguística Aplicada. Área de Concentração: Linguagem e Interação.

Aprovada em: 13 de abril de 2021.

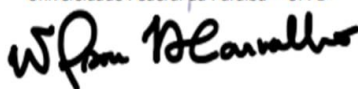
BANCA EXAMINADORA



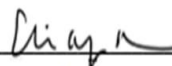
Prof. Dra. Vera Lúcia Santiago Araújo (Orientadora)  
Universidade Estadual do Ceará – UECE



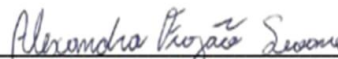
Prof. Dr. José Ferrari Neto  
Universidade Federal da Paraíba – UFPB



Prof. Dr. Wilson Junior de Araujo Carvalho  
Universidade Estadual do Ceará – UECE



Prof. Dra. Elisângela Nogueira Teixeira  
Universidade Federal do Ceará – UFC



Prof. Dra. Alexandra Frazão Seoane  
Universidade Estadual do Ceará – UECE

Ao meu pai e minha mãe, Neto e Ivone: foi por  
você e para você!

## AGRADECIMENTOS

Que jornada! São 29 anos de vida, dos quais 11 anos foram dedicados à academia. Muita gente foi importante nesse percurso e tenho receio de me esquecer de alguém, mas vamos lá:

A Deus, primeiramente, por sua benevolência e misericórdia na minha vida, mesmo sabendo que sou falho e não mereço.

À minha mãe, Ivone, por seu amor incondicional e pelo seu incentivo mesmo nas dificuldades; por ter feito de tudo para que eu pudesse estudar, inclusive pedindo dinheiro emprestado quando não tinha dinheiro para pagar nem a passagem para a faculdade. Mãe, sem você com certeza eu não teria chegado até aqui. Mulher guerreira, dona de casa, que não teve a chance de se dedicar aos estudos na juventude, mas sempre me mostrou que esse era o caminho, inclusive terminando o Ensino Médio depois dos 40 anos. Sua história me inspirou e me inspira! Te amo até o fim.

Ao meu pai, Neto, por ter sido o provedor da casa e ter permitido que eu me dedicasse aos estudos; por ter entendido quando eu disse que não queria trabalhar na oficina e queria estudar; por seu jeito próprio e bruto de demonstrar amor e cuidado. Tenho muito orgulho de ser seu filho! Tudo o que eu fiz até hoje foi para conseguir uma vida melhor para você e para a minha mãe. Te amo até o fim.

Ao meu irmão, Igor, por ser companheiro de vida; sem você eu também não estaria aqui, meu irmão. Te amo até o fim.

À minha orientadora, Vera, por ser muito mais que uma orientadora em minha vida; por me orientar não só nos estudos, mas na vida. Já são mais de 10 anos de conselhos, broncas (quando eu mereci – e mereci muitas vezes), afeto e amor. Não seria o que eu sou sem você também, professora. Minha eterna gratidão a você. Serei para sempre um dos seus *babies*. Seu trabalho transformou vidas e eu sou prova disso!

À minha coorientadora, Patrícia, por toda ajuda e apoio nesses 4 anos; por ser uma verdadeira amiga. Você que me escutou chorar por achar que eu não conseguiria. Eu não teria feito esse trabalho sem você. Obrigado por tudo!

À Thaís, por ter me encorajado quando eu achei que não conseguiria; por ter me dito que tudo ia dar certo no final; por ter sido tão importante para que eu completasse essa etapa da minha vida, mesmo à distância, por muitas vezes. Minha gratidão eterna a você. Você com certeza faz parte disso.

Ao Marcos Leandro, por ser um grande amigo nesses últimos anos. Fico feliz que os nossos caminhos tenham se cruzado de novo, meu amigo. As coisas boas que vivemos me deram fôlego para terminar este trabalho.

Ao meu querido amigo Matthias, por sua amizade verdadeira e por ser cúmplice de tantas aventuras da vida. Meu amigo, você também foi muito importante para que este trabalho fosse feito.

Ao meu amigo Demétrius, pela leveza de sua amizade, pelos momentos de diversão e pela profundidade de várias das nossas conversas. Fico feliz de termos nos aproximado ainda mais no último ano, e você também faz parte disso.

Aos meus amigos de Sobral, Lucas, Gabriel, Deyse e, especialmente, Kaoli e Daiani, por terem sido suporte nos últimos 4 anos da minha vida; por terem compreendido as tantas vezes que não pude estar presente por estar me dedicando a este trabalho. Obrigado por tantos momentos bons vividos e por estarem lá também nos momentos ruins.

À Carine, por ter me ajudado a aprender a utilizar o *RStudio* e por sempre estar disponível para minhas dúvidas.

À Gabi, pelo auxílio na realização da pesquisa e na transcrição dos relatos retrospectivos que serão utilizados em um momento posterior.

À Ana Carla, pelo auxílio com a revisão deste trabalho.

Aos meus amigos da vida, Wagner, Albert, Juliana, Jéssica, Williara, Jael, Rebeca, Diego, Jefferson, Jeilson, Kaliu, Laryssa, Hugo, João Victor, Victor e tantos outros que eu possa ter esquecido. Obrigado pelo carinho de sempre!

À Universidade Estadual Vale do Acaraú (UVA), especialmente ao colegiado do curso de Letras e à sua coordenação, por tanto me ensinarem desde que cheguei a Sobral como um “menino véi” de 25 anos. Amadureci muito profissionalmente nesses últimos anos, e vocês têm contribuição nisso. À coordenação, especialmente, por ter me permitido conciliar os horários do trabalho com os horários do doutorado.

A todos os meus alunos e ex-alunos da UVA, por terem sido compreensivos quando precisei me ausentar para alguma atividade do doutorado.

A todos os meus alunos e ex-alunos, por terem feito parte da minha vida. Espero que eu tenha feito em vocês pelo menos um terço do impacto que vocês tiveram em mim.

À Laryssa, minha amiga de doutorado e da vida! Já são mais de 10 anos de amizade e uma marca registrada da nossa relação é a leveza e a gaiatice. Em pensar que a gente se conheceu como professor e aluna e foi você quem me ensinou bem mais durante todo esse tempo.

Aos amigos da turma de doutorado, em especial, Róbson, Jana e Laryssa. Os prazos e cobranças às vezes tão sufocantes da pós-graduação seriam bem piores sem a companhia de vocês.

À família LEAD! Essa família é tão grande que com certeza eu esqueceria alguém. Obrigado por terem me ensinado tanto. Em especial, à Élide e à Alexandra, que me apresentaram à pesquisa em Tradução Audiovisual como professoras na graduação e aqui estou eu!

À Universidade Estadual do Ceará (UECE), por 11 anos de formação pública e de excelência. A todos os meus ex-professores e colegas que conheci nos muros dessa universidade que considero minha casa.

A todos do PosLA, desde o corpo docente, passando pela secretária e coordenação. Obrigado pela formação e pela contribuição na minha vida acadêmica.

À profa. Dra. Elisângela Nogueira Teixeira, que cedeu o equipamento e espaço do Laboratório de Psicolinguística e Ciências Cognitivas, sem os quais eu não teria conseguido realizar esta pesquisa. Por também ter orientado esse trabalho em diversas etapas.

Ao prof. Dr. José Ferrari Neto, por sua ajuda e disponibilidade constantes. Sem sua inspiração e auxílio essa pesquisa não existiria.

Aos Prof. Dr. Wilson Júnior de Araújo Carvalho e Profa. Dra. Alexandra Frazão Seoane, outros membros da banca, por sua disponibilidade para leitura e avaliação deste trabalho.

À Mariana, por ter auxiliado na interpretação em Libras nas sessões de coleta de dados.

Aos participantes desta pesquisa, por sua disponibilidade e contribuição para esse trabalho.

À comunidade surda, por sua luta por direitos de acessibilidade que inspira a nossa ciência; por sua disponibilidade em participar das pesquisas do LEAD. Esta pesquisa não teria propósito sem vocês.



## RESUMO

Pesquisas experimentais baseadas em rastreamento ocular têm ganhado força no estudo dos parâmetros de legendagem para ouvintes e Legendagem para Surdos e Ensurdecidos (LSE) nos últimos anos (DE LINDE; KAY, 1999; D'YDEWALLE *et al.*, 1987; D'YDEWALLE; DE BRUYCKER, 2007; MONTEIRO, 2016; VIEIRA, 2016, dentre outros). Dois desses parâmetros (velocidade da legenda e número de linhas) ainda carecem de maiores estudos, apesar de algumas das pesquisas supracitadas terem abordado esses aspectos. Nesse contexto, esta tese teve como objetivo investigar como a velocidade e o número de linhas influenciam o processamento de legendas por surdos e ouvintes testando duas hipóteses: a) Legendas de 1 linha induzem um processamento menos linear e, dessa forma, mais custoso do que legendas de 2 linhas; b) Legendas lentas (145 palavras por minuto – ppm) e rápidas (180 ppm) não diferem significativamente no processamento de surdos e ouvintes. Para tal, realizamos um estudo experimental de cunho quantitativo no qual selecionamos 8 vídeos (8 a 15 segundos de duração) do filme *A rede social* (2010), cujas legendas foram manipuladas de acordo com as quatro condições experimentais da pesquisa: 1 linha, 2 linhas, lentas (145 ppm ou 14 caracteres por segundo) e rápidas (180 ppm ou 17 caracteres por segundo) em um *design* fatorial 2 x 2 x 2. Os participantes da pesquisa (20 ouvintes e 20 surdos) assistiram a esses vídeos legendados em um aparelho de rastreamento ocular. Com auxílio de tratamento estatístico, os resultados da análise das fixações a partir de medidas de processamento indicaram que as legendas de 2 linhas induzem um processamento mais confortável por surdos e ouvintes, confirmando nossa hipótese de pesquisa. Sobre a velocidade da legenda, as medidas de processamento indicaram um processamento um pouco mais linear nas legendas de 180 ppm, refutando parcialmente nossa hipótese. Os resultados sobre velocidade precisam ser qualificados no futuro em relação ao papel de medidas como a de regressão e releitura para o processamento de legendas, um assunto ainda controverso na literatura em legendagem.

**Palavras-chave:** Tradução Audiovisual Acessível. Legendagem para Surdos e Ensurdecidos. Rastreamento Ocular. Velocidade da Legenda. Número de Linhas.

## ABSTRACT

Experimental research based on eye tracking has gained status in the study of subtitling for the hearing and Subtitling for Deaf and Hard-of-Hearing (SDH) parameters in the past years (DE LINDE; KAY, 1999; D'YDEWALLE *et al.*, 1987; D'YDEWALLE; DE BRUYCKER, 2007; MONTEIRO, 2016; VIEIRA, 2016, among others). Two of these parameters (subtitle speed and number of lines) still need more investigation, even though some of the aforementioned studies have tackled them. In this context, this dissertation aimed at investigating how the subtitle speed and the number of lines influence the processing of subtitles by deaf and hearing participants testing two hypotheses: a) 1-line subtitles induce a less linear processing and, likewise, a more demanding one than 2-line subtitles; b) Slow subtitles (145 words per minute – wpm) and fast ones (180 wpm) do not differ significantly in the processing of deaf and hearing participants. For that, we conducted a quantitative experimental study in which we selected 8 videos (duration between 8 and 15 seconds) of the movie *The social network* whose subtitles were manipulated in line with the four experimental conditions of the research: 1 line, 2 lines, slow (145 wpm or 14 characters per second) and fast (180 wpm or 17 characters per second) in a 2 x 2 x 2 factorial design. The participants of this research (20 hearing and 20 deaf) watched these manipulated videos on an eye-tracking device. With the help of statistical treatment, the results concerning the analysis of fixations based on processing measures indicated that 2-line subtitles induce a more comfortable processing by the deaf and hearing participants, confirming our research hypothesis. On subtitle speed, the processing measures indicated a slightly more linear processing of 180-wpm subtitles, partially denying our hypothesis. The results on subtitle speed need to be qualified in the future regarding the role of measures like regression and rereading on the processing of subtitles, a still controversial topic in subtitling literature.

**Keywords:** Accessible Audiovisual Translation. Subtitling for the Deaf and Hard-of-Hearing. Eye Tracking. Subtitle Rate. Number of Lines.

## LISTA DE FIGURAS

Figura 1 – Distribuição de duração de fixações .....	37
Figura 2 – Mapa de calor da mesma cena vista sem e com legenda .....	45
Figura 3 – <i>Scanpaths</i> de um surdo (à esquerda) e um ensurdecido (à direita) assistindo a um vídeo com legendas <i>verbatim</i> .....	47
Figura 4 – Vídeos e condições experimentais .....	65
Figura 5 – Exemplo do Vídeo 1 com legenda de 1 linha .....	82
Figura 6 – Exemplo do Vídeo 2 com legenda de 1 linha .....	83
Figura 7 – Exemplo do Vídeo 3 com legenda de 2 linhas .....	84
Figura 8 – Exemplo do Vídeo 4 com legenda de 2 linhas .....	85
Figura 9 – Exemplo do Vídeo 5 com legenda de 1 linha .....	86
Figura 10 – Exemplo do Vídeo 6 com legenda de 2 linhas .....	87
Figura 11 – Exemplo do Vídeo 7 com legenda de 2 linhas .....	88
Figura 12 – Exemplo do Vídeo 8 com legenda de 2 linhas .....	89
Figura 13 – Interface do <i>Subtitle Workshop</i> durante manipulação das legendas .....	98
Figura 14 – Interface do <i>Subtitle Workshop</i> na guia SETTINGS .....	99
Figura 15 – Interface do <i>Subtitle Workshop</i> aberto durante manipulação das legendas .....	100
Figura 16 – Interface do <i>Tobii Studio</i> aberta para início da realização do experimento .....	112
Figura 17 – Interface do <i>Tobii Studio</i> com Áreas de Interesse delimitadas em um dos estímulos .....	113
Figura 18 – Interface da ferramenta <i>Statistics</i> com dados sobre número e duração média de fixações .....	114
Figura 19 – Interface da ferramenta <i>GazePlot</i> .....	114
Figura 20 – <i>GazePlot</i> durante inspeção da movimentação do olhar .....	119

## LISTA DE QUADROS

<b>Quadro 1 – Duração média das fixações e comprimento das sacadas por tarefas .....</b>	<b>38</b>
<b>Quadro 2 – Legenda do Vídeo 1 manipulado na condição 1L145 .....</b>	<b>91</b>
<b>Quadro 3 – Legenda do Vídeo 2 manipulado na condição 1L180 .....</b>	<b>91</b>
<b>Quadro 4 – Legenda do Vídeo 3 manipulado na condição 2L145 .....</b>	<b>91</b>
<b>Quadro 5 – Legenda do Vídeo 4 manipulado na condição 2L180 .....</b>	<b>92</b>
<b>Quadro 6 – Legenda do Vídeo 5 manipulado na condição 1L145 .....</b>	<b>92</b>
<b>Quadro 7 – Legenda do Vídeo 6 manipulado na condição 1L180 .....</b>	<b>92</b>
<b>Quadro 8 – Legenda do Vídeo 7 manipulado na condição 2L145ppm.....</b>	<b>93</b>
<b>Quadro 9 – Legenda do Vídeo 8 manipulado na condição 2L180 .....</b>	<b>93</b>
<b>Quadro 10 – Características linguísticas das legendas .....</b>	<b>94</b>
<b>Quadro 11 – Legenda do Vídeo 5 manipulada nas 4 condições experimentais .....</b>	<b>102</b>
<b>Quadro 12 – Dados demográficos dos participantes do grupo controle de ouvintes .....</b>	<b>104</b>
<b>Quadro 13 – Dados demográficos dos participantes do grupo controle de surdos .....</b>	<b>105</b>
<b>Quadro 14 – Quadrado Latino: condições, vídeos e pares de participantes .....</b>	<b>107</b>

## LISTA DE GRÁFICOS

Gráfico 1 – Duração média de fixação por participante ouvinte .....	125
Gráfico 2 – Duração média de fixação por participante surdo .....	126
Gráfico 3 – Distribuição das fixações de ouvintes e surdos .....	127
Gráfico 4 – <i>Boxpot</i> da duração média das fixações .....	128
Gráfico 5 – Duração Média de Fixação por Vídeo .....	129
Gráfico 6 – Modelo de Regressão Linear com Efeito Fixo de Vídeo .....	132
Gráfico 7 – Modelo de Regressão Linear Misto de Duração Média de Fixação.....	135
Gráfico 8 – Número de fixações por participante ouvinte.....	140
Gráfico 9 – Número de Fixações por Participante Surdo.....	141
Gráfico 10 – Distribuição do número de fixações por ouvintes e surdos.....	142
Gráfico 11 – <i>Boxplot</i> do número de fixações antes e depois da filtragem .....	143
Gráfico 12 – Número de fixações por vídeo .....	143
Gráfico 13 – Modelo de Regressão Linear Misto de Número de Fixações por Legenda .....	145
Gráfico 14 – Modelo de Regressão Linear Misto de Número de Fixações por Legenda com Interação .....	146
Gráfico 15 – Média de duração de regressão do grupo de ouvintes .....	150
Gráfico 16 – Média de duração de regressão por participante surdo .....	151
Gráfico 17 – Duração média de regressão por vídeo.....	152
Gráfico 18 – Distribuição das regressões de ouvintes e surdos .....	153
Gráfico 19 – <i>Boxplots</i> de duração média das regressões antes e depois da filtragem ....	153
Gráfico 20 – Proporção de casos de regressão por grupo de participante.....	156
Gráfico 21 – Proporção de casos de regressão por número de linhas .....	157
Gráfico 22 – Proporção de casos de regressão por velocidade.....	158
Gráfico 23 – Proporção de casos de regressão por vídeo.....	160
Gráfico 24 – Modelo de Regressão Logística Misto de Regressão .....	162
Gráfico 25 – Média de duração de releitura por participante ouvinte.....	167
Gráfico 26 – Média de duração de releitura por participante surdo.....	168
Gráfico 27 – Média de duração de releitura por vídeo .....	169
Gráfico 28 – Distribuição de releitura em ouvintes e surdos.....	170
Gráfico 29 – <i>Boxplot</i> de duração média de releitura antes e após filtragem.....	171
Gráfico 30 – Proporção de casos de releitura por Grupo .....	174

<b>Gráfico 31 – Proporção de casos de releitura por Número de Linhas .....</b>	<b>175</b>
<b>Gráfico 32 – Proporção de casos de releitura por Velocidade .....</b>	<b>176</b>
<b>Gráfico 33 – Proporção de casos de releitura por Vídeo .....</b>	<b>177</b>
<b>Gráfico 34 – Modelo de Regressão Logística Misto de Releitura.....</b>	<b>179</b>
<b>Gráfico 35 – Interação do Modelo de Regressão Linear Misto de Releitura.....</b>	<b>180</b>
<b>Gráfico 36 – Duração média de encapsulamento por ouvintes .....</b>	<b>184</b>
<b>Gráfico 37 – Duração média de encapsulamento por surdos .....</b>	<b>185</b>
<b>Gráfico 38 – Duração média de encapsulamento por vídeo .....</b>	<b>186</b>
<b>Gráfico 39 – Distribuição de duração de encapsulamento por grupo .....</b>	<b>187</b>
<b>Gráfico 40 – <i>Boxplot</i> de duração média de encapsulamento antes e depois da filtragem .....</b>	<b>188</b>
<b>Gráfico 41 – Frequência de casos de deflexão por Grupo .....</b>	<b>192</b>
<b>Gráfico 42 – Frequência de casos de deflexão por Número de Linhas.....</b>	<b>193</b>
<b>Gráfico 43 – Frequência de casos de deflexão por Velocidade.....</b>	<b>194</b>
<b>Gráfico 44 – Frequência relativa de casos de deflexão por vídeo .....</b>	<b>195</b>
<b>Gráfico 45 – Modelo de Regressão Logística Misto de Deflexão.....</b>	<b>196</b>
<b>Gráfico 46 – Média de duração de latência por participante ouvinte .....</b>	<b>200</b>
<b>Gráfico 47 – Média de duração de latência por participante surdo .....</b>	<b>201</b>
<b>Gráfico 48 – Média de duração de latência por vídeo.....</b>	<b>202</b>
<b>Gráfico 49 – Distribuição de duração de latência entre os participantes.....</b>	<b>203</b>
<b>Gráfico 50 – <i>Boxplot</i> de duração média de latência dos participantes antes e após a filtragem.....</b>	<b>204</b>
<b>Gráfico 51 – Modelo de Regressão Linear Misto de Duração Média de Latência .....</b>	<b>206</b>
<b>Gráfico 52 – Modelo de Regressão Linear Misto de Duração Média de Latência com Interação .....</b>	<b>207</b>
<b>Gráfico 53 – Interação do Modelo de Regressão Linear Misto de Duração Média de Latência.....</b>	<b>207</b>

## LISTA DE TABELAS

Tabela 1 – Número de caracteres por segundo para a velocidade de 145 ppm .....	31
Tabela 2 – Duração média das fixações por Grupo.....	122
Tabela 3 – Duração média das fixações por Número de Linhas .....	122
Tabela 4 – Duração média das fixações por Velocidade .....	123
Tabela 5 – Duração média de fixação entre condições experimentais.....	123
Tabela 6 – Duração média de fixação por grupo após filtragem .....	130
Tabela 7 – Duração média de fixação por número de linha após filtragem .....	130
Tabela 8 – Duração média de fixação por velocidade após filtragem.....	130
Tabela 9 – Comparação entre Modelos Aninhados.....	134
Tabela 10 – Valores previstos pelo modelo para duração média de fixação.....	135
Tabela 11 – Número de fixações por legenda por Grupo .....	137
Tabela 12 – Número de fixações por legenda por Número de Linha.....	137
Tabela 13 – Número de fixações por legenda por Velocidade.....	138
Tabela 14 – Número de fixações por condição experimental .....	139
Tabela 15 – Modelo de Regressão Linear Misto de Número de Fixações por Legenda	145
Tabela 16 – Valores Previstos pelo Modelo de Regressão Linear Misto de Número de Fixações .....	147
Tabela 17 – Duração média de regressão por Grupo.....	148
Tabela 18 – Duração média de regressão por Número de Linhas .....	149
Tabela 19 – Duração média de regressão por Velocidade .....	149
Tabela 20 – Duração Média de Regressão por Condição Experimental.....	149
Tabela 21 – Duração média de regressão por Grupo (após filtragem).....	154
Tabela 22 – Duração média de regressão por Número de Linhas (após filtragem) .....	154
Tabela 23 – Duração média de regressão por Velocidade (após filtragem).....	154
Tabela 24 – Ocorrências de regressão por condição experimental.....	159
Tabela 25 – Valores previstos pelo modelo para ocorrência de regressão .....	162
Tabela 26 – Duração média de releitura por Grupo .....	164
Tabela 27 – Duração média de releitura por Número de Linhas.....	165
Tabela 28 – Duração média de releitura por Velocidade.....	165
Tabela 29 – Duração média de releitura por condição experimental.....	166
Tabela 30 – Duração média de releitura por Grupo após filtragem.....	171
Tabela 31 – Duração média de releitura por Número de Linha após filtragem.....	172

<b>Tabela 32 – Duração média de releitura por Número de Linha após filtragem.....</b>	<b>172</b>
<b>Tabela 33 – Modelo de Regressão Linear Misto de Duração Média de Releitura .....</b>	<b>173</b>
<b>Tabela 34 – Ocorrências de releitura por condição experimental .....</b>	<b>176</b>
<b>Tabela 35 – Valores previstos pelo modelo para ocorrência de releitura .....</b>	<b>180</b>
<b>Tabela 36 – Duração média de encapsulamento por Grupo .....</b>	<b>182</b>
<b>Tabela 37 – Duração média de encapsulamento por Número de Linha .....</b>	<b>182</b>
<b>Tabela 38 – Duração média de encapsulamento por Velocidade.....</b>	<b>182</b>
<b>Tabela 39 – Duração média de encapsulamento por condição experimental .....</b>	<b>183</b>
<b>Tabela 40 – Duração média de encapsulamento por Grupo após filtragem.....</b>	<b>188</b>
<b>Tabela 41 – Duração média de encapsulamento por Número de Linha após filtragem .....</b>	<b>189</b>
<b>Tabela 42 – Duração média de encapsulamento por Velocidade após filtragem .....</b>	<b>189</b>
<b>Tabela 43 – Valores previstos pelo modelo para duração média de encapsulamento ...</b>	<b>191</b>
<b>Tabela 44 – Ocorrências de deflexão por condição experimental.....</b>	<b>194</b>
<b>Tabela 45 – Valores previstos para a ocorrência de deflexão .....</b>	<b>197</b>
<b>Tabela 46 – Duração média de latência por Grupo.....</b>	<b>198</b>
<b>Tabela 47 – Duração média de latência por Número de Linhas .....</b>	<b>199</b>
<b>Tabela 48 – Duração média de latência por Velocidade .....</b>	<b>199</b>
<b>Tabela 49– Duração média de latência por condição experimental em surdos e ouvintes.....</b>	<b>199</b>
<b>Tabela 50 – Duração média de latência por Número de Linha após filtragem .....</b>	<b>204</b>
<b>Tabela 51 – Duração média de latência por Velocidade após filtragem.....</b>	<b>204</b>
<b>Tabela 52 – Duração média de latência por Grupo após filtragem .....</b>	<b>205</b>
<b>Tabela 53 – Valores previstos para duração média de latência a partir do modelo .....</b>	<b>208</b>



## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>18</b>
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA .....</b>	<b>27</b>
<b>2.1</b>	<b>Legendagem para surdos e ensurdecidos (LSE).....</b>	<b>27</b>
<b>2.2</b>	<b>Sobre a velocidade da legenda.....</b>	<b>28</b>
<b>2.2</b>	<b>Legendagem e rastreamento ocular.....</b>	<b>35</b>
2.2.1	O rastreamento ocular .....	35
2.2.2	O processamento de legendas.....	43
2.2.3	Pesquisa em legendagem com rastreamento ocular .....	47
<b>3</b>	<b>METODOLOGIA.....</b>	<b>78</b>
<b>3.1</b>	<b>Tipo de pesquisa .....</b>	<b>78</b>
<b>3.2</b>	<b>Contexto da pesquisa.....</b>	<b>79</b>
<b>3.3</b>	<b>Estímulos .....</b>	<b>80</b>
3.3.1	Descrição imagética dos vídeos legendados .....	82
3.3.1.1	<i>Vídeo 1 – “Bebendo em casa”.....</i>	82
3.3.1.2	<i>Vídeo 2 – “O julgamento” .....</i>	83
3.3.1.3	<i>Vídeo 3 – “O roubo”.....</i>	84
3.3.1.4	<i>Vídeo 4 – “O encontro”.....</i>	85
3.3.1.5	<i>Vídeo 5 – “A negociação” .....</i>	86
3.3.1.6	<i>Vídeo 6 – “A festa”.....</i>	87
3.3.1.7	<i>Vídeo 7 – “Sobre Christy...”.....</i>	88
3.3.1.8	<i>Vídeo 8 – “A nova sede”.....</i>	89
3.3.2	Descrição linguística das legendas .....	90
<b>3.4</b>	<b>Procedimentos metodológicos .....</b>	<b>96</b>
3.4.1	Preparação e manipulação dos estímulos .....	96
3.4.1.1	<i>Subtitle Workshop .....</i>	97
3.4.2	Seleção dos participantes.....	103
3.4.3	Desenho experimental .....	105
3.4.4	Variáveis e medidas.....	109
3.4.5	Instrumentos .....	111
3.4.5.1	<i>O rastreador ocular .....</i>	111
3.4.6	Procedimentos éticos.....	115
3.4.7	Procedimentos de coleta do comportamento ocular.....	115

3.4.8	Procedimentos de análise .....	116
<b>4</b>	<b>RESULTADOS E DISCUSSÃO DOS DADOS.....</b>	<b>118</b>
<b>4.1</b>	<b>Resultados quantitativos sobre a movimentação ocular dos participantes.....</b>	<b>118</b>
4.1.1	Duração média das fixações .....	121
4.1.2	Número de fixações.....	137
4.1.3	Regressão.....	148
4.1.4	Releitura .....	164
4.1.5	Encapsulamento.....	181
4.1.6	Deflexão .....	191
4.1.7	Latência .....	198
<b>4.2</b>	<b>Discussão dos resultados.....</b>	<b>209</b>
<b>5</b>	<b>CONCLUSÃO .....</b>	<b>233</b>
	<b>REFERÊNCIAS.....</b>	<b>243</b>
	<b>APÊNDICE A – PERFIL PRÉ-COLETA DOS PARTICIPANTES</b>	
	<b>OUVINTES.....</b>	<b>248</b>
	<b>APÊNDICE B – PERFIL PRÉ-COLETA DOS PARTICIPANTES SURDOS ...</b>	<b>250</b>
	<b>APÊNDICE C – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO</b>	
	<b>PARA OUVINTES .....</b>	<b>253</b>
	<b>APÊNDICE D – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO</b>	
	<b>PARA SURDOS .....</b>	<b>254</b>

## 1 INTRODUÇÃO

A LSE – legendagem para surdos e ensurdecidos – tem se tornado um recurso possível para que as pessoas que se encaixam no seu público-alvo desfrutem da experiência de assistir a um produto audiovisual, como novelas, peças de teatro, telejornais, programas de TV, filmes etc. Nesse contexto, a norma complementar 01/2006, aprovada pela portaria 310 do Ministério das Comunicações, de 27 de junho de 2006, assevera, para um prazo de até 11 anos a partir da sua data de publicação, que a LSE e a janela de Libras para pessoas surdas, além da audiodescrição para pessoas com deficiência visual (PcDVs), estejam presentes em toda a programação veiculada pelas estações transmissoras ou retransmissoras de radiodifusão.

Contudo, no caso específico da LSE, pesquisas de recepção realizadas pela UECE (ARAÚJO, 2008; ARAÚJO, 2012; ARAÚJO; NASCIMENTO, 2011; FRANCO; ARAÚJO, 2003) com surdos cearenses e de todo o Brasil têm apontado para uma necessidade de ajustes no que tange aos parâmetros utilizados na confecção da LSE do tipo *closed caption*, disponível na programação da TV brasileira, sugerindo que esses parâmetros não satisfazem a demanda da comunidade surda. Os resultados alcançados pelo Projeto Modelo de Legendagem para Surdos (MOLES), uma pesquisa de recepção nacional com 34 surdos de 4 regiões brasileiras que teve como objetivo apresentar parâmetros de LSE confortáveis para os surdos brasileiros (ARAÚJO, 2012), sugeriram que legendas bem segmentadas promovem uma recepção eficaz do programa televisivo legendado mesmo que sejam rápidas (entre 160 e 180 palavras por minuto – ppm). Isso nos surpreendeu, pois acreditávamos, entre os pesquisadores do Grupo LEAD (Legendagem e Audiodescrição), com base em pesquisas anteriores (ARAÚJO, 2008; FRANCO; ARAÚJO, 2003), que o principal entrave para uma recepção eficiente de filmes e programas legendados por parte de surdos estaria na velocidade da legenda.

Esses resultados alcançados pelas pesquisas exploratórias supracitadas, no que diz respeito à segmentação, encontraram ressonância na literatura em Tradução Audiovisual (TAV). Díaz Cintas e Remael (2007, p. 173), por exemplo, afirmam que: “Uma segmentação cuidadosa da informação pode ajudar a reforçar a coerência e a coesão na legendagem<sup>1</sup>”. Ainda neste sentido, Karamitroglou (1998), ao apontar a necessidade de legendas serem divididas em texto de uma forma que sejam autossuficientes sintática e semanticamente, diz que: “Quando segmentamos uma sentença, forçamos o cérebro a pausar o seu processamento linguístico por

---

<sup>1</sup> “A careful segmentation of the information can help reinforce coherence and cohesion in subtitling”. Todas as traduções ao longo do trabalho são de nossa autoria.

um momento até que os olhos captem a próxima informação<sup>2</sup>” (KARAMITROGLOU, 1998, p. 10). Para uma boa segmentação, dessa forma, os legendistas devem seguir diretrizes linguísticas, ou seja, acontecer no mais alto nível sintático possível, segundo o autor.

No que tange à descrição e à classificação dos problemas de segmentação linguística na LSE do tipo *closed caption*, resultados em diferentes gêneros de LSE demonstraram certas regularidades em relação à sua natureza e aos seus padrões relacionados a esses problemas (ARAÚJO, 2015; ARAÚJO; ASSIS, 2014; ASSIS, 2016; GABRIEL, 2015), como uma grande quantidade de problemas de segmentação indevida, principalmente nos sintagmas verbal e nominal, em legendas rápidas e de três linhas. Essas pesquisas foram realizadas no âmbito do Projeto CORSEL (*Corpus, Segmentação e Legendagem*), vinculado ao Grupo de Pesquisa Tradução e Semiótica da Universidade Estadual do Ceará, mais especificamente à linha de pesquisa Legendagem e Audiodescrição (LEAD). Esse projeto foi realizado entre os anos de 2013 e 2015 e teve como objetivo, como supracitado, a identificação, descrição e classificação dos problemas de segmentação na LSE da TV brasileira.

Além do CORSEL e de sua dimensão prioritariamente descritiva, o LEAD, concomitantemente, iniciou pesquisas experimentais com a utilização do rastreador ocular para a análise de parâmetros da legendagem. O projeto denominado Estudos Experimentais em Tradução Audiovisual para Surdos: Análise da Velocidade e da Segmentação<sup>3</sup> (ExLeg) tem como objetivo investigar se os parâmetros apontados em pesquisas anteriores como problemáticos para uma recepção eficiente da LSE – segmentação e velocidade – de fato dificultavam a experiência do usuário surdo na apreciação do filme legendado. Apesar do reconhecimento da influência do parâmetro da segmentação apregoada na literatura em TAV, como discutido acima, poucos foram os trabalhos que se utilizaram de metodologias experimentais para, de fato, verificar a extensão e influência da segmentação no processamento e recepção de legendas por ouvintes (PEREGO *et al.*, 2010; RAJENDRAN *et al.*, 2013); com participantes surdos, então, não havia nenhum estudo que se assemelhasse em objetivo e metodologia, sendo essa a lacuna a ser preenchida pelo projeto ExLeg. Além disso, o parâmetro da velocidade, controverso na literatura em TAV quanto a qual velocidade de legenda seria ideal (DE LINDE; KAY, 1999; JENSEMA, 1998), seria analisado de forma simultânea à

---

<sup>2</sup> “When we segment a sentence, we force the brain to pause its linguistic processing for a while, until the eyes trace the next piece of linguistic information.”

<sup>3</sup> O Projeto ExLeg teve os trabalhos de Monteiro (2016) e Vieira (2016) como parte de sua fase 1. Este trabalho pretende iniciar a fase 2 do projeto, o qual pretende continuar testando o parâmetro já testado por Monteiro e Araújo (velocidade) em outro gênero (neste caso, o fílmico), assim como a influência do número de linhas no processamento de legendas por surdos e ouvintes.

segmentação de forma a tentar esclarecer a influência de cada um desses parâmetros separadamente, assim como em conjunto, a partir dos resultados sugeridos pelas pesquisas exploratórias sobre LSE já realizadas pelo Grupo LEAD<sup>4</sup>.

Dessa forma, dois trabalhos representaram a fase 1 das pesquisas no âmbito do Projeto ExLeg, os quais apresentam grande semelhança com a presente pesquisa: Vieira (2016) e Monteiro (2016), cujo objetivo em comum foi analisar a influência dos parâmetros da velocidade e da segmentação linguística no processamento de LSE por surdos em programas da TV brasileira. Esses dois trabalhos são fundamentais para a realização da presente pesquisa pelo *corpus* em comum (LSE) e metodologia experimental baseada na utilização de rastreador ocular para análise de parâmetros da LSE e seu processamento por participantes surdos.

Por um lado, os resultados de Vieira (2016), ao utilizar trechos do programa de TV no formato de documentário *Globo Repórter* como estímulo do seu experimento, sugeriram que legendas mal segmentadas causam maior desconforto e custo ao processamento dos participantes, sendo as legendas rápidas (180 palavras por minuto) e bem segmentadas as mais confortáveis, principalmente para os surdos. Monteiro (2016), por sua vez, trilhando percurso metodológico semelhante ao de Vieira (2016), analisou o processamento da LSE de propagandas políticas veiculadas na TV; assim como Vieira (2016), a autora observou em seus dados a relevância de uma boa segmentação tanto em legendas rápidas quanto lentas (145 palavras por minuto) para um processamento menos custoso por surdos e ouvintes. Por outro lado, diferentemente de Vieira (2016), os resultados de Monteiro (2016) sugeriram que o parâmetro da velocidade não foi preponderante para o processamento dos participantes do experimento.

A pesquisa de Monteiro (2016), no entanto, analisou apenas legendas de 1 linha, características da legenda de propaganda política veiculada na TV brasileira, algo que pode ter influenciado os seus resultados. Isso porque alguns estudos, como d'Ydewalle *et al.* (1987), Praet *et al.* (1991) e d'Ydewalle e De Bruycker (2007), apontam que o processamento e leitura de legendas de 2 linhas é mais regular e menos custoso por proporcionarem, dentre outras questões, uma sintaxe mais completa ao espectador.

Apesar de as duas pesquisas terem suas similaridades em termos de percurso metodológico e *corpus*, algumas questões podem ter causado diferenças nos resultados obtidos. Primeiramente, a forma como os dados foram tratados nas duas pesquisas. Apesar de Vieira

---

<sup>4</sup> Apesar de o grupo ser cadastrado oficialmente como Grupo de Pesquisa Tradução e Semiótica, ele é mais conhecido como Grupo LEAD (Legendagem e Audiodescrição). Sendo assim, esse é o nome que mais se repete ao longo deste trabalho.

(2016) e Monteiro (2016) terem proposto uma análise a partir de uma triangulação de dados provenientes de relatos retrospectivos e guiados (parte exploratória da pesquisa) e dados referentes ao rastreador ocular (parte experimental), ou seja, pesquisas de natureza quali-quantitativa, apenas Vieira (2016) analisou os dados quantitativos estatisticamente. A partir disso, a autora ofereceu resultados mais sólidos a respeito do efeito significativo das variáveis independentes da pesquisa tanto entre os grupos de surdos (grupo experimental) e de ouvintes (grupo controle), quanto dentro dos participantes do próprio grupo.

Outro ponto que pode ter diferenciado os resultados das pesquisas em questão foi o número de medidas (variáveis dependentes) utilizadas por ambas. Enquanto Vieira (2016) utilizou um conjunto de medidas para mensurar, através dos dados do rastreador ocular, um possível desconforto dos participantes submetidos às condições experimentais (número de fixações, duração das fixações, deflexões, regressões, atraso da primeira fixação nas legendas, perda de tudo ou parte da legenda), Monteiro (2016) utilizou apenas o número e a duração de fixações em sua análise.

Além desses dois pontos, uma outra questão que pode ter influenciado os resultados distintos de seu experimento está relacionada à própria constituição do seu gênero de LSE: a LSE de propagandas políticas. Ao contrário da legendagem profissional para ouvintes (e da LSE de DVDs também), que apresenta legendas com 1 e 2 linhas, a LSE de campanhas políticas no Brasil normalmente apresenta apenas legendas de 1 linha, o que pode ter gerado certo desconforto aos participantes da pesquisa (MONTEIRO, 2016).

Díaz Cintas e Remael (2007, p. 82) apontam que, enquanto a legendagem interlinguística se limita a 2 linhas de texto, a legendagem para surdos e ensurdecidos pode apresentar também legendas de 3 ou 4 linhas, ao mesmo tempo que legendas bilíngues, em alguns casos, também podem apresentar legendas de 4 linhas. Mesmo reconhecendo a existência de legendas com um número maior do que 2 linhas, os autores não fazem julgamento de valor em relação aos tipos de legendagem que apresentam um número maior de linhas.

No que diz respeito à preferência de legendas de 1 ou 2 linhas, os autores (p. 86) ponderam que essa escolha normalmente se baseia em fatores linguísticos (como no caso da segmentação) e técnicos (tempo de exposição em tela, número de caracteres etc.), sendo que há situações<sup>5</sup> ou meios de distribuição que preferem, em determinados casos, legendas de 1 ou 2

---

<sup>5</sup> Uma das situações citadas pelos autores é de modo que se evite a obstrução de outras informações lexicais que possam aparecer em tela; especialmente em documentários e entrevistas, ao invés de se colocar a legenda em outros lugares da tela que não na parte inferior (como ocorre comumente), a utilização de legendas de 1 linha seria preferível (DÍAZ CINTAS; REMAEL, 2007, p. 83).

linhas. Legendas muito longas, por exemplo, preferencialmente deveriam ser segmentadas em duas linhas de legenda de modo a diminuir o percurso do movimento ocular e fazer com que os usuários consigam confortavelmente fazer um movimento entre a área da legenda e a da imagem.

A preferência por legendas de duas linhas à de uma linha<sup>6</sup> já foi discutida nos estudos de d'Ydewalle *et al.* (1987), d'Ydewalle *et al.* (1991), Praet *et al.* (1990), d'Ydewalle e De Bruycker (2007) e Kruger e Steyn (2014). Os resultados, entretanto, não foram uniformes e, dessa forma, não são conclusivos sobre qual tipo de legenda (1 ou 2 linhas) seria menos custosa para os espectadores. Dentre estes, os que mais se aproximaram de um estudo do processamento (e não apenas do estudo da alocação de atenção entre a área da imagem e área da legenda) foram d'Ydewalle e De Bruycker (2007) e Kruger e Steyn (2014). De acordo com os resultados de d'Ydewalle e De Bruycker (2007), os participantes do seu estudo estiveram por mais tempo com o olhar fixado nas legendas de 2 linhas e apresentaram um tempo maior de latência (intervalo entre a última fixação na imagem e a primeira na legenda) nessas legendas. De uma forma geral, contudo, os autores observaram um processamento mais confortável de ouvintes nas legendas de 2 linhas, pois estas proporcionariam um padrão de leitura mais fluido, reforçando os dados sugeridos por d'Ydewalle *et al.* (1987), d'Ydewalle *et al.* (1991) e Praet *et al.* (1990). Já Kruger e Steyn (2014), por sua vez, ao utilizarem o que chamaram de Índice de Leitura de Textos Dinâmicos para analisar o processamento de legendas em língua inglesa como L2 em um contexto acadêmico, não observaram diferenças estatísticas significantes entre legendas de 1 e 2 linhas.

Uma questão sobre o trabalho de d'Ydewalle e De Bruycker (2007) que deve ser ressaltada, no entanto, é a de que os autores testaram esse processamento sem isolar essas duas variáveis, ou seja, exibiram aos participantes o mesmo vídeo com legendas de 2 e 1 linha. Neste estudo, contudo, observamos o olhar dos espectadores surdos e ouvintes assistindo a vídeos com uma sintaxe completa (2 linhas) e não tão completa (1 linha), assim como nas condições de legendas lentas e rápidas.

As recomendações sobre o número de linhas na LSE profissional do Brasil não parecem estar sendo seguidas pelas empresas responsáveis pela produção desse tipo de legenda em DVDs. O estudo de Chaves (2017) verificou que, ao analisar as características técnicas e os problemas de segmentação de um *corpus* composto de legendas para ouvintes de filmes de

---

<sup>6</sup> Ao longo do trabalho, utilizaremos tanto a notação “uma e duas linhas”, quanto “1 e 2 linhas” de maneira indistinta.

DVD (em inglês e português), 60% das legendas desse *corpus* eram compostas por legendas de 1 linha.

A partir disso, levando em consideração o que a literatura da área fala sobre o número de linhas em legendagem, assim como os resultados apontados por Monteiro (2016) em sua pesquisa com rastreador ocular e por Chaves (2017) através da sua descrição das características da legenda de filmes em DVD, qual seria a influência de número de linhas na recepção de LSE por espectadores surdos e ouvintes?

De forma a verificar e ampliar o contexto no qual esse parâmetro seria analisado, decidimos adicionar o elemento da velocidade no estudo, tanto de forma a respaldar ou refutar a questão já discutida a respeito da influência da velocidade nos estudos de d'Ydewalle *et al.* (1987), Jensema (1998), De Linde e Kay (1999), Araújo e Nascimento (2011), Monteiro (2016), Vieira (2016) e Souza (2020) quanto analisar a influência do número de linhas em contextos de velocidade lenta (145 ppm) e rápida (180 ppm). Dentre estes estudos, alguns tiveram como participantes ouvintes (D'YDEWALLE *et al.*, 1987), alguns surdos (ARAÚJO; NASCIMENTO, 2011) e outros ambos os grupos (ARAÚJO; NASCIMENTO, 2011; DE LINDE; KAY, 1999; JENSEMA, 1998; MONTEIRO, 2016; VIEIRA, 2016), fazendo uso de metodologias de pesquisa distintas, desde metodologias exploratórias baseadas em instrumentos qualitativos a metodologias mistas baseadas no uso de rastreamento ocular. Os resultados dessas pesquisas divergiram em relação ao que seria uma velocidade mais confortável (ou passível de ser processada) para o processamento dos espectadores: 133 ppm para ouvintes (D'YDEWALLE *et al.* 1987), 145 ppm para surdos e ouvintes (JENSEMA, 1998), ambas as velocidades (145 e 180 ppm) para surdos (ARAÚJO; NASCIMENTO, 2011; ARAÚJO, 2012; MONTEIRO, 2016) e 180 ppm para surdos (SOUZA, 2020; VIEIRA, 2016; VIEIRA; TEIXEIRA; CHAVES, 2017). De Linde e Kay (1999), por sua vez, observaram que o processamento dos participantes surdos e ouvintes variou de acordo com a velocidade da legenda: quanto mais rápida a legenda (139 ppm), mais rápidas as fixações, ao passo que legendas lentas (74 ppm) podem induzir a releituras, o que pode ter indicado um processamento mais confortável de legendas de 139 ppm no estudo em questão.

De uma forma geral, os resultados a respeito da relação entre a velocidade da legenda e o processamento por espectadores surdos não são uniformes. Entretanto, de acordo com os resultados de Vieira (2016), surdos tendem a processar melhor legendas de velocidade rápida (180 ppm) por questões relacionadas à sua visão parafoveal mais desenvolvida. Dessa forma, para investigar a influência da velocidade e do número de linhas a partir das problemáticas acima delineadas, fizemos uso de uma metodologia experimental quantitativa



baseada na utilização de rastreamento ocular de modo a observar o processamento de participantes surdos e ouvintes em 8 trechos do filme *A rede social* (2010).

Os resultados de pesquisas anteriores (ARAÚJO, 2004; ARAÚJO, 2005; ARAÚJO, 2007; ARAÚJO, 2008; ARAÚJO, 2009; FRANCO; ARAÚJO, 2003), como já mencionado, apontam a necessidade de um número maior de pesquisas que proponham uma reflexão teórica e prática acerca da LSE veiculada no país, sendo que essa modalidade de TAV acessível (TAVa) ainda se encontra distante das demandas e necessidades do público surdo/ensurdecido. Sendo assim, a justificativa social da pesquisa em questão se apresenta na procura de melhoria para uma modalidade de TAVa que, junto à Libras, trabalha a inserção do indivíduo surdo na sociedade oferecendo acesso à informação e ao lazer.

Em relação à justificativa acadêmica, ou seja, o que esse trabalho pretende acrescentar em termos de lacuna no que diz respeito ao estado da arte da pesquisa em legendagem e, mais especificamente, em legendagem para surdos e ensurdecidos, está o estudo da influência do número de linhas de legendas (aliado à velocidade da legenda) na recepção de surdos brasileiros. Dentre as características técnicas e linguísticas da LSE já investigadas no Brasil por pesquisas exploratórias e experimentais, o parâmetro do número de linhas ainda não foi investigado. Em relação ao aspecto da velocidade, Vieira (2016) e Souza (2020) apontaram que legendas rápidas (180 ppm) proporcionam um processamento mais linear aos espectadores surdos dos seus estudos. Contudo, ressaltamos que esse parâmetro, ao ser analisado em outro gênero audiovisual (ficção, ao invés de documentário jornalístico e propaganda política), pode proporcionar diferenças em termos de processamento.

Outra questão está relacionada à pouca quantidade e à divergência dos dados de estudos experimentais (baseados em rastreamento ocular, principalmente), os quais foram supracitados, que se ocuparam da análise do parâmetro de número de linhas com foco no processamento e na questão da velocidade da legenda. Vale ressaltar que os estudos sobre o número de linhas (1 ou 2) foram realizados com participantes ouvintes apenas, o que coloca também o perfil dos participantes desta tese (surdos) como algo ainda não realizado.

Além disso, a maior parte desses estudos esteve mais focado em observar a alocação de atenção durante a exibição de vídeos legendados, e não o processamento destes *per se*<sup>7</sup>, ou seja, nenhum dos estudos isolou o parâmetro do número de linhas para investigação. Dessa

---

<sup>7</sup> Os estudos sobre rastreamento ocular e legendagem com foco na alocação de atenção se preocupam mais com o modo como a atenção (fixação) do espectador se distribui entre as áreas de interesse da imagem e da legenda do que o processamento linguístico em si.

forma, esta pesquisa também se propõe a isolar essa variável de modo a reforçar ou refutar os dados reportados pelas pesquisas supracitadas anteriormente.

Ademais, outra questão é o gênero de LSE aqui investigado, como já ressaltado acima: LSE de filmes, ou seja, ficção, que apresenta uma constituição linguística distinta dos estudos sobre LSE com auxílio do rastreamento ocular já realizados no Brasil (documentários e propagandas políticas). É esperado, por exemplo, que as legendas de documentários apresentem um número maior de sintagmas nominais em documentários do que em ficção, já que a natureza do documentário é descritiva. Mesmo não sendo objetivo específico da presente pesquisa verificar essas questões de gênero e LSE, é algo que deve ser levado em consideração quando os resultados forem contrapostos com o que já foi discutido na literatura da área, assim como estes resultados já serão levados em conta no Projeto ExLeg II, o qual pretende analisar a influência da recepção desses parâmetros a partir de diferentes gêneros audiovisuais com LSE.

Em relação à justificativa institucional da pesquisa, ou seja, a importância do estudo para o nicho acadêmico na Universidade Estadual do Ceará (UECE) e no Programa de Pós-Graduação em Linguística Aplicada (PosLA), é interessante salientar a relevância do Grupo LEAD (Legendagem e Audiodescrição), coordenado pela profa. Dra. Vera Lúcia Santiago Araújo e do qual faço parte desde a graduação, no que tange ao estudo e à promoção da LSE no país. Desde 2002, o grupo realiza pesquisas que visam à melhoria da experiência audiovisual de surdos por meio de legendas específicas para a comunidade surda, assim como a organização de eventos culturais, como mostras de cinema e peças de teatro acessíveis (tanto para surdos quanto para pessoas com deficiência visual por meio de audiodescrição), e a produção de DVDs acessíveis. Os resultados desta pesquisa fazem parte desse contexto, os quais irão reafirmar e fortalecer os estudos feitos pelo grupo em relação à LSE no país. Dessa forma, esta tese tem os seguintes objetivos:

Objetivo geral:

- a) analisar a influência do número de linhas e da velocidade no processamento de legendas por surdos e ouvintes.

Objetivos específicos:

- a) investigar o custo de processamento de legendas de 1 linha e 2 linhas em surdos e ouvintes;
- b) investigar o custo de processamento de legendas rápidas (180 palavras por minuto) e lentas (145 palavras por minuto) em surdos e ouvintes.

Esses objetivos específicos têm por objetivo responder às seguintes questões de pesquisa:

- a) Até que ponto legendas de 1 linha ou 2 linhas influenciam o processamento de legendas por surdos e ouvintes?
- b) Até que ponto legendas rápidas (180 palavras por minuto) ou lentas (145 palavras por minuto) influenciam o processamento de surdos e ouvintes?

Para essas perguntas e objetivos, baseados na literatura da área, formulamos as seguintes hipóteses:

- a) legendas de 1 linha induzem um processamento menos linear e, dessa forma, mais custoso do que legendas de 2 linhas;
- b) legendas rápidas (180 palavras por minuto) e lentas (145 palavras por minuto) não diferem significativamente no processamento de surdos e ouvintes.

A tese está organizada da seguinte forma para além desta introdução: na seção 2, discorremos sobre aspectos gerais da legendagem, enfatizando o que diz a literatura da área em relação aos parâmetros de número de linhas e velocidade, assim como sobre o processamento de legendas e a interface entre rastreamento ocular e estudos experimentais em legendagem. Na seção 3, relatamos o percurso metodológico da pesquisa, fornecendo detalhes acerca do contexto da pesquisa, da seleção e constituição dos materiais utilizados, do *design* experimental e dos procedimentos de análise. Na seção 4, apresentamos os resultados quantitativos por meio da estatística descritiva e inferencial para, então, confrontar os nossos achados com a literatura da área na discussão dos resultados. Por último, na seção 5, apresentamos a conclusão do presente estudo, enfatizando os resultados da pesquisa e como estes impactam o conhecimento sobre legendagem para surdos e ensurdecidos e sua prática. Além disso, apontamos possíveis desdobramentos e possibilidades de pesquisa futuras em relação ao tema estudado.

## 2 FUNDAMENTAÇÃO TEÓRICA

Nesta seção, tecemos considerações a respeito das áreas que conjugam para o embasamento teórico-metodológico de nosso trabalho: as concepções teóricas acerca da LSE de uma forma geral e a pesquisa com rastreador ocular na legendagem.

### 2.1 Legendagem para surdos e ensurdecidos (LSE)

A LSE é uma modalidade de TAVa destinada principalmente ao público que se filia a essa descrição (apesar de ouvintes também fazerem uso desse tipo de tradução quando não é possível ter acesso ao produto), permitindo que esta audiência tenha acesso, pela LSE, à informação veiculada por meio da língua oral e das informações adicionais, que são a identificação de falante e efeitos sonoros. A LSE pode ser tanto intralinguística, ou seja, da língua verbal oral para a língua verbal escrita, quanto interlinguística, de uma língua para outra, como nos casos em que é feita a partir de filmes e programas de TV em língua estrangeira (ARAÚJO; ASSIS, 2014). Ela também pode ser considerada uma modalidade de tradução intersemiótica quando falamos da tradução de sons (ou seja, um meio semiótico organizado que significa e gera interpretações recorrentes em sua audiência) para o texto escrito da legenda (ASSIS, 2016), pensado por Jakobson (1959) enquanto classificação para os tipos de tradução, assim como a interlinguística e a intralinguística já citados.

Díaz Cintas e Remael (2007), no que diz respeito aos parâmetros linguísticos da legendagem, tanto para ouvintes quanto para surdos, citam fatores como: redução do texto por condensação ou reformulação, omissões ou exclusões, coesão e coerência, segmentação e quebra de linhas, segmentação retórica e segmentação visual, que são importantes e devem ser considerados na confecção da legenda. Sobre as questões técnicas da legendagem, os autores as dividem em duas dimensões principais: espacial e temporal. Os aspectos espaciais estão relacionados ao número de linhas, à posição na tela, ao tipo de fonte e ao número de caracteres por linha. Os aspectos temporais, por sua vez, abarcam a marcação e duração das legendas, sincronização, intervalo entre legendas consecutivas, velocidade de leitura e tempo de exposição das legendas.

Por estarem mais relacionados à problematização deste estudo, darei um pouco mais de enfoque nos elementos que dizem respeito à velocidade de leitura da legenda e ao número de linhas.

## 2.2 Sobre a velocidade da legenda

A velocidade da legenda depende da velocidade da fala original e da capacidade dos leitores de lerem essa legenda confortavelmente (DÍAZ CINTAS; REMAEL, 2007, p. 95), assim como das necessidades do público-alvo e gênero (ROMERO-FRESCO, 2009), o que coloca a velocidade da legenda em uma relação com a velocidade de leitura do público ao qual a legenda está direcionada. Ou seja, a quantidade de texto que um legendista pode escrever em uma inserção de legenda a partir do tempo disponível para inserção (determinado pelo diálogo original do vídeo e pela marcação inicial e final da legenda) depende da suposta capacidade de leitura do público-alvo (DÍAZ CINTAS; 2008, p. 96). Contudo, a partir de fatores diversos, como idade e escolaridade, definir uma velocidade de legenda que contemple todos os telespectadores é complicado (DÍAZ CINTAS, 2008, p. 97). Além disso, diferentemente de textos estáticos, os espectadores não têm controle sobre a velocidade de apresentação das legendas (KRUGER; STEYN, 2014), o que faz com que dificuldades de processamento possam não ser esclarecidas devido ao fato de não poderem retornar ao texto da legenda.

Em legendas com padrão de velocidade mais alto, ou seja, mais rápidas, o legendista pode “se dar ao luxo” de escrever mais caracteres em uma inserção de legenda. Esse fato ocorre porque a quantidade de texto presente é uma relação entre o tempo de duração da legenda (definida pela marcação inicial e final da inserção) e o padrão de velocidade utilizado<sup>8</sup> (medida baseada em palavras por minuto, mas comumente transposta para caracteres por segundo em programas de legendagem) (ROMERO-FRESCO, 2009). E qual a relação dessa velocidade da legenda com a velocidade com que a fala ocorre em produtos audiovisuais? Romero-Fresco (2009) aponta que há uma diferença entre a velocidade da fala espontânea e a

---

<sup>8</sup> A relação tempo, quantidade de texto e padrão de velocidade em legendagem significa que legendas mais rápidas podem conter todas as palavras ditas em cena, enquanto em legendas mais lentas, dependendo também, claro, do padrão de velocidade da fala em vídeo (caso seja alto), o legendista tem que fazer uso de parâmetros de redução linguística (ROMERO-FRESCO, 2009, p. 110). Apesar de a redução linguística ser um parâmetro comum na legendagem interlinguística para ouvintes, é um assunto controverso quando se trata de LSE. Enquanto que as emissoras de TV, por exemplo, advogam pelo uso de legendagem *verbatim* (sem edição) por necessitar de menos custos para sua produção, assim como as associações de surdos (estas com motivação política, pois consideram que a redução é, também, uma forma de censura e nega ao espectador surdo a possibilidade de acessar integralmente o conteúdo audiovisual), os acadêmicos, por sua vez, são os únicos que defendem a edição das legendas por considerarem que legendas muito rápidas não proporcionam, normalmente, que os surdos leiam essas legendas por completo (ROMERO-FRESCO, 2009, p. 111-112). A visão acadêmica é avalizada por estudos de recepção que investigaram a preferência de espectadores surdos em relação às velocidades. Romero-Fresco (2009, p. 2012) cita um estudo conduzido por Sancho-Aldridge e o IFF Research LTD (1996, p. 24), o qual aponta que, na pesquisa em questão, ao serem perguntados sobre sua preferência, mais da metade dos surdos (54%) disse preferir legendas *verbatim*, enquanto 33% responderam preferir legendas editadas e 13% não manifestaram preferência. Contudo, após os pesquisadores pedirem aos respondentes que refletissem sobre as dificuldades potencialmente causadas pela leitura de legendas *verbatim*, as preferências se igualaram (43% para as legendas *verbatim* e 43% para legendas editadas).

velocidade da fala na TV. Citando os estudos realizados em língua inglesa, Romero-Fresco aponta os trabalhos de Steinfeld (1999) e Wingfield *et al.* (2006), os quais sinalizaram que essa fala espontânea é de 160 e 140-160 palavras por minuto (ppm), respectivamente. No caso da TV, normalmente a velocidade de fala é maior, chegando a 200 ppm em programas de notícias americanos (UGLOVA; SHEVCHENKO, 2006 *apud* ROMERO-FRESCO, 2009), sendo ainda mais rápida em noticiários do tempo (WINGFIELD *et al.*, 2006 *apud* ROMERO-FRESCO, 2009).

Romero-Fresco (2009, p. 113), citando Carver (1974), aponta que a taxa de leitura de um indivíduo não parece restrita a um número estático definido em escala, visto que variáveis como o nível de leitura dos materiais, o propósito da leitura e ainda o contexto conceitual do material são fatores que podem influenciá-la. Ainda segundo Carver (1974), o fator mais relevante a ser considerado para a definição dessa taxa deve ser a precisão e eficiência da compreensão.

Sobre as velocidades de leitura de textos impressos, Romero-Fresco cita Carver (1976), cujo estudo sugeriu uma velocidade que varia entre 315 e 200 palavras por minuto de acordo com a dificuldade do texto. Em relação à velocidade de leitura de legendas, Jensema (1998), ao realizar um estudo com 578 pessoas, entre surdos, ensurdecidos e ouvintes, observou que a velocidade preferida dos participantes foi a de cerca de 145 ppm. Para obter tal resultado, o autor expôs os participantes a legendas de 96, 110, 126, 140, 156, 170, 186 e 200 ppm, extraíndo respostas baseadas em uma escala de conforto que ia de 1 a 5, de “muito lento” a “muito rápido”. A maior parte das respostas que sugeriu uma velocidade de legenda por volta de 145 ppm como confortável esteve no espectro do “ok”, ou seja, a velocidade da legenda estava confortável. Contudo, são dados que precisam ser mais bem qualificados, visto que: “A velocidade leitora difere grandemente entre o público ouvinte, e isso se aplica a uma extensão ainda maior para os surdos e ensurdecidos<sup>9</sup>” (IVARSSON; CARROLL, 1998, p. 131), porque esse público é mais dado a sua heterogeneidade em termos de constituição e posicionamento, o que torna difícil conciliar esses aspectos em direção a uma velocidade de legenda (DE LINDE; KAY, 1999).

Antes de Jensema (1998), d’Ydewalle *et al.* (1987), em um estudo basilar da literatura em legendagem e que será detalhado mais adiante, testou três velocidades de leitura associadas à ppm com auxílio de rastreamento ocular: duas linhas de 32 caracteres em 4 segundos, cerca de 192 ppm, em 6 segundos, aproximadamente 130 ppm e 8 segundos, cerca

---

<sup>9</sup> “Viewer reading speeds differ greatly among hearing audiences and this applies to an even greater extent to the deaf and hard of hearing”.

de 96 ppm. Neste estudo, os autores tinham como objetivo investigar se o padrão de velocidade utilizado na Europa, também conhecido como “Regra dos 6 Segundos” (*The Six-Seconds Rule*), poderia se sustentar com uma pesquisa empírica, além de ser uma prática comum na área. Os seus resultados ratificaram esse tempo máximo de exposição<sup>10</sup> para legendas de duas linhas com 32 caracteres cada. Hoje em dia, porém, um número maior de caracteres é utilizado com base nessa regra dos seis segundos, cerca de 74 a 78 caracteres, com o uso de 74 caracteres divididos em 6 segundos de exposição promovendo um processamento de cerca de 145 palavras por minuto (DÍAZ CINTAS; REMAEL, 2007).

Tendo em mente a regra dos seis segundos, Díaz Cintas e Remael (2007) elaboraram, com a ajuda do programa *Wincaps* de legendagem, tabelas que relacionam caracteres por segundo (cps) e *frames* – quadros em sequência que criam a sensação de movimento nas gravações em vídeo – de forma que os legendistas fossem instrumentalizados a utilizar as três velocidades normalmente preconizadas pela indústria da legendagem, de 145, 160 e 180 ppm. Abaixo, na Tabela 1, a conversão da velocidade de 145 palavras por minuto em caracteres por segundo, sendo esta a medida utilizada por *softwares* de legendas e por profissionais da área. Pontuação e espaços também são considerados caracteres na contagem de uma legendagem.

---

<sup>10</sup> Ao passo que ficar por muito tempo em tela pode atrapalhar a recepção do telespectador, ficar por menos de 1 segundo pode fazer com que o telespectador não tenha tempo suficiente para captar a mensagem e fazer o cotejo com a imagem visual em tela, de modo que a legenda apenas “pisque” em tela (DÍAZ CINTAS; REMAEL, 2007). Além disso, um intervalo de pelo menos 100 milissegundos entre legendas é comumente utilizado de forma que as legendas não fiquem “encavaladas”.

**Tabela 1 – Número de caracteres por segundo para a velocidade de 145 ppm**

<b>145 palavras por minuto</b>	<b>Segundos: Quadros</b>	<b>Caracteres</b>	<b>Segundos: Quadros</b>	<b>Caracteres</b>
	01:00	16	02:00	29
	01:04	17	02:04	32
	01:08	18	02:08	34
	01:12	20	02:12	36
	01:16	23	02:16	38
	01:20	25	02:20	40
	<b>Segundos: Quadros</b>	<b>Caracteres</b>	<b>Segundos: Quadros</b>	<b>Caracteres</b>
	03:00	44	04:00	58
	03:04	46	04:04	60
03:08	48	04:08	62	
03:12	50	04:12	64	
03:16	52	04:16	65	
03:20	54	04:20	67	

Fonte: Adaptada de Díaz Cintas e Remael (2007, p. 97).

A velocidade de 145 ppm era a considerada ideal para a LSE nos produtos legendados por nosso grupo de pesquisa, estando de acordo com os resultados apresentados até então pelas pesquisas de recepção realizadas com o público surdo cearense (ARAÚJO, 2008). Contudo, o estudo de Araújo e Nascimento (2011) apontou que, com uma legenda bem segmentada, uma pessoa surda se torna capaz de ler em uma velocidade de até 180 ppm, sugerindo que o parâmetro da velocidade na compreensão do filme legendado pelos surdos não seria o de velocidade lenta.

Sobre as tabelas propostas por Díaz Cintas e Remael (2007), o trabalho de Carvalho e Seoane (2018) teve como objetivo investigar a correspondência entre as medidas de caracteres por segundo (cps), presentes nas orientações sobre a prática de legendagem da plataforma de *streaming* Netflix, e a de palavras por minuto, como aparece na literatura em legendagem na LSE, da segunda temporada da série *3%*. Os resultados apontaram uma não linearidade dessa conversão de medidas. A razão para isso seria o fato de que essa correspondência é baseada no tamanho médio das palavras em uma oração, e o tamanho médio das palavras em uma legenda pode variar. Dessa forma, com base nessa constatação de Carvalho e Seoane (2018), não utilizamos as tabelas propostas por Díaz Cintas e Remael (2007) à risca. Levando em conta o tamanho médio da palavra em português para o uso de caracteres em *softwares* de legendagem, chegamos ao número de 14 caracteres por segundo para a velocidade de 145 ppm e ao número



de 17 caracteres por segundo para a velocidade de 180 ppm<sup>11</sup>, apesar de reconhecermos que, em determinada legenda, pode haver tamanhos de palavras diferentes e essa lógica não ser exata.

Em relação ao número de linhas por legenda, Díaz Cintas e Remael (2007, p. 82), conforme citado anteriormente, afirmam que o número máximo a ser utilizado na legendagem interlinguística é o de duas linhas por legenda, sendo que a legendagem para surdos e ensurdecidos diferencia-se por permitir o uso de 3 ou até mesmo 4 linhas, como no caso do Teletexto, sistema espanhol de LSE. Essa característica também é seguida pelo sistema de legendagem fechada para surdos utilizado no Brasil – o *closed caption*, que faz uso de um alto número de legendas com 3 linhas, como visto em Assis (2013), Araújo e Assis (2014) e Gabriel (2015). Além deste, há também legendas bilíngues que usam 4 linhas de legenda, assim como legendas de fã de *animes* (e outros produtos audiovisuais asiáticos) que utilizam várias linhas de legenda em determinadas situações, como abertura, encerramento e algumas referências culturais<sup>12</sup> (DÍAZ CINTAS; SÁNCHEZ, 2006).

---

<sup>11</sup> A correspondência de número de caracteres utilizado nos *softwares* de legendagem é confusa. Segundo Díaz Cintas (2008), visto que o tamanho médio da palavra em língua inglesa é de 5 caracteres, uma velocidade de 140 a 150 palavras por minuto daria uma média de 12 caracteres por segundo, ao passo que a velocidade de 180 ppm seria de 17 caracteres por segundo. É possível observar, contudo, na sua tabela sobre a velocidade de 145 ppm, que o autor começa sua contagem de 1 segundo com 16 caracteres (DÍAZ CINTAS; REMAEL, 2007). Como observado por Carvalho e Seoane (2018), entretanto, essa tabela não é linear; ao passo que o tempo de exposição da legenda aumenta, o número de caracteres por segundo considera tamanhos diferentes de palavra para essa transposição. É fato que Díaz Cintas e Remael (2007) consideram também o número de espaços e pontuação nessa métrica, mas, como isso é calculado em relação ao tamanho médio da palavra, não é uniforme e não é explicitado pelos autores. Também sobre isso, Romero-Fresco (2009) menciona que a velocidade de 180 ppm é na verdade relativa a 15 caracteres por segundo (15 cps), enquanto a velocidade de 160 ppm seria relativa a 13 cps e a de 190, a 16 cps. Além disso, o autor cita uma publicação do Escritório de Comunicações (tradução literal de *Office of Communications*), órgão britânico de regulação sobre telecomunicações no Reino Unido, que considera a velocidade de 180 palavras por minuto como sendo 15 cps. Essa pequena digressão sobre a relação caracteres por segundo × palavras por minuto nos deixou confusos para a manipulação do experimento do trabalho. Como tínhamos que tomar uma decisão, levamos em consideração que o tamanho médio da palavra em português é similar, de 4,64 caracteres (QUARESMA; PINHO, 2007) e arredondamos para 5. Dessa forma, obtivemos 12 caracteres por segundo para 145 ppm (2,41 palavras por segundo) e 15 caracteres por segundo para 180 ppm (3 palavras por segundo). Além disso, também levando em consideração que os *softwares* de legendagem contabilizam os espaços e pontuação como caracteres, projetamos que pelo menos dois espaços serão adicionados em um intervalo entre 3 palavras na velocidade 180 ppm, da primeira palavra para a segunda e da segunda para a terceira (passando assim para 17 caracteres por segundo), sem contabilizar possíveis pontuações que finalizam uma sentença. Para a velocidade de 145 ppm, a partir da noção de que corresponde a 2,41 palavras por segundo, aumentar em dois caracteres a partir do intervalo entre palavras não é uma conta tão precisa como em 180 ppm. Esse aumento foi também considerado por efeitos de aproximação (dado que não existe a possibilidade de colocar um caráter e meio que continuaria com a mesma proporção). Dessa forma, para a velocidade de 145 ppm, adotamos 14 caracteres por segundo, aproximadamente.

<sup>12</sup> Em aberturas e encerramentos de *animes*, por exemplo, é comum que os *fansubbers* (fãs legendistas, ou seja, legendistas amadores) coloquem o efeito de karaokê nas músicas, além das versões das músicas em *romaji* (transliteração do japonês para o alfabeto latino) e no sistema de escrita japonês. Além disso, também é comum a adição de notas sobre referências culturais, que normalmente aparecem na parte de cima da tela ao mesmo tempo que a legenda à qual faz referência, tornando a leitura da legenda dificultosa (DÍAZ CINTAS; SANCHEZ, 2006).

D'Ydewalle *et al.* (1987) foi o primeiro estudo a investigar a influência do processamento de legendas de 1 linha e 2 linhas. Apesar de não utilizarem tantas medidas de processamento, e sim mais de alocação de atenção, nem isolarem a variável em questão, os autores observaram que a leitura de legenda de 2 linhas era mais fluida a partir do tempo de latência e duração média de fixações dos participantes. Esses dados foram reforçados em pesquisa posterior, relatada em d'Ydewalle *et al.* (1991). Os resultados sobre o tempo de latência nos dois estudos, contudo, chamaram a atenção dos pesquisadores: apesar de um tempo menor proporcionalmente de leitura de legendas em 2 linhas, os participantes demoravam mais tempo para mudar seu foco de atenção da área da imagem para a área da legenda (tempo de latência) nas legendas de 2 linhas do que de 1 linha. Sobre isso, d'Ydewalle *et al.* (1991) especularam que haveria algum tipo de interferência lateral nas legendas de 2 linhas, que faria com que os espectadores demorassem mais a começar o processamento da primeira palavra da primeira linha, assim como demorariam mais por causa do “pulo” entre o final da segunda linha e o começo da primeira.

Praet *et al.* (1990), contudo, observaram que não há essa interferência lateral. Ademais, relatam em seu trabalho que as legendas de 2 linhas são menos custosas ao processamento por serem mais completas sintaticamente do que legendas de 1 linha. Entretanto, os resultados divergiram quando as legendas foram apresentadas com e sem a imagem no experimento, fazendo com que os participantes tivessem um desempenho melhor com legendas de 2 linhas sem a presença da imagem e com legendas de 1 linha acompanhadas da imagem. Isso porque, segundo os autores, há uma maior redundância de informação nas legendas de 1 linha com as imagens e das legendas de 2 linhas sem a imagem, ocasionando uma maior facilidade de leitura. Assim, os autores apontam que a redundância de informação seria o facilitador desse processo de leitura.

Kruger e Steyn (2014), por sua vez, apesar de não terem como objetivo específico o estudo do processamento de legendas com número de linhas distintos, observaram uma covariância em relação a esse aspecto em apenas um dos vídeos analisados em seu trabalho (dentre 6), questionando o impacto do número de linhas no processamento de legendas, o qual “[...] pode ter sido exagerado ou ser aplicável apenas em situações em que participantes estavam acostumados com a leitura de um grande número de legendas<sup>13</sup>” (KRUGER; STEYN, 2014, p. 114). Esse trabalho será retomado mais adiante, quando falarmos da metodologia para o estudo de processamento ocular em legendagem. Os resultados de d'Ydewalle e De Bruycker

---

<sup>13</sup> “May have been overstated or only applicable to a situation in which participants were used to reading a lot of subtitles”.

(2007), por sua vez, ressaltaram um processamento mais fluido em legendas com 2 linhas, assim como em Praet *et al.* (1990). Os resultados desse trabalho, assim como os de Praet *et al.* (1990), serão detalhados mais adiante.

Nesse contexto de parâmetros de legendagem, o Grupo LEAD-UECE advoga a favor de uma LSE pautada pelo uso dos mesmos parâmetros técnicos e linguísticos da legendagem para ouvintes – diferenciando-se apenas pela adição de informações adicionais já mencionadas. Além do mais, sendo respaldado pelos achados de d’Ydewalle *et al.* (1987) sobre o número máximo de linhas por inserção de legenda, essa característica (de até 3 linhas por legenda) tornaria a LSE praticada na TV do Brasil inadequada em relação a esse aspecto técnico específico, dentre outros. A LSE de filmes em DVD no país, contudo, costuma respeitar o limite de duas linhas sugerido pela literatura da área, conforme visto, por exemplo, no trabalho de Chaves (2012).

De um modo mais geral, a partir das pesquisas de recepção com surdos feitas pela UECE (ARAÚJO, 2008; ARAÚJO, 2012; ARAÚJO; NASCIMENTO, 2011), algumas considerações sobre os parâmetros de legendagem foram estabelecidas para um modelo de LSE no Brasil: i) a cor deve ser, preferencialmente, amarela em um fundo transparente; ii) a identificação de falantes e os efeitos sonoros devem aparecer entre colchetes; iii) a voz filtrada (voz fora da tela, vinda da televisão, vinda do rádio etc.) deve apresentar-se em itálico; e iv) a linguagem utilizada deve ser o português formal. Contudo, esses parâmetros não são comumente colocados em prática na LSE que é veiculada pelos canais de TV brasileiros. No entanto, quando a mídia muda, como no caso dos filmes em DVD, as LSEs seguem o padrão da legendagem para ouvinte. Provavelmente isso acontece porque essas legendas de programas de TV – geralmente ao vivo – são feitas por profissionais legendistas, e não estenotipistas<sup>14</sup> ou por programas de refalamento<sup>15</sup> (*respeaking*), os quais não são geralmente realizadas por tradutores.

---

<sup>14</sup> Profissional que digita em tempo real, em posse de um estenótipo, teclado especial de 24 teclas em que a pressão de várias destas ao mesmo tempo pode formar uma palavra.

<sup>15</sup> *Software* de reconhecimento de voz que transforma a voz de profissionais treinados para tal tarefa em texto escrito (no caso, legendas).

## 2.2 Legendagem e rastreamento ocular

### 2.2.1 O rastreamento ocular

Mas por que o rastreamento ocular é importante? Respondendo a essa pergunta, Duchowski (2007, p. 3) afirma que movemos o nosso olho em direção a algum ponto de modo a aumentar a resolução daquele objeto/lugar e captar mais detalhes do que está na centralidade do nosso olhar, esta chamada de área foveal. Desse modo, também transferimos nossa atenção àquele objeto ou área de interesse. Sendo assim, captar a movimentação do percurso ocular de alguém “[...] pode nos dar algum *insight* sobre o que o observador achou interessante, ou seja, o que chamou sua atenção e talvez até prover uma pista de como aquela pessoa percebeu qualquer que seja a cena que ela estava vendo<sup>16</sup>” (DUCHOWSKI, 2007, p. 3), mesmo que essas fixações não signifiquem, necessariamente, um processo cognitivo (TOBII STUDIO, 2010, p. 11).

O campo de visão humana se estende por 220 graus e está dividido em três regiões principais: a foveal, a parafoveal e a periférica (TOBII STUDIO, 2010, p. 4). A área foveal é a principal parte responsável por registrar a informação visual e, apesar de compreender apenas cerca de 8% do campo visual, constitui 50% do que é mandado ao cérebro através dos nervos óticos. O rastreador ocular só pode identificar o movimento explícito do olho, pois o centro do foco que está na região foveal não registra a região periférica da visão, sendo este um problema da pesquisa com o rastreador ocular: a atenção pode estar nessa região periférica, como nos casos dos astrônomos que propositalmente observam as estrelas com baixa luminosidade com a área periférica da visão, já que esta parte se faz mais precisa na percepção de objetos opacos e escuros. Desse modo, os pesquisadores assumem a suposição de que o movimento ocular foveal está ligado à atenção, mas reconhecem que às vezes isso pode não acontecer (DUCHOWSKI, 2007, p. 12).

A utilização de rastreadores oculares começou a se popularizar no meio dos anos 1970, marcando uma nova fase da pesquisa em movimentação ocular no processamento de textos alavancada por melhorias tecnológicas nos sistemas de monitoramento da movimentação ocular, os quais se tornaram mais precisos e mais fáceis de ser obtidos. Segundo Rayner (1998), esses avanços também conseguiram integrar as interfaces de computadores com os sistemas de *eye-tracking*, fazendo com que uma grande quantidade de dados pudesse ser analisada. Além

---

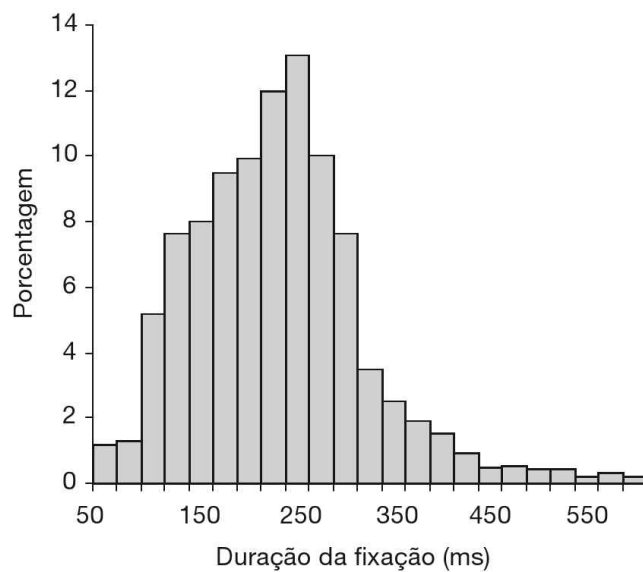
<sup>16</sup> “[...] may give us some insight into what the observer found interesting, that is, what drew their attention, and perhaps even provide a clue as to how that person perceived whatever scene she or he was viewing”.

disso, o desenvolvimento de teorias de processamento também fez com que os registros de movimentação ocular pudessem ser usados para a compreensão dos processos cognitivos ativados durante a leitura (RAYNER, 1998).

Ainda segundo Rayner (1998), quando olhamos determinada cena, nossos olhos não param, estão em movimento continuamente. A esses movimentos chamamos **sacadas**. Estas são movimentos rápidos que possuem uma velocidade de até 500° por segundo, duram cerca de 20-40 milissegundos e têm um tamanho médio de aproximadamente 7 a 8 espaços durante a leitura (RAYNER; BARBARA; POLLATSEK, 2013). Normalmente, não conseguimos obter informações novas durante esse processo, já que a nossa acuidade visual durante essa movimentação não é tão precisa e, graças a essa velocidade alta, apenas um borrão seria percebido. Contudo, por causa da informação que normalmente está presente antes e depois da sacada, não percebemos esse borrão. A todo esse fenômeno chamamos de **supressão sacádica** (RAYNER, 1998).

Segundo Rayner (1998), a função das sacadas é trazer alguma área para a visão foveal, fazendo dessa área objeto de análise mais detalhada. Entretanto, nem todas as palavras são fixadas durante a leitura. De acordo com os estudos de Carpenter e Just (1983) e Rayner e Duffy (1983), citados por Rayner (1998), palavras de conteúdo, por exemplo, são fixadas em 85% das vezes, enquanto palavras de função, apenas 35% das vezes. Rayner e McConkie (1976), também citados por Rayner (1998), apontam que palavras de função normalmente são menores e que há uma relação entre o tamanho da palavra e sua fixação, visto que quanto maior a palavra, menor a chance de que ela seja pulada; palavras de 2-3 caracteres, por exemplo, são fixadas apenas em 25% das ocasiões, enquanto palavras de 8 letras ou mais são quase sempre fixadas e, em muitas vezes, mais de uma vez.

Chamamos de **fixação** o momento em que os olhos estão relativamente estáveis entre as sacadas e é durante esse momento que os leitores extraem informação das palavras e textos. Essas fixações duram cerca de 200-250 milissegundos, mas existe uma variação nesse número, conforme pode ser visto na Figura 1 abaixo, em que um histograma mostra uma variação de 50 a 550-600 milissegundos na frequência relativa de fixações.

**Figura 1 – Distribuição de duração de fixações**

Fonte: Rayne, Barbara e Pollatsek (2013, p. 99).

Em inglês, assim como na maior parte das línguas ocidentais, movemos nossos olhos da esquerda para a direita (RAYNER; BARBARA; POLLATSEK, 2013), seguindo a ordem sintática da língua. Segundo Rayne, Barbara e Pollatsek (2013), ao chegarmos ao final da linha, realizamos um movimento de **varredura de retorno** (*return sweep*) para a linha seguinte. Geralmente, esse movimento é impreciso e faz com que o pouso do olhar recaia em um lugar além do começo da linha, o que é seguido por uma sacada de correção da esquerda para a direita. O movimento de varredura de retorno também está presente na legendagem, por exemplo, quando observamos o retorno da linha de cima para a linha de baixo em uma legenda de 2 linhas, podendo ser um movimento vertical de cima para baixo (de uma linha para outra da legenda) que seja maior que a distância entre as duas linhas da legenda. Isso fez com que essas fixações de retorno imprecisas fossem desconsideradas em nossa pesquisa, pois como não são seguidas de uma movimentação da esquerda para a direita, não representam um processo de leitura, mas sim de ajuste, um pouso na área da legenda. Dessa forma, essas fixações foram descartadas dos nossos dados, o que só pode ser feito através de um acompanhamento do olhar dos participantes expostos ao experimento (VIEIRA, 2016).

No Quadro 1, abaixo, Rayner (1998) aponta a duração média aproximada das fixações em algumas atividades de leitura e reconhecimento de cena, assim como o tamanho médio dessas sacadas.

**Quadro 1 – Duração média das fixações e comprimento das sacadas por tarefas**

Tarefas	Média de duração das fixações (milissegundos)	Tamanho médio das sacadas (graus)
Leitura silenciosa	225	2° (cerca de 8 letras)
Leitura oral	275	1,5° (cerca de 6 letras)
Pesquisa visual	275	3°
Percepção de cena	330	4°
Leitura de música	375	1°
Digitação	400	1° (cerca de 4 letras)

Fonte: Rayner (1998, p. 373). Traduzido por Vieira (2016).

De acordo com a descrição de tarefas no Quadro 1, a leitura silenciosa, com média de duração de 225 milissegundos, seria a mais próxima ao contexto da leitura dinâmica da legendagem. Especificamente em relação à legendagem, d'Ydewalle *et al.* (1985) estimaram um tempo médio de 124 ms por palavra de adultos expostos a vídeos legendados, um tempo abaixo até do tempo estimado por Rayner (1998) para leitura silenciosa de textos estáticos. Há vários aspectos textuais que influenciam esse tempo de processamento durante a leitura de textos estáticos, como a complexidade do próprio texto, o propósito de leitura, o público-alvo, dentre outros. No caso de um texto dinâmico e multimodal como um vídeo legendado, a duração de fixação ainda pode ser influenciada por outros fatores, como o gênero audiovisual, a movimentação de cena, os fatores relacionados à imagem e ao som, os quais também acompanham o texto verbal escrito na legenda, dentre outros.

Os estudos sobre processamento de surdos e ouvintes por meio de rastreamento ocular no Brasil tem indicado tempos na faixa de 250 a 300 milissegundos para essa medida (MONTEIRO, 2016; SOUZA, 2020; VIEIRA, 2016).

Uma outra característica da movimentação ocular durante a leitura é a de **regressão**. Rayner, Barbara e Pollatsek (2013) afirmam que as regressões são movimentos para trás durante a leitura, voltando para palavras anteriormente processadas. Ainda segundo os autores, as regressões compreendem de 10 a 15% das fixações que os leitores proficientes realizam e geralmente são curtas, sendo o movimento de sacada para a esquerda de 2 a 3 palavras anteriores. Segundo Rayner e Liversedge (2004), essas regressões podem compreender palavras que foram anteriormente fixadas ou puladas e seriam importantes para a compreensão do texto.

Rayner, Barbara e Pollatsek (2013) apontam que as regressões ainda não são muito bem compreendidas, mas que normalmente estão relacionadas a um custo de processamento. Sobre o papel da regressão na legendagem, Kruger e Steyn (2014) afirmam que as regressões não devem ser tão consideradas na investigação do processamento de legendas, pois, diferentemente de textos estáticos, os telespectadores não possuem controle sobre a

apresentação das legendas. Dessa forma, os telespectadores não possuem controle sobre a regressão em palavras que devem ser desambiguadas ou então palavras que não foram fixadas. De qualquer forma, decidimos incluir a medida de regressão no nosso trabalho, pois há na literatura registros de estudos sobre o processamento de legendas que utilizaram a regressão como forma de obter indícios sobre o processamento de legendas. Estamos conscientes de que essa contextualização sobre a relação entre essa medida e a especificidade da legendagem enquanto texto multimodal dinâmico também devem ser levadas em conta.

A **releitura** também é uma medida utilizada na psicolinguística como indicador de custo de processamento (RAYNER; LIVERSEDGE, 2004), assim como na legendagem (VIEIRA, 2016). Segundo Rayner e Liversedge (2004), a releitura, como o nome já deixa bem explícito, diz respeito ao tempo que o leitor leva para reler uma porção de texto já lida; ou seja, após finalizar um segmento, o leitor faz uma sacada à esquerda para o início desse segmento para reler todo o trecho ou partes específicas do texto, indicando um custo de processamento daquele trecho. Na legendagem, apesar de o telespectador não ter controle sobre o tempo de apresentação da legenda, a releitura acontece quando o telespectador lê toda a legenda e ela permanece na tela. Também é possível que esse movimento de releitura aconteça após o telespectador voltar à imagem e, pelo fato de a legenda ainda estar lá, seu olhar volta à legenda que já havia sido lida. Este outro exemplo de movimento de releitura também é conhecido como **revisita** (*revisit*) ou um dos tipos de deflexão.

Existem duas perspectivas sobre as *revisits* no contexto dos estudos de processamento. Se por um lado a revisita é vista como algo que denota a informatividade semântica de uma área (HOLMQVIST *et al.*, 2011 *apud* SZARKOWSKA; GERBER-MORÓN, 2018), por outro ela é vista como algo que pode estar associado a uma dificuldade de processamento, dada a necessidade de se voltar a uma área para checar novamente (CONKLIN; PELLICER-SÁNCHEZ, 2016 *apud* SZARKOWSKA; GERBER-MORÓN, 2018). Sobre isso, Szarkowska e Gerber-Morón (2018) afirmam que essa revisita no contexto de vídeos legendados mostra como os telespectadores dividem sua atenção entre o ato de processar a legenda e a imagem. Eles afirmam ainda que um alto número de revisitas pode indicar um processamento menos fluente e eficiente.

O movimento de **deflexão**, que pode ter outros nomes na literatura, conforme será detalhado na próxima subseção, diz respeito ao movimento de mudança de atenção entre legenda e imagem durante a leitura de um mesmo bloco de legenda. Segundo Kruger, Szarkowska e Krejtz (2015), esse movimento, apesar de ser parte do padrão de leitura de ensurdecidos, por exemplo, é normalmente um indicativo de uma leitura não tão fluida. Isso



porque, segundo Jensema *et al.* (2000) a partir de contribuições dos estudos com rastreamento ocular, a tendência é que os olhos comecem a observar a imagem na parte central da tela, direcionem-se ao começo da legenda e, após a finalização da leitura, retornem à imagem. Vieira (2016), entretanto, ao triangular diversas medidas de processamento de surdos e ouvintes expostos a um documentário com LSE em português, observou que a condição experimental na qual os surdos mais fizeram deflexões<sup>17</sup> – 180 ppm e bem segmentada – também foi a que apresentou o melhor índice de duração média de fixação, apontando um conforto maior no processamento desses participantes nesse contexto.

Entretanto, visto que a maior parte da literatura aponta para o caráter digressivo da deflexão em contraste com o apontado por Vieira (2016), também partiremos do pressuposto inicial de que se trata de um fator de dificuldade para o processamento, ao passo que também iremos contextualizar esse fenômeno com as outras medidas analisadas e a literatura da área de uma forma geral. Todavia, diferentemente da autora, utilizaremos apenas a noção de deflexão enquanto movimento de volta à legenda após ida à imagem antes da finalização da leitura da legenda.

O leitor de textos estáticos, diferentemente de alguém que está apenas escutando algo ou, no nosso caso, lendo um texto dinâmico, pode controlar o ritmo da leitura, ou seja, pular partes do texto, reler alguma seção ou pausar em alguma palavra (JUST; CARPENTER, 1980). Além disso, de acordo com Just e Carpenter (1980), o leitor absorve a informação do texto em um ritmo que se encaixe com seus processos internos de compreensão. Dessa forma, é possível investigar os processos de compreensão de leitores ao examinarmos onde pausam durante a leitura.

Esse modelo de compreensão leitora delineado por Just e Carpenter (1980) foi feito a partir da duração de fixações de estudantes universitários lendo textos científicos. Apesar de ter sido realizado com textos estáticos, o seu modelo estabeleceu aspectos importantes que norteiam o estudo das fixações do olhar como meio de se compreender os processos de leitura e, por isso, foi considerado como relevante para este trabalho.

O Modelo de Just e Carpenter (1980) a partir desse estudo se baseia em dois princípios: i) Princípio da Imediaticidade (*Immediacy Assumption*) e ii) Princípio da Ligação Olho-Mente (*Eye-Mind Assumption*). Por um lado, o Princípio da Imediaticidade postula que o leitor tenta interpretar as palavras de conteúdo de um texto à medida que as encontra, mesmo

---

<sup>17</sup> A autora, no caso, interpretou as deflexões como movimentos de volta à legenda após ida à imagem tanto antes da finalização da leitura das legendas, como após a finalização dessa leitura, o que poderia ser encaixado como revisita, de acordo com Szarkowska e Gerber-Morón (2018).

que, às vezes, essa interpretação esteja errada. O significado de interpretação, nesse caso, está relacionado a esse processamento em vários níveis, como na decodificação de uma palavra, na seleção do seu significado, na ligação com o seu referente, na determinação do seu *status* na sentença e no texto de uma maneira mais ampla. Sendo assim, esse modelo aponta que tudo isso acontece ao mesmo tempo em que a palavra é processada. Por outro lado, o Princípio da Ligação Olho-Mente refere-se à noção de que “[...] o olho permanece fixado em uma palavra enquanto ela está sendo processada<sup>18</sup>” (JUST; CARPENTER, 1980, p. 1980). Nesse sentido, a duração de fixação em uma palavra indica o tempo do seu processamento, e, obviamente, essa compreensão também depende da informação no texto que precede aquela palavra; ou seja, não existe atraso entre o que está sendo fixado e o que está sendo processado. Tendo isso em mente, palavras menos frequentes tendem a gerar um tempo maior de duração de fixação, assim como palavras de função são menos fixadas do que palavras de conteúdo.

O presente modelo de leitura está dividido nas seguintes etapas: chegada do olhar ao próximo *input* (palavra), decodificação e acesso lexical, estabelecimento de caso da palavra, integração entre orações e encapsulamento da sentença<sup>19</sup>. Just e Carpenter conceituam o encapsulamento (*sentence wrap-up*), uma medida também trabalhada aqui, como um estágio do processamento no final das sentenças. Segundo os autores, o encapsulamento envolve uma busca por referentes que não foram designados e por relações entre as orações, sendo uma tentativa de resolver algo que não ficou claro dentro da sentença. Esse estágio da leitura tem dois aspectos importantes para o processamento: i) ambiguidades que possam surgir dentro da sentença são resolvidas no final da sentença, normalmente; e ii) o final da sentença sinaliza o fim de um pensamento e o começo de outro. Os estudos sobre o tempo de leitura de textos estáticos têm enfatizado que há a necessidade de um tempo extra para esse encapsulamento. Além disso, estudos sobre fixação relatam que “[...] quando uma inferência baseada no léxico deve ser feita para que se relacione uma nova sentença a alguma parte anterior do texto, há uma forte tendência de que haja uma pausa no último item lexical em questão e no final da sentença que o contém<sup>20</sup>” (JUST; CARPENTER, 1978 *apud* JUST; CARPENTER, 1980, p. 346). Contudo, também é possível que esse encapsulamento aconteça em segmentos menores ou

<sup>18</sup> “[...] eye remains fixated on a word as long as the word is being processed”.

<sup>19</sup> Segundo Castilho, a “designação da sentença não é pacífica na literatura” (CASTILHO, 2012, p. 58), sendo assim possível encontrar “[...] termos tais como oração, frase, período (conjunto de orações) etc.” (CASTILHO, 2012, p. 58). Ao longo do trabalho, entretanto, usamos sentença como sinônimo de período por acreditarmos ser uma melhor tradução para o sentido de *sentence* utilizado por Just e Carpenter (1980).

<sup>20</sup> “[...] when a lexically based inference must be made to relate a new sentence to some previous portion of the text, there is a strong tendency to pause at the lexical item in question and at the end of sentence that contains it”.

maiores do que uma sentença, como no final de orações ou de parágrafos. Essa decisão sobre o encapsulamento, onde e se fazê-lo, pode ser controlada pelo nível de processamento desejado, ou seja, depende do propósito de leitura.

No contexto da legendagem, o telespectador não possui o controle sobre a apresentação do texto, o que faz com que não lhe seja permitido pausar apenas no final das sentenças para resolver alguma ambiguidade interna da própria sentença ou relacionada a alguma informação anterior, no texto. Por essa razão, acreditamos, no contexto da legendagem, que esse encapsulamento é, na verdade, parcialmente induzido pelo legendista a partir da divisão do texto apresentado em tela de acordo com os parâmetros técnicos e linguísticos da legendagem. Percebemos, durante o acompanhamento da movimentação ocular dos participantes no presente estudo, que um tempo maior de fixação ocorria com frequência na última fixação da legenda – sendo esta de 1 ou 2 linhas. Como nem sempre uma legenda de 1 linha – ou até mesmo 2 linhas – é uma sentença completa, acreditamos que o encapsulamento em legendagem pode ocorrer em unidades menores do que uma sentença, conforme possibilidade apontada por Just e Carpenter (1980).

Nesse sentido, também percebemos que a fixação de encapsulamento ultrapassa o tempo de exposição da legenda em tela. Isso quer dizer que a última fixação na legenda, seja ela de 1 ou 2 linhas, significa um tempo de fixação após o desaparecimento da legenda. Durante a análise das fixações em nosso estudo, esse contexto de encapsulamento em legendagem sugeriu que legendas de 1 linha induzem um tempo maior de encapsulamento porque há necessidade de se realizar mais encapsulamentos do que as legendas de 2 linhas, principalmente porque legendas de 1 linha, por muitas vezes, não são sentenças completas. Isso ocasiona um tempo extra de fixação após o desaparecimento da legenda e retarda o movimento de sacada para o começo de uma nova legenda que surge. No caso de uma sequência de legendas de 1 linha, como na nossa manipulação, pode ocasionar um custo maior de processamento e, dessa forma, um efeito *spillover*<sup>21</sup>. Esse conceito, contudo, entra um pouco em conflito com o modelo de leitura de Just e Carpenter (1980) na medida em que este coloca a ênfase do custo de

---

<sup>21</sup> *Spillover* (“transbordamento”) em psicolinguística diz respeito ao tempo de fixação de uma palavra relacionado ao contexto linguístico em que aparece, ou seja, uma palavra anterior que ocasiona um tempo maior de fixação na palavra seguinte (RAYNER; LIVERSEDGE, 2004). No contexto da legendagem, a partir da nossa hipótese, legendas de 1 linha, por conterem uma completude sintática menor e, dessa forma, causarem um encapsulamento em unidades sintáticas menores e, possivelmente, mais ambíguas, teriam o potencial de ocasionar esse efeito de *spillover*. Dessa forma, legendas de 1 linha poderiam aumentar o custo de processamento das unidades de texto que as seguem. Esse conceito, entretanto, não será aprofundado em nosso trabalho.

processamento na palavra, enquanto o efeito *spillover* aponta a influência de palavras precedentes no tempo de fixação.

A **latência**, por sua vez, no contexto da legendagem, é mais uma medida de alocação de atenção do que um índice de processamento de leitura *per se*. De uma forma geral, a latência diz respeito ao tempo de intervalo de alocação de atenção entre estímulos distintos; por isso, também é conhecido como latência sacádica (REICHLE; RAYNER; POLLATSEK, 2003). Segundo Rayner (1998), esse tempo associado à sacada é explicado pelos movimentos motores necessários para sua realização, os quais necessitam de planejamento para serem executados. Ainda sobre isso, Rayner (1998) aponta que, mesmo que o fator de incerteza seja desconsiderado, um movimento de latência dura no mínimo entre 150 e 175 milissegundos. Westheimer (1954), citado por Just e Carpenter (1980), aponta um tempo estimado entre 150 e 200 milissegundos em relação a um estímulo visual que tem uma incerteza temporal e espacial. Essa descrição pode ser encaixada com o nosso objeto de estudo – a legendagem, visto que o momento em que a legenda irá aparecer, apesar de certa forma esperado, é sempre incerto. Nesse sentido, um tempo grande de latência entre imagem e legenda pode prejudicar o processamento do texto dinâmico em questão porque, ao demorar muito, o tempo de exposição da legenda em tela pode não ser suficiente para o telespectador, especialmente no caso de legendas curtas e rápidas. Foi essa a questão que fez d’Ydewalle *et al.* (1987) apontarem a legenda de 2 linhas como mais confortável do que as legendas de 1 linha. No quesito latência, as legendas de 2 linhas tiveram um tempo proporcional de latência menor do que as legendas de 1 linha.

### 2.2.2 O processamento de legendas

A tecnologia baseada em rastreamento ocular tem sido efetiva ao dizer para “[...] onde as pessoas estão olhando, mesmo em situações semioticamente densas, como a recepção de um filme legendado<sup>22</sup>” (PEREGO, 2012, p. 7). E por que semioticamente denso? Devido ao seu caráter multimodal, o processamento e percepção de legendas demanda do telespectador atenção alocada a vários dos componentes presentes em um produto audiovisual simultaneamente (GHIA, 2012, p. 158).

---

<sup>22</sup> “[...] where people are looking, even in situations which are semiotically dense, such as the viewing of a subtitled film”.

Dada essa característica fundamentalmente específica do vídeo legendado, Kruger, Szarkowska e Krejtz (2015) apontam que, apesar da tradição no estudo de leitura através de comportamento ocular em textos estáticos:

Estudos sobre a leitura de legendas, contudo, não atingiram o mesmo rigor científico em grande parte por questões práticas: legendas não são estáticas por mais do que alguns segundos por vez; elas competem pela atenção visual com a imagem em movimento; e elas competem pelos recursos cognitivos no geral com sons verbais e não verbais<sup>23</sup> (KRUGER; SZARKOWSKA; KREJTZ, 2015, p. 1).

Segundo os autores, a questão da imagem dinâmica realmente é um dos principais complicadores no que diz respeito ao estudo do processamento de legendas. Esse contexto faz com que as estatísticas automáticas de leitura geradas por *softwares* integrados a rastreadores oculares sejam quase inúteis, fazendo com que a análise de dados nesse tipo de pesquisa se torne um processo doloroso e lento, cheio de inspeções manuais. Diferentemente dos textos estáticos, em que o ritmo de leitura é em grande parte ditado pelo leitor, a leitura de legenda é imposta pela própria legenda ao espectador, de forma contínua, sem que haja muito tempo para regressões – mesmo com algumas exceções, em que o leitor consegue controlar mais ou menos o surgimento das legendas, como no caso de DVDs e outros tipos de mídia (KRUGER; SZARKOWSKA; KREJTZ, 2015).

Ainda segundo os autores, apesar de a integração de canais de informação facilitar a contextualização e, dessa forma, o processamento do vídeo legendado, não deixa de haver uma competição por alocação<sup>24</sup> de atenção e cognição. Enquanto a atenção de surdos e ensurdecidos é dividida entre imagem, legenda e informações não verbais como leitura labial, gestuais etc. (uma carga adicional para cognição), o ouvinte tem que dividir sua atenção entre imagem, som e legenda, além de outras características não verbais da fala.

Jensema *et al.* (2000) apontam que a aparição das legendas transforma o processamento de um vídeo de um processo de visualização para, principalmente, um processo de leitura. Dessa forma, os olhos se fixam na imagem não só porque o texto é fonte de

---

<sup>23</sup> “Studies on subtitle reading, however, have not achieved the same level of scientific rigour largely for practical reasons: subtitles are not static for more than a few seconds at a time; they compete for visual attention with a moving image; and they compete for overall cognitive resources with verbal and non-verbal sounds”.

<sup>24</sup> Essa demanda cognitiva por haver mais *inputs* de informação é comprovada. Sohl (1989 *apud* D’YDEWALLE; GIELEN, 1992) investigou, utilizando uma técnica de tarefa dupla (“*double task*”), se os participantes realmente estavam processando legenda, som e imagem ao mesmo tempo. Para isso, o autor fez uso de luzes que piscavam, às quais os participantes tinham que reagir ao longo do experimento. Essas luzes apareciam quando 1) não havia legenda; 2) a legenda e os falantes estavam na tela; e 3) nem a legenda, nem os falantes estavam em tela. O tempo de reação mais lento obtido pelo estudo foi quando legenda e atores estavam em tela, o que mostra que o processamento de som, imagem e legenda estava acontecendo e, dessa forma, era mais custoso cognitivamente.

informação relevante, mas por causa de uma mudança que a aparição da legenda causa. Isso é notável quando observamos a mudança de alocação de atenção de um filme visto sem e com legendas, como na Figura 2. No caso, na imagem da direita, percebe-se que o espectador vai menos à imagem quando comparado à movimentação do olhar da esquerda.

**Figura 2 – Mapa de calor da mesma cena vista sem e com legenda**



Fonte: Kruger, Szarkowska e Krejtz (2015, p. 4).

Essa questão foi primeiramente levantada a partir dos dados do trabalho de d'Ydewalle *et al.* (1987), os quais produziram duas hipóteses: 1) devido à experiência e familiaridade com produtos legendados, principalmente em comunidades em que há mais programas legendados, a leitura de legendas é mais ou menos automática (**hipótese da automaticidade**); 2) a leitura de legendas é mais eficiente e leva menos tempo para ser processada se comparada ao som do vídeo; além disso, as legendas podem ser lidas antes que os interlocutores falem, assim como podem ser relidas (**hipótese da eficiência**) (D'YDEWALLE; GIELEN, 1992, p. 17). Esses resultados foram qualificados em d'Ydewalle *et al.* (1991), os quais confirmaram a hipótese da eficiência e reformularam a hipótese da automaticidade. Mesmo falantes de países que não possuem tradição com o uso de legendas processam as legendas durante a exibição de filmes legendados.

Segundo Jensema *et al.* (2000, p. 284), “[...] parece haver uma tendência a se começar a olhar no meio da tela e então mover o olhar para o começo da legenda em uma fração de segundo. Espectadores leem a legenda e então olham a ação do vídeo após terminarem a leitura<sup>25</sup>”. No entanto, esse padrão suave de leitura não acontece sempre. Segundo Kruger, Szarkowska e Krejtz (2015), exemplos desses movimentos não tão suaves de leitura são as deflexões (*deflections*) (DE LINDE; KAY, 1999). As deflexões, de uma maneira geral, dizem

<sup>25</sup> “[...] there appears to be a general tendency to start by looking at the middle of the screen and then moving the gaze to the beginning of a caption within a fraction of a second. Viewers read the caption and then glance at the video action after they finish reading”.

respeito ao ato de o telespectador: 1) fazer sacadas entre a legenda e a imagem durante a leitura de uma legenda; ou 2) fazer sacadas entre imagem e legenda após a finalização da leitura dessa mesma legenda. Esse fenômeno, inclusive, recebe outros nomes na literatura, como “movimentos de para lá e para cá”, no que seria uma tradução literal de *back-and-forth shifts* (D’YDEWALLE; DE BRUYCKER, 2007), *saccadic crossovers*, no que seria uma tradução literal de “cruzamentos sacádicos” (RAJENDRAN *et al.*, 2013) e “revisitas” (*revisits*) (SZARKOWSKA; GERBER-MORÓN, 2018).

Essa pluralidade terminológica também ocasiona certa falta de clareza sobre qual dos 2 tipos de deflexão mencionados no parágrafo anterior as pesquisas utilizam em suas análises sobre o processamento de legendas. Enquanto autores como Vieira (2016) analisam os dois movimentos – durante e após a leitura de um bloco de legenda – sob o termo “deflexão”, assim como Rajendran *et al.* (2013) e Souza (2020), Szarkowska e Gerber-Morón (2018) utilizam o termo “revisita” apenas para o movimento de retorno à mesma legenda após leitura e ida subsequente à imagem. Os outros trabalhos mencionados (DE LINDE; KAY, 1999; D’YDEWALLE; DE BRUYCKER, 2007; KRUGER; SZARKOWSKA; KREJTZ, 2015) não deixam claro se fazem distinção entre os dois tipos de movimento deflexivo ou se a sua nomenclatura abarca esses dois tipos. A última opção, contudo, parece a mais provável.

Todos esses nomes para o termo “deflexão”, o qual utilizaremos aqui com mais frequência, denotam uma ruptura do processo de leitura, tornando o padrão de leitura não tão fluido e ocasionando um maior custo de processamento; ou seja, quão maior o número de deflexões, menos fluente o processamento da legenda (DE LINDE; KAY, 1999; D’YDEWALLE; DE BRUYCKER, 2007; SZARKOWSKA; GERBER-MORÓN, 2018). Contudo, alguns autores, como Vieira (2016) e Souza (2020), pensam essa questão de forma diferente, principalmente em relação ao processamento de legendas por surdos, conforme será exposto mais à frente.

Ainda sobre a questão da deflexão, Kruger, Szarkowska e Krejtz (2015) apontam que o público-alvo pode influenciar a forma como um vídeo legendado é processado, especificamente comparando o processamento de um surdo e um ensurdecido<sup>26</sup> (*hard-of-hearing*). Na Figura 3, podemos perceber que o surdo (à esquerda) lê toda a legenda e volta ao centro da imagem, para a face do personagem. O ensurdecido, por sua vez, à direita, faz movimentos de deflexão constantes entre legenda e imagem, provavelmente em uma tentativa de leitura labial ou de escutar algo a partir de algum resíduo de escuta.

---

<sup>26</sup> Para além das discussões de identidade, as quais não pretendemos abordar aqui, entendemos “ensurdecido” como uma pessoa com baixa audição e que tem uma língua oral como língua materna.

**Figura 3 – Scanpaths de um surdo (à esquerda) e um ensurdecido (à direita) assistindo a um vídeo com legendas *verbatim***



Fonte: Kruger, Szarkowska e Krejtz (2015, p. 5).

Em virtude de questões como esta é que, em nossa pesquisa, só iremos analisar o processamento de participantes surdos, além dos ouvintes.

Apesar de já termos citado algumas questões que envolvem o processamento de legendas e o rastreamento ocular, na próxima subseção citaremos alguns trabalhos dentro dessa interface e, ao reportarmos seus objetivos, metodologia e resultados, estabeleceremos um diálogo com a presente pesquisa.

### 2.2.3 Pesquisa em legendagem com rastreamento ocular

Nesta seção, apresentamos alguns trabalhos que versam sobre o processamento ocular e a legendagem e se alinham à proposta do nosso trabalho, seja por objetivo, metodologia ou estudo dos parâmetros de número de linhas e de velocidade.

A tradição de se estudar como as legendas eram recebidas por espectadores remontam aos estudos do psicólogo Gery d'Ydewalle e seus colegas no âmbito do seu laboratório na década de 1980. Apesar de a maioria dos seus trabalhos ter focado em como a atenção do espectador é alocada durante a exibição do filme legendado, o pesquisador e seu grupo forneceram dados interessantes para que se entendesse a natureza das legendas, tendo inclusive influenciado fortemente o mundo profissional que cerca essa área, anteriormente pautado apenas em tradições e práticas compartilhadas.

Em coautoria com outros pesquisadores, muito do conhecimento a respeito do processamento cognitivo de legendas que se tem até hoje veio a partir desses trabalhos, como, por exemplo, a influência do número de linhas no processo de leitura (D'YDEWALLE *et al.*, 1987; D'YDEWALLE; DE BRUYCKER, 2007) e a regra dos 6 segundos (D'YDEWALLE *et*



*al.*, 1987). Essa regra, estabelecida na prática da legendagem na Europa e testada em um de seus experimentos, reza que uma legenda completa ou cheia de duas linhas deveria permanecer na tela por no máximo 6 segundos, tempo ideal para se ler toda a legenda, não devendo essa marca ser ultrapassada por motivo de releitura da legenda. Esses e outros trabalhos, como já dito, regem muitas das pressuposições da literatura sobre legendagem até hoje, embasando o uso do rastreador ocular no estudo da TAV e dando margem à replicação dessa metodologia no estudo de outros aspectos da área ligados, muitas vezes, mais a um conhecimento oriundo da prática profissional do que a algo comprovado por pesquisas, como é o caso da influência da segmentação na recepção.

O estudo de maior influência nesse sentido foi d'Ydewalle *et al.* (1987). Nesse trabalho, os autores reportam dois experimentos. No Experimento 1, os participantes reportaram sobre a velocidade de três velocidades de legenda distintas (“regras” de 4, 6 e 8 segundos). Neste experimento, os participantes assistiram ao mesmo vídeo utilizado no estudo anterior com legendas nas mesmas velocidades de leitura (D'YDEWALLE *et al.*, 1985). Após cada legenda, eles deveriam reportar se a velocidade da legenda em questão era muito rápida, apropriada ou muito lenta, tratando-se, assim, de um método *on-line* para investigar o processamento que tinha como base a consciência dos participantes acerca da velocidade de apresentação das legendas. Dessa forma, os participantes sabiam que esse aspecto estava sendo estudado.

A partir do experimento 1 de d'Ydewalle *et al.* (1987), dois resultados foram alcançados: os participantes reclamaram mais em relação à velocidade das legendas quando elas ficavam mais rápidas. Sobre o número de linhas, os autores observaram que as legendas de 1 linha eram percebidas como mais rápidas do que as de 2 linhas. Além disso, o estudo em questão teve como objetivo investigar sobre qual combinação de linhas (1 ou 2) e velocidade (regras de 4, 6 e 8 segundos) os participantes não reclamariam. Segundo os autores, a única combinação que os participantes escolheram como apropriada foi a combinação 2 linhas/6 segundos (D'YDEWALLE *et al.*, 1987, p. 314).

Duas conclusões desse primeiro experimento puderam ser delineadas: 1) os participantes são capazes de distinguir a velocidade da legenda de acordo com a regra do tempo utilizada; 2) legendas de 1 linha, não importando a medida de tempo utilizada, são vistas como mais rápidas do que legendas de 2 linhas. Inicialmente, os autores acreditavam que o processamento de legendas de 2 linhas seria mais custoso do que de 1 linha, contudo, os resultados das observações dos respondentes da pesquisa, os quais apresentaram mais reclamação sobre o tempo com as legendas de 1 linha, refutaram essa hipótese inicial.

Inicialmente, os autores pensavam que talvez fosse uma questão relacionada ao tempo gasto para pular da imagem para a legenda quando ela é apresentada, também chamado de tempo de latência:

Nossa suspeita é direcionada ao tempo de latência para pular da imagem para a legenda quando uma nova legenda é apresentada. Supondo que essa latência é independente em relação às legendas de uma/duas linhas, obviamente sobra menos tempo proporcionalmente para o processamento da legenda com uma linha<sup>27</sup> (D'YDEWALLE *et al.*, 1987, p. 314).

Ou seja, se sempre há um tempo de intervalo entre a última fixação na área da imagem e a primeira na área da legenda, a conclusão dos autores era de que, como há mais intervalos desse tipo considerando legendas de 1 linha, menos tempo os telespectadores têm para processar essas legendas. Por essa razão, quanto mais tempo gasto na imagem, menos tempo há para o processamento das legendas, sendo o processamento, nesse caso, menos confortável.

Essa concepção, contudo, não é clara no trabalho dos autores, assim como não é clara na literatura em legendagem. Mais tempo gasto na área da legenda em relação à imagem representa um desconforto? Ou seria conforto? À primeira vista, como os próprios autores reconhecem, menos tempo gasto na legenda significa que mais tempo está sendo alocado para a imagem e, dessa forma, havendo uma média menor de duração das fixações, o processamento é menos custoso, caso sejam levadas em consideração as fixações em que o telespectador realmente está lendo (sobre o acompanhamento do olhar falaremos mais adiante). Contudo, a maneira como os autores expressam esse tempo maior entre a última fixação na imagem e a primeira na legenda desconsidera o fato de que legendas de 1 linha poderiam proporcionar mais tempo de visita à imagem (principalmente no intervalo entre blocos de legendas de 1 linha).

Em relação ao Experimento 2, os autores tiveram como objetivo investigar as “regras” de velocidade (vide Regra dos 6 segundos) em amostras independentes de participantes, examinando o comportamento de leitura destes em relação a velocidades normais de apresentação de legenda. Esse estudo foi motivado por um estudo anterior de d'Ydewalle *et al.* (1985), em que os autores haviam misturado as velocidades de legenda em um mesmo vídeo e cujos participantes tinham que reportar, logo após a exibição do trecho legendado, se a legenda estava muito rápida, apropriada ou lenta. Neste estudo, os autores não conseguiram

---

<sup>27</sup> “Our suspicion is directed to the time latency to jump from the visual image to the subtitle when a new subtitle is presented. Assuming that this latency is independent from a one-/two-lines presentation, obviously there remains proportionally less time available for processing the subtitle with one line”.

definir com clareza o padrão de leitura dos participantes possivelmente pelo fato de terem misturado essas velocidades de legenda em um só filme.

Em d'Ydewalle *et al.* (1987), os autores levaram em consideração o fato de os participantes serem usuários constantes de vídeos legendados e serem conscientes sobre variações na velocidade de apresentação de uma legenda (como apontou o Experimento 1),

A partir disso, para o Experimento 2, três versões (2 segundos, 4 segundos, 6 segundos) do mesmo filme foram feitas. A regra de 8 segundos foi retirada por apresentar problemas e dificuldades para sincronização. Como um dos requisitos para a manipulação dos vídeos, o começo das legendas encaixava ao começo da fala dos atores em tela. Como padrão, as legendas utilizadas faziam uso de até 64 caracteres em legendas de duas linhas (32 caracteres em legendas de 1 linha). Para o experimento, foi definido um *design* de fator 3 (com som, sem som e com participantes com conhecimento da língua do filme), x 3 (velocidade de legenda em 2, 4 e 6 segundos) e x 2 (sexo masculino e feminino) (D'YDEWALLE *et al.*, 1987).

O equipamento utilizado no trabalho foi o DEBIC 80, baseado no método centro-da-pupila/reflexão-na-córnea. A capacidade desse rastreador ocular era de 50 hertz (hz), sendo capaz de gerar amostras a cada 20 milissegundos.

Para a análise, os autores examinaram quanto tempo foi gasto na legenda pelos participantes. Além disso, foram analisados três preditores de variância no estudo: 1) características da palavra (*word characteristics*), que incluía número de sílabas, número de palavras, frequência das palavras em holandês; 2) estrutura de superfície da legenda (*surface structure of the subtitle*), relacionada ao número de linhas e velocidade da legenda; e 3) características episódicas da legenda (*episodic characteristics of the subtitle*), o qual está relacionado a fatores como: (primeiro grupo) se o interlocutor muda de uma legenda para outra, se a legenda começa um novo episódio no filme, se (segundo grupo) se trata de uma pergunta, afirmativa ou os dois em uma mesma legenda, (terceiro grupo) quais interlocutores estão falando. Em suma, as características de superfície da legenda foram os principais preditores de variância, principalmente na condição em que a legenda era acompanhada do som original do filme (D'YDEWALLE *et al.*, 1987).

Sobre o tempo gasto na área da legenda em função da velocidade e do número de linhas, os resultados mostraram que os participantes passaram mais tempo proporcionalmente nas legendas de 2 linhas, quando considerados os três grupos expostos às legendas (com som, sem som e falantes da língua). Contudo, na condição de seis segundos, não houve diferença, sendo bem menor também na condição "língua" (D'YDEWALLE *et al.*, 1987).

Em velocidades de leitura mais rápidas, como nas regras de 2 e 4 segundos, houve um maior tempo gasto nas legendas de duas linhas, especialmente nas condições “com som” e “sem som” (ou seja, quanto maior a velocidade, mais tempo é gasto nas legendas de 2 linhas proporcionalmente). A pequena variação dos participantes que conheciam a língua indica que a velocidade não parece ser um fator tão determinante para o tempo gasto na área da legenda tanto em legendas de 1 linha quanto de 2 linhas (D’YDEWALLE *et al.*, 1987).

Sobre o tempo de latência, os autores ponderam que os telespectadores precisam de um determinado tempo (que é constante) para ir da imagem para a legenda. Sendo assim, mais tempo estaria disponível para o processamento de legendas de duas linhas (D’YDEWALLE *et al.*, 1987).

Em relação às condições “com som” e “sem som” quando relacionadas ao número de linhas na medida de tempo de latência, não houve diferenças significativas em legendas de 1 linha e 2 linhas, tendo sido constatada uma diferença mais acentuada apenas no tempo maior de latência da legenda de 2 linhas na condição “língua” quando comparada com as latências de 2 linhas nas outras condições e com as latências de 1 linha em todas as condições. Proporcionalmente, entretanto, o tempo de latência nas condições de 2 linhas foi menor do que nas condições de 1 linha. Em relação a isso, visto que o tempo de latência para as condições “com som” e “sem som” foi parecido em legendas de 2 linhas e 1 linha, havia mais tempo de leitura proporcionalmente para as legendas de 2 linhas, o que entrou em consonância com as respostas dos participantes que apresentaram mais reclamações para legendas de 1 linha (D’YDEWALLE *et al.*, 1987).

A diferença de tempo gasto na área da legenda em legendas de 1 linha e de 2 linhas foi bem menor na condição “língua”, o que é resultado do fato de eles terem dedicado um tempo maior para iniciar a leitura de legendas de 2 linhas quando comparados com as outras condições (D’YDEWALLE *et al.*, 1987).

Outra consideração do estudo foi que a diferença entre legendas de 1 linha e 2 linhas foi bem menor quando a regra de 6 segundos foi utilizada, o que não pode ser associado à questão do tempo de latência, visto que a velocidade de leitura da legenda não apresentou significância e não esteve envolvida de forma significativa em uma interação.

A regra dos seis segundos foi a que mostrou um padrão de leitura mais homogêneo, considerados os grupos que assistiram ao filme com som, sem som e com conhecimento da língua veiculada em cena, como no tempo gasto na área da legenda, por exemplo, tanto em porcentagem quanto em tempo absoluto. Nas outras velocidades de legenda, 2 e 4 segundos, foi possível observar um tempo maior gasto nas legendas de 2 linhas (comparados com as de 1

linha); ou seja, quanto maior a velocidade da legenda, mais custoso foi o processamento de legendas de 2 linhas (D'YDEWALLE *et al.*, 1987).

O estudo de d'Ydewalle *et al.* (1987) é um dos mais relevantes na literatura em legendagem por ter sido o primeiro a testar experimentalmente parâmetros utilizados na área baseados apenas na tradição da tradução audiovisual, principalmente na Europa. O estudo em questão se faz relevante para esta tese por, além do uso de uma metodologia baseada em rastreamento ocular, fazer uso da análise de parâmetros de velocidade e de número de linhas. Contudo, algumas das suas escolhas em termos de *design* experimental falseiam até demais uma validade ecológica do processamento de um produto audiovisual (como a questão de perguntar após cada inserção de legenda se achavam que estavam mais rápidas). Ademais, um controle maior acerca do número de legendas com 1 linha e de 2 linhas no experimento não é explicitado, assim como o número de caracteres (o artigo menciona apenas que o máximo de 32 caracteres em legendas de 1 linha e de 64 em legendas de 2 linhas foi utilizado). Além da medida de tempo de latência utilizada pelos autores para explicar como uma legenda de 1 linha seria menos confortável do que de 2 linhas, os autores também usaram o número maior de reclamações dos respondentes em relação às legendas de 1 linha para fortalecer a sua argumentação. Entretanto, a questão do número baixo de caracteres em legendas de 1 linha também pode ter influenciado essa percepção dos participantes, como no caso de legendas com exclamações ou interjeições, dando a impressão de que são mais rápidas por ficarem menos tempo em tela. Esse aspecto não é discutido no trabalho e nem qualificado durante a discussão dos resultados.

Outra questão que não fica clara é como essas áreas de interesse (AOIs – *areas of interest*) foram delimitadas, assim como características técnicas e linguísticas das legendas. Hoje em dia sabemos que essas características também influenciam o processamento, como a coesão, a segmentação baseada em corte de cenas, segmentação linguística etc.

A medida de tempo de latência foi utilizada nesta pesquisa, como será explicitado mais adiante, mas o tempo gasto na área da legenda, não. Isso porque nem todo tempo gasto na área da legenda significa um processo de leitura. E apesar de a noção de latência ser mais uma medida de alocação de atenção do que de processamento em si, é uma medida que historicamente tem sido analisada nos estudos de processamento de legendas de forma mais consistente.

Outro estudo interessante para esta tese foi Verfaillie e d'Ydewalle (1987), reportado por d'Ydewalle e Gielen (1992). O estudo em questão teve como objetivo investigar dois aspectos do processamento de legendas por participantes surdos: 1) a preferência de

modalidade (leitura labial, tradução em língua de sinais e legenda); e 2) a compreensão da mensagem a partir dessa preferência. Os participantes foram expostos a vídeos retirados de uma ficção de TV em 4 modalidades distintas (1 – falante, intérprete e legendas ao mesmo tempo; 2 – legendas e intérprete; 3 – legendas e falante; e 4 – falante e intérprete). Após responderem a perguntas de verdadeiro/falso sobre detalhes e perguntas sobre a compreensão geral da história nessas modalidades em questão, os resultados evidenciaram uma preferência pelas legendas, apresentando melhores resultados nas perguntas quando expostos às legendas. Esses resultados mostraram que, aparentemente, os participantes não conseguiram acompanhar a história com presença apenas da leitura labial e da língua de sinais. A língua de sinais, inclusive, aparentemente influenciou negativamente a recepção das legendas enquanto assistiam ao vídeo. Segundo os autores, a preferência pelas legendas parece se dar pelo fato de que, quando acompanham a interpretação em língua de sinais, os participantes surdos acompanham-na tão extensivamente que acabam por ter receio de olhar para a imagem do vídeo e perder alguma informação importante na tradução. O melhor resultado das legendas também pode ser causado pelo seu tipo de processamento ser quase obrigatório quando está em tela, processamento automático já discutido anteriormente.

Segundo d'Ydewalle e Gielen (1992), em quase todos os experimentos diferenças de processamento entre legendas de 1 linha e 2 linhas foram obtidas. Ainda segundo os autores, legendas de duas linhas são olhadas por mais tempo e, em alguns casos, apresentaram também um tempo maior de latência (da imagem para a legenda). Mas por que isso ocorre? De acordo com d'Ydewalle *et al.* (1991), devido à preferência por fixar o olhar na imagem, e, tendo em mente que legendas de 1 linha normalmente são mais rápidas, os participantes tendem a desviar o olhar mais rapidamente para elas. Isso está em acordo com o que os autores chamam de “*length-expectation hypothesis*”, ou seja, os espectadores decidem a rapidez com que irão desviar o seu olhar para a área da legenda a partir do seu tamanho e da sua velocidade. Isso foi sugerido nos resultados do segundo experimento de d'Ydewalle *et al.* (1989), que observou que, a partir dos resultados de tempo de latência dos participantes jovens da pesquisa, estes eram capazes de observar o tamanho e o formato da legenda enquanto olhavam a imagem.

Outra explicação seria a de que as legendas de 2 linhas são fixadas por mais tempo por haver algum tipo de interferência lateral (diferentemente das legendas de 1 linha), conforme já colocado anteriormente, o que faria com que as pessoas demorassem mais a começar o processamento da primeira palavra da primeira linha. Da mesma forma, esses participantes demorariam a realocar o olhar entre o final da primeira linha e o começo da segunda. Esses

tipos de interferência aumentariam a probabilidade de que movimentações de correção acontecessem.

Praet *et al.* (1990), por sua vez, investigaram essa interferência da seguinte forma: expuseram os participantes a legendas sem imagem ou som, chegando a manipular o espaçamento entre as linhas da legenda. Como resultado, observaram que não há interferência lateral, assim como as legendas de 2 linhas foram lidas mais rapidamente devido à sua complexidade semântica, ao passo que legendas de 1 linha normalmente apresentam sentenças curtas. Dessa forma, essa complexidade semântica e sintática proporcionada pelas legendas de 2 linhas ocasionariam um processamento mais fluido por causar mais redundância na estrutura interna da legenda. Contudo, um achado do mesmo estudo, quando as legendas foram veiculadas junto às imagens, foi um pouco surpreendente. Quando havia imagens, os resultados foram opostos (as legendas de 2 linhas foram lidas mais vagarosamente nesse contexto). Segundo os autores, uma possível explicação para isso é o fato de que a falta da imagem coloca as legendas de 2 linhas (e sua característica de sintaxe e semântica mais completas) em primeiro plano, aumentando assim a fluidez do processamento a partir de uma redundância textual. Já quando as imagens estão em cena (que é o que acontece normalmente), estas fornecem uma maior redundância e encaixe para legendas de 1 linha. Dessa forma, as legendas de 2 linhas já conteriam grande parte da informação veiculada na imagem. Apesar de não serem tão assertivos quanto ao melhor processamento (se de legendas de 1 linha ou de 2 linhas), os autores concluem que a redundância facilitaria o processamento das legendas, apontando, assim, para as legendas de 2 linhas como mais fluidas.

Outro estudo que se relaciona ao presente trabalho em termos de metodologia e parâmetro analisado (velocidade) com auxílio de rastreamento ocular foi De Linde e Kay (1999). Nesta pesquisa, os autores reportam cinco experimentos que tinham como objetivo analisar os efeitos de determinados parâmetros no comportamento leitor de 20 participantes (metade surda e metade ouvinte) de média de idade de 29,7 anos. Esses parâmetros foram: i) velocidade da legenda; ii) tempo de entrada da legenda após o começo da fala; iii) cortes de cena; iv) tipo de edição realizada (quantidade e tipo de omissões realizadas); e v) visibilidade do falante (se está ou não em cena). Além dessas variáveis independentes relacionadas às legendas, o grupo e o tipo de programa legendado também se configuravam como variáveis manipuladas.

Para cada um dos parâmetros, dois vídeos, contendo 2 minutos cada (os 30 segundos iniciais e finais de cada vídeo foram descartados da análise, totalizando um tempo de 1 minuto por vídeo), foram utilizados, tendo sido extraídos de canais de TV britânicos. Com a

intenção de verificar se os vídeos que testavam cada parâmetro só se diferenciavam no aspecto que estava sendo estudado, testes T foram realizados, os quais confirmavam que as outras características presentes estavam distribuídas de forma aleatória. Esses mesmos vídeos tiveram seu som desligado tanto para os ouvintes quanto para os surdos. Após cada vídeo, os participantes respondiam também a uma série de perguntas sobre a compreensão do vídeo e o experimento *per se* (DE LINDE; KAY, 1999).

Com a sua movimentação ocular registrada por um monitor EMR-V, os participantes foram observados a partir das seguintes variáveis dependentes<sup>28</sup>: i) tempo de leitura (tempo entre a primeira fixação e a última fixação na legenda divididos pelo número de palavras); ii) deflexões (número de vezes que os participantes desviaram da legenda para a imagem); iii) duração das deflexões; iv) fixações (número de palavras por fixação); v) regressões (número de participantes que releram palavras ou caracteres); vi) releituras (número de vezes que os participantes releram uma legenda inteira). Alguns movimentos oculares não foram incluídos na análise dos experimentos, como **fixações** e **tamanho das sacadas** (DE LINDE; KAY, 1999).

O **estudo 1**, que teve como objetivo examinar o efeito de legendas rápidas e lentas no comportamento leitor, presumia que mais informação seria retida em legendas mais lentas. Além disso, era esperado que mais releituras iriam acontecer com legendas lentas devido ao maior tempo de exibição. Dois vídeos foram utilizados com as seguintes velocidades de legenda: 1) programa 1a – velocidade de legenda alta (média de 139 palavras por minuto); 2) programa 1b – velocidade de legenda baixa (média de 74 palavras por minuto). Os resultados mostraram efeito de interação nas variáveis programa e velocidade nas medidas de **fixação**, **releitura** e **tempo de leitura**. Em relação ao **tempo de leitura**, um menor tempo de fixação por palavra foi observado nas palavras do programa 1a (velocidade rápida), cerca de 250 milissegundos por palavra, do que no 1b, o qual teve cerca de 300 milissegundos por palavra. Segundo os autores, isso pode sugerir que o tempo de leitura é um reflexo do ritmo das legendas (DE LINDE; KAY, 1999).

---

<sup>28</sup> Segundo os autores, alguns parâmetros de investigação do comportamento ocular durante a leitura não podem ser simplesmente categorizados como ruins ou bons. Um exemplo seria o de releitura. Inicialmente, é presumível que o fato de a legenda necessitar ser relida seja um fator de perturbação e, dessa forma, seja negativo no processamento da legenda; contudo, pode haver um caso em que, após visualizar a imagem, um determinado significado emerge com a releitura de uma legenda. Dessa forma, é necessário que essa análise do comportamento ocular se dê em contexto e não de forma automática, de modo que o julgamento a respeito do efeito de determinadas características da legenda nesse processamento seja mais preciso (DE LINDE; KAY, 1999, p. 62).



Na variável dependente de **fixação**, os autores registraram que os participantes processaram mais palavras por fixação no programa 1a (1,36 palavras por fixação) do que no programa 1b (1,33 palavras por fixação), alinhando os resultados dessa medida com o tempo de leitura. No que diz respeito à **releitura**, mais releituras ocorreram no programa 1b (0,615) do que no programa 1a (0,315), também estando de acordo com as expectativas relacionadas ao programa com velocidade mais lenta possibilitar mais releituras. Contudo, uma crítica a esse respeito é uma legenda de cerca de 8,56 segundos exemplificada no estudo. Uma legenda exibida por tanto tempo na tela, desrespeitando até mesmo a regra dos 6 segundos (D'YDEWALLE *et al.*, 1987) aplicada na Europa pode ter causado tantas releituras no programa com legendas mais lentas (DE LINDE; KAY, 1999).

As notas obtidas nos testes de compreensão foram mais altas no programa 1a (71%), com legendas rápidas, do que nos programas 1b (46%), com legendas lentas, contrariando a hipótese de que mais informação seria resultante do programa com velocidade mais baixa. Contudo, como os autores escolheram programas semelhantes (documentário), mas com características textuais e imagéticas distintas (no programa 1a, o aspecto textual era mais relevante do que no 1b), eles próprios consideram que essa composição distinta dos estímulos pode ter influenciado o resultado (DE LINDE; KAY, 1999).

Esse primeiro estudo relatado em De Linde e Kay (1999) é relevante para nossa pesquisa por tratar do estudo de velocidade da legenda alinhado ao uso de medidas provenientes de rastreamento ocular, semelhante ao que nos propomos fazer aqui. Contudo, algumas escolhas metodológicas do estudo colocam em suspeição alguns dos dados obtidos: 1) o fato de não terem feito uso de um quadrado latino (MARQUER, 2003) de modo que os vídeos fossem vistos em condições diferentes e, dessa forma, a relação entre imagem, conteúdo e condição experimental pudessem ser minimizados nos resultados; 2) o uso de uma velocidade bem baixa, como no caso da condição com 74 ppm, o que acabou gerando legendas bem longas, como no caso da legenda de 8 segundos, relacionado ao 3) não uso da regra de 6 segundos, prática recomendada por pesquisas, pela comunidade europeia de profissionais legendistas e pelos próprios telespectadores (D'YDEWALLE *et al.*, 1998).

Apesar disso, a questão apontada em relação ao número maior de releituras nas legendas lentas (no caso do estudo, muito mais lentas do que o padrão de legendas consideradas lentas na literatura), assim como um tempo menor de leitura nas legendas apontadas como rápidas pelos autores, sugeriu caminhos para o estudo da velocidade da legenda no processamento. Legendas muito lentas atrapalham o processamento, ocasionando mais releituras e um tempo maior de duração média das fixações por palavra.

O **estudo 2** investigou como o atraso (0,02 segundos, praticamente nulo, no programa 2b) ou antecipação (0,8 segundos no programa 2a) de uma legenda afetaria o processamento desta a partir da hipótese de que mais deflexões ocorreriam se a legenda antecipasse o falante em tela (o espectador sairia da legenda para procurar um falante), assim como era esperado que releituras (mais releituras) ocorressem nesse contexto. Os resultados apresentaram efeito significativo de releitura por grupo, programa e legenda, além de uma interação entre programa e legenda na medida **número de deflexões**. Em termos de **releitura**, os ouvintes leram duas vezes mais do que os participantes surdos (como leram mais rápido, tinham mais tempo para reler a legenda). Além disso, mais releituras ocorreram no programa 2b, contudo, aparentemente mais pela natureza dos vídeos do que pela sincronização das legendas (devido a características de algumas legendas no final do vídeo 2a, um diálogo com falas próximas e com dois falantes em cena, menores releituras ocorrem e mais deflexões, estas motivadas pela presença dos falantes em tela) (DE LINDE; KAY, 1999).

O **estudo 3**, por sua vez, investigou os efeitos de cortes de cena no processamento de legendas utilizando dois vídeos (programa 3a, com menos cortes de cenas, média de 1,3; e programa 3b, com mais cortes, média de 3,5). A hipótese era de que esses cortes causariam perturbação na leitura, assim ocasionando mais deflexões para a imagem. Um efeito relacionado ao número de palavras por fixação, regressões e releituras também era esperado. Os resultados mostraram efeito na variável “programa” em relação à duração das deflexões, assim como uma interação de efeitos de programa e legenda com as variáveis de tempo de leitura, fixações, regressões e releitura. A duração das deflexões, no caso, confirmou, aparentemente, a hipótese no programa 3b, que tinha um maior número de cortes e teve um tempo médio maior de deflexão. O número de palavras por fixação também foi maior no programa 3b (quanto maior o número de cortes, mais o olhar desviava para a imagem e menos tempo passava na área da legenda) para ambos os grupos. A releitura também mostrou ter sido influenciada pelos cortes de cena. O tempo de leitura, contudo, não pareceu ser afetado pelos cortes de cena, enquanto as regressões também foram similares em ambos os programas (DE LINDE; KAY, 1999).

O **estudo 4**, por sua vez, teve como objetivo investigar os efeitos das omissões de palavras no processamento das legendas a partir da hipótese de que um maior número de omissões seria prejudicial, dada a possível falta de sincronia com o que é dito pelos atores em cena, causando menos palavras por fixação. Nos dois vídeos utilizados, um continha um alto número de omissões (programa 4a, com média de 9,6) e o outro, um número baixo de omissões (programa 4b com uma média de 2,4). Os resultados mostraram interação de efeitos de grupo

e programa, grupo e legenda e programa e legenda. O efeito mais significativo nesse estudo foi o número menor de palavras fixadas nas legendas mais editadas em ambos os grupos de participantes, o que acabou por causar, também, um menor número de releituras nas legendas do programa 4b. Contudo, a duração menor das fixações pode também ter sido causada pela estranheza de algumas bocas estarem se mexendo e não haver legenda na tela, além de um pouco mais de ação no vídeo em questão. A diferença no tempo de leitura, contudo, parece ter sido causada mais pelo tamanho da legenda (e por sua velocidade) do que pela questão da omissão em si.

O **estudo 5**, por sua vez, testou os efeitos que o falante presente em tela (*on-screen*) ou fora dela (*off-screen*) têm no processamento de legendas. Segundo a hipótese do estudo, os participantes teriam um comportamento leitor menos desordenado caso o falante estivesse *on-screen* (programa 5b), pois poderiam focar apenas nas legendas e no movimento dos lábios. Dessa forma, o vídeo com falante *off-screen* (programa 5a) causaria mais deflexões e, dessa forma, menos tempo de leitura. Os resultados mostraram efeito significativo em grupo e duas interações significativas entre grupo e programa e programa e legenda. De uma forma geral, a hipótese do estudo foi confirmada; o vídeo com o falante *on-screen* demandou menos do processamento das legendas em ambos os grupos. Os surdos, especialmente, tiveram mais dificuldade com os vídeos em que o falante estava fora da tela, o que resultou em um efeito *garden-path*, ou seja, os surdos procuravam quem estava com o turno de fala na imagem. Além disso, como houve mais deflexões no programa 5a (*off-screen*), os participantes também apresentaram um tempo de leitura maior, assim como um número maior de palavras por fixação. As regressões também foram menores, visto que havia menos tempo de leitura.

Todos os 5 experimentos realizados por De Linde e Kay (1999) são relevantes para a presente pesquisa por, além de trabalharem com rastreamento ocular e o processamento de legendas, terem utilizado como público-alvo surdos e ouvintes e como uma de suas variáveis a questão da velocidade (**estudo 1**). Outras variáveis independentes testadas, como o corte de cena, a sincronização, o grau de omissão e a visibilidade do falante em cena, as quais mostraram em menor ou maior extensão influência no processamento de legendas, apesar de não serem testadas no presente estudo, também foram consideradas durante a montagem e manipulação dos vídeos presentes no experimento. Além disso, algumas das variáveis dependentes utilizadas, como o uso de algumas medidas adotadas por De Linde e Kay (1999), como o número de deflexões, fixações, regressões e releituras, também são importantes e foram considerados em nossa análise.

Algumas das decisões metodológicas na manipulação e distribuição dos estímulos que aparentemente demonstram influência no processamento (para além das variáveis que estavam sendo testadas) foram evitadas até certo ponto na presente pesquisa. Apesar de também utilizarmos vídeos distintos para cada condição, todos os vídeos foram legendados nas condições experimentais testadas (velocidade e número de linhas); todos os vídeos continham o mesmo número de legendas e tempo aproximado. Contudo, mesmo assim, a influência do aspecto imagético no processamento das legendas testadas ainda é uma incógnita; mesmo utilizando um filme baseado em diálogos mais estáticos (apesar de, determinadas vezes, com mais de um falante em tela) e com trechos selecionados a partir de pouco movimento de câmera e determinada clareza sobre quem possuía o turno de fala, era possível que, mesmo não desejado, aparecesse um efeito de vídeo nos nossos resultados, como de fato apareceu. Esse talvez seja um dos próximos passos na pesquisa de processamento de legendas: como neutralizar o caráter imagético do produto audiovisual legendado sem comprometer a validade ecológica do experimento? Sobre isso, De Linde e Kay (1999, p. 73) reafirmam que: “A pesquisa em movimentação ocular até agora tem se concentrado nos efeitos de características da legenda, mas é provável que o conteúdo visual de um filme seja igualmente importante<sup>29</sup>”.

Um dos estudos que abordam a interface legendagem/rastreador ocular, o que o torna interessante para a presente pesquisa, seja por objetivos, metodologia ou objeto de estudo, é Perego *et al.* (2010). A pesquisadora testou a hipótese de que legendas de duas linhas que tenham segmentação sintaticamente “incoerente”, ou seja, mal segmentadas, poderiam atrapalhar o processo cognitivo de recepção daquelas legendas por parte dos espectadores. Além desta, outra hipótese testada foi sobre o processamento de um produto legendado. A questão de pesquisa residia no fato de este ser cognitivamente efetivo, a partir de um bom entendimento do filme sem comprometimento dos elementos visuais e verbais. Para tal, levaram em consideração a movimentação do olhar (alocação de atenção), a compreensão do filme como um todo e o reconhecimento de palavras específicas das legendas e de cenas do filme.

Analisando a relação entre alocação da atenção, reconhecimento da palavra e reconhecimento da cena, o *corpus* do experimento consistiu em um trecho de um filme húngaro de 15 minutos com legendas em italiano. O número total de legendas criadas foi de 171, todas feitas conforme os preceitos técnicos da área, a não ser a variável independente do experimento: a segmentação. Dessas 171 legendas, 28 foram manipuladas – sem e com problemas de

---

<sup>29</sup> “Eye-movement research has up till now concentrated on the effects of subtitle features, but it is likely that the visual content of a film is equally important”.

segmentação, estes somente no nível do sintagma nominal – e apenas os dados do rastreador ocular provenientes da recepção a essas legendas manipuladas foram considerados. Participaram do experimento 41 estudantes – 23 mulheres e 18 homens – de graduação e pós-graduação das Universidades de Trieste e Pavia, na Itália. Todos falantes nativos de italiano e com visão normal ou corrigida que não tinham costume de assistir a filmes legendados e não possuíam conhecimento da língua do áudio original – o húngaro. Eles foram divididos em dois grupos de modo a contrabalancear os dados do grupo exposto a legendas mal segmentadas e do grupo exposto a legendas bem segmentadas. Apenas 16 dos participantes foram sujeitos à aplicação do rastreador ocular, devido às limitações de tempo (PEREGO *et al.*, 2010).

Segundo seus resultados, a má segmentação nas legendas não impossibilitou ou dificultou a compreensão do filme pelos participantes do experimento, tendo eles atingido bons índices de reconhecimento de cena e de palavras. Em relação ao rastreador ocular, também não houve diferenças significativas em relação ao número de fixações nas legendas bem e mal segmentadas, ao número de fixações na parte legendada e na imagem, à proporção de tempo gasto na área legendada e ao tempo médio de cada fixação nos trechos com condições experimentais distintas – bem e mal segmentada. Também não houve diferenças no *path length* – soma de toda a distância percorrida, em pixels, dos segmentos que ligam cada fixação durante o intervalo em que cada legenda é apresentada, ou no número de deflexões, este o número de transições entre a área legendada e a imagem. A única discrepância, apesar de pequena, foi no tempo médio gasto na fixação dentro da área legendada nas legendas mal segmentadas, que, no caso, foi um pouco maior do que nas legendas bem segmentadas (PEREGO *et al.*, 2010).

O trabalho de Perego *et al.* (2010) é importante para a presente pesquisa por ser um estudo experimental realizado sobre legendagem com o auxílio do rastreador ocular. Ele também se faz relevante a partir das medidas do rastreador ocular utilizadas para a análise quantitativa do percurso do olhar do participante, sendo que algumas dessas medidas foram replicadas neste estudo. Contudo, reitero, assim como o fez Chaves (2012), a inadequação em um ponto crucial do experimento em questão: a quantidade e diversidade dos problemas de segmentação linguística, já que apenas problemas dentro do sintagma nominal foram manipulados e criados. Isso é incoerente com o que tem sido observado nas pesquisas descritivas sobre o parâmetro da segmentação linguística, que, por sua vez, identificaram, a partir dos pressupostos teóricos da legendagem, possíveis problemas de segmentação em todos os níveis sintagmáticos – sintagma verbal, nominal, adverbial, adjetival e preposicionado, além de quebras entre orações coordenadas e subordinadas (ARAÚJO; ASSIS, 2014; ASSIS, 2013; ASSIS, 2016; CHAVES, 2012; DINIZ, 2012; GABRIEL, 2015). Para a presente pesquisa,

todas as legendas contidas no experimento foram condições relacionadas às duas variáveis independentes da pesquisa: o número de linhas e a velocidade da legenda; dessa forma, todas as legendas apresentaram uma boa segmentação.

D'Ydewalle e De Bruycker (2007), por sua vez, é outro estudo que se alinha à metodologia e ao parâmetro investigado em nosso trabalho (número de linhas). Em seu trabalho, os autores analisaram a movimentação ocular de crianças (5<sup>o</sup> e 6<sup>o</sup> anos) e adultos quando expostos às legendas em duas condições: i) filme com áudio em língua estrangeira e legendagem na língua materna dos participantes (ao que eles denominaram “legendagem padrão”); e ii) filme com áudio em língua materna e legendagem na língua estrangeira (ao que eles denominaram como “legendagem reversa”). O estudo foi motivado por ter sido pouco estudada a forma como os espectadores realizavam uma mudança de atenção entre área destinada à imagem e área destinada à legenda em uma tarefa tão complexa.

Segundo os autores, citando d'Ydewalle e Gielen (1992), a essência do processamento de um vídeo legendado consiste na alocação de atenção para o começo (*onset*) da legenda assim que esta aparece em tela, algo quase obrigatório, independentemente de a disponibilidade da trilha sonora ser em língua estrangeira ou até mesmo das ações que ocorrem em tela. Segundo os autores, esse processamento de gêneros que contêm texto verbal e escrito eram condizentes com outros estudos dentro da psicologia experimental.

Diferentemente de estudos prévios relacionados à legendagem, que se ocupavam apenas da atenção e do processamento da mudança de imagem para texto escrito (deflexão), a pesquisa em questão teve como objetivo o estudo do processamento palavra por palavra da legenda. Para tal, utilizou como medidas: 1) número de fixações na legenda; 2) duração das fixações; 3) a amplitude da sacada; 4) o número de regressões à legenda; 5) o número de vezes em que a legenda foi completamente “pulada”; assim como 6) o número de mudanças de atenção entre a área da imagem e a área da legenda (deflexão); e 7) a probabilidade de fixação da palavra. Além de estudarem o processamento palavra por palavra de legendas por crianças e adultos nas condições supracitadas, os autores também tiveram como objetivo observar o processamento de legendas de uma linha contra legendas de duas linhas. O estudo tinha como hipótese que a condição que apresentasse legendas de 2 linhas teria menos legendas puladas, mais tempo gasto na legenda, maior fixação de palavras etc. (D'YDEWALLE; DE BRUYCKER, 2007).

Os participantes (crianças e adultos falantes nativos de holandês) foram expostos a excertos de um desenho animado sueco, língua da qual não tinham conhecimento. O *design* do experimento foi 2x2x2, com três variáveis independentes: idade, o tipo de legendagem (padrão

x reversa) e o número de linhas. Quanto aos resultados relativos ao número de linhas, os quais têm mais relevância para a presente pesquisa, o **número de legendas “puladas”** foi significativo, ocorrendo em 15% das legendas de 1 linha contra 10% das legendas de 2 linhas.

Sobre o **tempo de latência** (tempo entre a aparição da legenda e a primeira fixação), os resultados foram significativos em relação ao número de linhas (375 ms para legendas de 1 linha e 501 ms para legendas de 2 linhas). Houve também uma interação significativa entre as variáveis independentes de número de linhas e tipo de legendagem. Nas condições com legendagem reversa, os participantes demoraram 416 ms para realizar a fixação em legendas de uma linha, contra 601 ms em legendas com 2 linhas, tempo considerável.

No que diz respeito à **porcentagem de tempo gasto na área da legenda**, os resultados mostraram novamente uma interação significativa entre o número de linhas e o tipo de legendagem. Na condição que apresentava estímulo com legendagem padrão, houve maior tempo gasto na área da legenda em legendas de 1 linha (45%) do que em legendas de 2 linhas (37%). Na condição com legendagem reversa, os resultados foram opostos, com as legendas de 2 linhas apresentando um maior tempo (28%) gasto na área da legenda do que as legendas de 1 linha (24%).

Sobre a **duração das fixações**, houve significância no número de linhas novamente, sendo as fixações em legendas de 1 **linha mais curtas** (212 ms) do que em legendas de 2 linhas (228 ms). Também houve interação significativa entre o número de linhas e o tipo de legendagem. A duração de fixações foi maior nas legendas de 2 linhas na condição com legendagem reversa (242 ms) do que na legendagem padrão (215 ms), sendo que não houve significância nesse sentido em relação às legendas de 1 linha. Contudo, houve uma pequena significância em relação às legendas de 1 e 2 linhas na legendagem reversa, com menor tempo gasto nas de 2 linhas (212 ms) do que nas de 1 linha (242 ms).

Além disso, a interação entre o número de linhas e a idade dos participantes mostrou que houve diferença significativa na duração das fixações pelas crianças (241 ms em legendas de 1 linha e 268 ms em legendas de 2 linhas), o que não ocorreu com os adultos.

Em relação à **probabilidade de fixação na palavra** (número de fixações dividido pelo número de palavras na legendagem), houve significância na interação entre número de linhas e tipo de legendagem, sendo que os participantes fizeram menos fixações em legendas de 2 linhas do tipo reversa do que nas de 1 linha reversa e padrão, assim como nas de 2 linhas padrão.

No que tange à **porcentagem de movimentos regressivos**, o número de linhas também foi significativo, com as legendas de uma linha apresentando um número maior de

regressões (46%) do que as legendas de duas linhas (34%). Da mesma forma, o número de linhas também foi significativo no número de deflexões, com os participantes fazendo um número maior de deflexões nas legendas de 2 linhas do que nas legendas de 1 linha.

Em suma, a interação entre tipo de legendagem e número de linhas foi significativa, com as legendas de 2 linhas na condição padrão mostrando menos linhas puladas, maior tempo gasto na área da legendagem e menor porcentagem de regressões. Na condição reversa, as legendas de 1 linha causaram maior conforto, obtendo um menor tempo de latência, maior porcentagem gasta na área da legenda, menor duração das fixações e maior probabilidade de fixação de palavra.

Os resultados sobre a influência das legendas no processamento da legendagem padrão são consistentes com os resultados de d'Ydewalle e Gielen (1992) no que diz respeito ao maior conforto apresentado pelas legendas de 2 linhas do que de uma, apesar de os resultados não terem sido sempre significativos. D'Ydewalle e De Bruycker (2007) enfatizam que são resultados que precisam ser qualificados, já que, por exemplo, há mais deflexões nas legendas com duas linhas do que nas legendas com 1 linha, o que, para os autores, seria um sinal de desconforto. Isso seria explicado devido à regra dos 6 segundos, a partir da qual uma legenda de 2 linhas dá tempo o suficiente para que, após a leitura da legenda, o espectador volte para a imagem.

Essa questão, no entanto, analisada pelos autores como um sinal de desconforto, vai de encontro ao que é defendido por Monteiro (2016) e Vieira (2016), especialmente para o espectador surdo. Para as autoras, um maior número de deflexões significa um maior conforto do espectador do produto legendado. Isso ocorre devido ao fato de que, quanto mais o espectador, surdo, principalmente, por ter uma língua visual como L1, puder fazer o cotejo entre imagem e legenda e processar os dois quase que simultaneamente, melhor será sua recepção, pois essa é uma das premissas da legendagem, segundo as autoras.

Em suma, na condição de legendagem padrão, as legendas de 2 linhas apresentaram padrões de leitura mais regulares do que as de 1 linha, enquanto na legendagem reversa os resultados foram opostos. A questão de a legendagem reversa apresentar resultados contrários do que o esperado pela literatura na área está relacionada à própria natureza desse tipo de legendagem: ter o áudio da língua materna à disposição faz com que o comportamento leitor seja menos regular.

O trabalho de d'Ydewalle e De Bruycker (2007) é relevante para esta pesquisa por ter investigado um tema semelhante ao nosso: a influência do número de linhas no processamento dos espectadores. Desse trabalho, por exemplo, extraímos várias das medidas



para o estudo aqui empreendido, como regressão, tempo de latência e duração média das fixações. Além disso, é um dos poucos trabalhos que, até então, focaram mais o estudo do processamento de legendas do que a alocação de atenção em relação a legendas de 1 linha e 2 linhas. Esse trabalho foi o que forneceu com mais precisão os dados sobre os quais formulamos nossas hipóteses: legendas de 1 linha são mais custosas para o processamento de legendas não só para ouvintes, mas também para surdos.

Outro estudo realizado que foi de bastante importância para o presente trabalho foi Vieira (2016), o qual utilizou trechos de documentários televisivos para avaliar a influência da velocidade e da segmentação linguística na legendagem por surdos e ouvintes. A primeira parte do estudo analisou o processamento na movimentação ocular de legendas com velocidades baixa (145 ppm) e alta (180 ppm) com e sem problemas de segmentação linguística. A segunda parte visou analisar as respostas dos participantes em relação aos parâmetros em questão nos relatos e questionários durante e após o experimento. Contando com 16 participantes (8 surdos e 8 ouvintes), os participantes assistiram a quatro vídeos com legendas que refletiam quatro condições experimentais distintas: i) lenta bem segmentada (LBS); ii) lenta mal segmentada (LMS); iii) rápida bem segmentada (RBS); e iv) rápida mal segmentada (RMS).

Para construção do *corpus*, a autora retirou as legendas do tipo *roll-up* transmitidas via satélite do programa em questão e as manipulou para transformá-las no formato *pop-on* para que a variável relacionada ao modo de exibição da legenda na tela fosse evitada. Além desse aspecto, as legendas também foram manipuladas para que ficassem na velocidade de 145 ppm e 180 ppm, assim como foram manipulados os problemas de segmentação linguística, conforme objetivo da pesquisa. Essa manipulação, na prática, significou a relegendagem dos vídeos feita a partir do *software* de legendagem *Subtitle Workshop 2.51*, respeitando os outros parâmetros técnicos e linguísticos inerentes à prática da legendagem.

Tendo como objetivo o processamento de legendas com problemas de segmentação por surdos, Vieira formou um dos grupos dos participantes com pessoas ouvintes, as quais seriam expostas às legendas sem som, para que houvesse uma replicação (ou tentativa de) da condição semelhante à que o surdo seria exposto.

Os participantes foram submetidos ao rastreador ocular da marca *Tobii TX300*, o qual gerou dados que foram analisados com o auxílio do *software Tobii Studio* em sua versão 3.2.2. Divididos em um grupo de controle (ouvintes) e um experimental (surdos), todos os participantes foram submetidos a todas as condições experimentais (LMS, LBS, RMS e RBS) a partir de uma divisão chamada quadrado latino (MARQUER, 2003), conforme a Figura 4 abaixo:

**Figura 4 – Vídeos e condições experimentais**

Quadrado latino	
<p>Vídeo 1:</p> <p>LBS – ouvinte 1/surdo 1; ouvinte 5/surdo 5;  LMS – ouvinte 2/surdo 2; ouvinte 6/surdo 6;  RBS - ouvinte 3/surdo 3; ouvinte 7/surdo 7;  RMS - ouvinte 4/surdo 4; ouvinte 8/surdo 8;</p>	<p>Vídeo 3:</p> <p>LBS – ouvinte 3/surdo 3; ouvinte 7/surdo 7;  LMS – ouvinte 4/surdo 4; ouvinte 8/surdo 8;  RBS - ouvinte 1/surdo 1; ouvinte 5/surdo 5;  RMS - ouvinte 2/surdo 2; ouvinte 6/surdo 6;</p>
<p>Vídeo 2:</p> <p>LBS – ouvinte 2/surdo 2; ouvinte 6/surdo 6;  LMS – ouvinte 3/surdo 3; ouvinte 7/surdo 7;  RBS - ouvinte 4/surdo 4; ouvinte 8/surdo 8;  RMS - ouvinte 1/surdo 1; ouvinte 5/surdo 5;</p>	<p>Vídeo 4:</p> <p>LBS – ouvinte 4/surdo 4; ouvinte 8/surdo 8;  LMS – ouvinte 1/surdo 1; ouvinte 5/surdo 5;  RBS - ouvinte 2/surdo 2; ouvinte 6/surdo 6;  RMS - ouvinte 3/surdo 3; ouvinte 7/surdo 7;</p>

Fonte: Vieira, Teixeira e Chaves<sup>30</sup> (2017, p. 510).

Após a manipulação, cerca de 360 legendas de duas linhas foram analisadas no total (80 na condição LBS; 80 na condição LMS; 100 na condição RBS e 100 na condição RMS), ou seja, cerca de 720 linhas de legenda. Por outras questões relacionadas à legendagem (fluxo de fala, tempo de exposição em tela etc.), os vídeos contaram com algumas legendas de apenas uma linha, legendas estas que foram descartadas para análise devido ao objetivo da pesquisa: análise da influência da segmentação linguística entre linhas na recepção; ao contrário de Monteiro (2016), que analisou a influência de problemas de segmentação linguística entre legendas, trabalho que será discutido adiante.

As variáveis dependentes (medidas) utilizadas na pesquisa foram: 1) número de fixações; 2) duração média das fixações; 3) deflexões; 4) regressões; 5) atraso da primeira fixação nas legendas (latência); 6) perdas do todo ou de partes da legenda. Os resultados da análise quantitativa mostraram que, quando a variável independente de velocidade aumentava (RBS e LBS), o número de fixações também aumentava em ambos os grupos. Contudo, quando os resultados foram comparados entre as condições, notou-se que o número de fixações dos participantes surdos diminuía nas condições que apresentavam uma má segmentação linguística (LBS: 564 x LMS: 537; RBS: 715 x RMS: 713), enquanto nos ouvintes esse número aumentava (LBS: 672 x LMS: 718; RBS: 867 x RMS: 882), sugerindo uma possível relação do parâmetro de segmentação linguística com o desconforto dos surdos em relação a legendas mal segmentadas.

Esses dados numéricos foram validados com a aplicação do teste estatístico não paramétrico de Kruskal-Wallis, que revelou diferença significativa na comparação de número

<sup>30</sup> Vieira, Teixeira e Chaves (2017) traz os resultados da tese de Vieira (2016) em formato de artigo. Deste artigo retiramos a Figura 3 em questão.

de fixações entre grupos [ $H(1) = 511,5$   $p < 0.05$ ]; ou seja, os ouvintes fazem mais fixações do que os surdos, no entanto, os surdos apresentam duração das fixações maiores do que os ouvintes. Isso está em acordo com o que asseveram Rayner, Barbara e Pollatsek (2013), que afirmam que aprendizes de segunda língua apresentam esse comportamento. O teste também ratificou a influência das variáveis de segmentação [ $H(1) = 11,09$   $p < 0.05$ ] e velocidade [ $H(1) = 13,03$   $p < 0.05$ ] no número de fixações.

Em relação à duração das fixações, os resultados apontaram que os surdos gastam mais tempo nas legendas consideradas lentas do que nas rápidas. Contudo, entre as condições que apresentavam velocidade baixa, não houve diferença significativa entre as condições bem e mal segmentadas (LBS: 300 ms x LMS: 299 ms), apenas nas rápidas (RBS: 268 ms x RMS: 292 ms). Diferentemente, os resultados dos participantes apresentaram diferença entre as condições com velocidade considerada lenta (LBS: 257 x LMS: 273) e não apresentaram diferença nas condições com velocidade rápida (RBS: 259 x RMS: 253). Esses resultados apontam na mesma direção de De Linde e Kay (1999), que observaram um maior tempo gasto em legendas com velocidade considerada lenta no seu trabalho (apesar de, a partir da nossa perspectiva, as duas velocidades testadas pela autora, 139 ppm e 74 ppm, serem lentas). Segundo Vieira (2016), esse processamento mais custoso em legendas lentas é comum, especialmente em surdos, não sendo apenas um reflexo da velocidade da legenda, como acreditam De Linde e Kay (1999). Os autores se embasam em Emmorey *et al.* (2008) para tal afirmação, pois “devido ao uso de uma língua visual e sinalizada, os surdos nativos tendem a ter mais acuidade na parafóvea, permitindo previsões sobre os próximos caracteres durante a leitura” (VIEIRA; TEIXEIRA; CHAVES, 2017, p. 513).

Esses dados foram submetidos ao teste estatístico ANOVA, o qual mostrou uma tendência de os dois grupos lerem mais rapidamente as legendas consideradas rápidas (surdos: 281 ms e ouvintes: 266 ms), apesar de os surdos gastarem mais tempo processando as legendas de uma forma geral. Contudo, esses dados não foram considerados significativos após o teste quando considerada a interação *grupo vs velocidade*. Diferentemente, a interação *grupo vs segmentação vs velocidade* ratificou a influência da segmentação na velocidade desse processamento.

No que diz respeito às regressões durante a exposição ao experimento, os participantes ouvintes apresentaram um número maior de regressões nas legendas rápidas, sendo, contudo, a diferença mais acentuada nas legendas com problemas de segmentação. Em relação à duração média dessas regressões, esses participantes passaram mais tempo lendo as legendas com problemas de segmentação. Os participantes surdos, por sua vez, apresentaram

um número semelhante de regressões em todas as condições; entretanto, gastaram mais tempo relendo as legendas consideradas lentas. Ainda nos dados quantitativos, na condição RBS os surdos apresentaram uma menor duração de regressões.

Em relação aos movimentos regressivos, o teste Kruskal-Wallis não mostrou efeito significativo entre o número de fixações e as variáveis de velocidade e segmentação. Os resultados estatísticos do teste ANOVA mostraram uma significância na interação *grupo vs segmentação vs velocidade*, apontando, assim, para uma influência da segmentação na duração dos movimentos regressivos. Além disso, a condição RBS foi a que causou um tempo menor de regressões nos participantes surdos. Citando Reichle, Rayner e Pollatsek (2003), a autora enfatiza que as regressões são originadas normalmente no processamento linguístico de “sentenças estruturalmente difíceis” (VIERA, 2016, p. 48), o que levantaria a questão dos problemas de segmentação e sua interferência nesses movimentos regressivos.

Ao observar em que lugar da oração esses movimentos regressivos eram mais comuns, Vieira (2016) observou que havia mais movimentos regressivos dentro dos sintagmas, especialmente nas condições mal segmentadas, em ambos os grupos de participantes. Quanto maior o tempo gasto com regressões durante o processamento da legenda, maior o tempo total gasto no processamento (RAYNER; LIVERSEDGE, 2004).

No que diz respeito à variável deflexão, que remete ao movimento de mudança da área da legenda para a área da imagem em produtos legendados, algo já estudado por De Linde e Kay (1999) e d’Ydewalle e De Bruycker (2007), Vieira (2016) considerou dois movimentos: 1) o participante lia a legenda como um todo e depois ia para a área da imagem e, caso houvesse ainda tempo, voltava a fixar na área da legenda e a relia; e 2) o participante lia uma parte da legenda, ia para a área da imagem e voltava para terminar de ler o texto da legenda (nesta, gastava alguns milissegundos tentando localizar onde havia parado). Esta segunda opção foi considerada confortável pelo estudo, diferentemente da percepção de De Linde e Kay (1999) e d’Ydewalle e De Bruycker (2007) sobre os movimentos de deflexão. Os resultados da pesquisa demonstraram que o número de deflexões na imagem nas condições bem segmentadas foi maior do que nas condições mal segmentadas (com exceção da condição RMS para os surdos).

Por meio do teste estatístico Kruskal-Wallis, percebeu-se o efeito significativo do grupo em relação ao número de deflexões, ou seja, os grupos apresentaram comportamentos diferentes. Além disso, a variável velocidade apresentou efeito significativo, com as legendas lentas apresentando mais fixações deflexivas, assim como a variável segmentação, com mais fixações nas legendas com boa segmentação. O teste ANOVA, por sua vez, demonstrou que houve diferença dos grupos em relação à quantidade das deflexões no que diz respeito à variável

de velocidade, tendo os surdos apresentado um maior número de deflexões nas legendas rápidas e os ouvintes, nas lentas. Contudo, esse mesmo teste não apresentou efeito significativo da segmentação e da velocidade na performance dos participantes em relação à duração das deflexões (VIEIRA, 2016).

Considerando o atraso nas fixações causados por legendas com problemas de segmentação, o que chamamos aqui de **latência**, Vieira (2016) observou que estas realmente afetam essa variável dependente em questão. Foram as legendas nas condições rápidas as que causaram mais atraso para os ouvintes, enquanto, para os surdos, foram as condições lentas que causaram mais dificuldade.

De um modo geral, os resultados da parte experimental e da exploratória sugeriram que legendas mal segmentadas, além de causarem desconforto, trazem maior dificuldade no processamento da legenda, sendo a condição rápida e bem segmentada a mais confortável para os participantes, especialmente para os surdos. Como a parte experimental do trabalho de Vieira (2016) trouxe resultados significativos, não irei relatar em mais detalhes os dados provenientes dos questionários aplicados e relatos dos participantes, algo que farei a seguir em Monteiro (2016). Diferentemente de Vieira (2016), Monteiro não fez tratamento estatístico dos seus dados, conforme será detalhado mais adiante.

O trabalho de Vieira (2016) é fundamental para a realização da presente pesquisa dado o seu objeto de estudo (LSE) alinhado à metodologia baseada em rastreamento ocular, algo que poucos trabalhos realizaram até então. Além disso, a maior parte do seu *design* experimental foi replicada na presente pesquisa: perfil dos participantes, divisão de tarefas a partir do quadrado latino, assim como uma boa parte das medidas utilizadas (variáveis dependentes).

Monteiro (2016), por sua vez, utilizou os mesmos preceitos metodológicos que Vieira (2016), além dos mesmos objetivos e hipóteses de pesquisa<sup>31</sup>, tendo como objetivo a análise da recepção de LSE de campanhas políticas veiculadas na TV. A autora teve como estímulos do experimento 4 vídeos de programas políticos de partidos distintos transmitidos durante as eleições de 2010 na cidade de Fortaleza, tendo escolhido cada vídeo de um partido diferente, de modo que os participantes fossem expostos a diferentes situações de campanha na TV. Os vídeos utilizados apresentavam as campanhas dos candidatos a governador do estado e a presidente por “fornecerem mais informações que pudessem ser exploradas durante a pesquisa” (MONTEIRO, 2016, p. 64). Os 4 vídeos apresentavam, assim como em Vieira

---

<sup>31</sup> As duas pesquisas também utilizaram os mesmos participantes por terem acontecido concomitantemente dentro do mesmo grupo de pesquisa, o LEAD.

(2016), as seguintes condições experimentais: 1) legendas lentas e mal segmentadas (LMS); 2) legendas lentas e bem segmentadas (LBS); 3) legendas rápidas e mal segmentadas (RMS); e 4) legendas rápidas e bem segmentadas (RBS).

Os vídeos escolhidos para a pesquisa foram editados até que perfizessem 1 minuto. Além do fato de o rastreador ocular gerar uma grande quantidade de dados a ser analisada, trabalhar com o vídeo em sua duração original poderia ser cansativo para os participantes. Ao serem manipuladas as quatro condições experimentais, algo que perpassou a manipulação dos vídeos foi a necessidade de eles conterem começo, meio e fim claros (MONTEIRO, 2016).

As perguntas que embasaram a pesquisa foram as seguintes: 1) Até que ponto legendas de campanhas políticas lentas bem segmentadas facilitam a recepção de surdos e ouvintes? 2) Até que ponto legendas de campanhas políticas lentas mal segmentadas dificultam a recepção de surdos e ouvintes? 3) Até que ponto legendas de campanhas políticas rápidas bem segmentadas facilitam a recepção de surdos e ouvintes? 4) Até que ponto legendas de campanhas políticas rápidas mal segmentadas dificultam a recepção de surdos e ouvintes? (MONTEIRO, 2016).

O processo metodológico da pesquisa foi o mesmo utilizado por Vieira<sup>32</sup> (2016) no que tange ao perfil dos participantes, assim como aos instrumentos utilizados (relato retrospectivo livre; questionário pós-coleta; relato retrospectivo guiado). Em relação aos procedimentos metodológicos, após cada participante assistir aos vídeos em todas as condições, ele fazia seu relato livre sobre o conteúdo do vídeo e respondia às perguntas do questionário pós-coleta. Após isso, o participante era convidado a se retirar e as pesquisadoras analisavam o percurso do seu olhar. Em seguida, o participante era chamado para observar o percurso do seu olhar a partir dos dados obtidos por meio do rastreador ocular, respondendo a algumas perguntas que a pesquisadora achava interessante a respeito do seu comportamento ocular (exemplo: Por que fez fixações muito longas nesta palavra? Por que olhou mais para a legenda do que para a imagem em tal cena?) (MONTEIRO, 2016).

O foco da parte quantitativa do trabalho foi a análise do número de fixações e do tempo de duração destas na legenda e na imagem, já que houve uma quantidade grande de fixações e uma grande duração de fixação. Os resultados obtidos pela pesquisa mostraram que, a partir do número de fixações, as condições RMS e LMS foram as mais problemáticas em termos de recepção para os ouvintes no número de fixações na imagem (LMS: 44,62/ RMS:

---

<sup>32</sup> Uma coisa que diferenciou o trabalho de Monteiro (2016) do de Vieira (2016), entretanto, foi que esta não realizou a análise de fixação por fixação através do acompanhamento do olhar dos participantes. Monteiro (2016), por sua vez, utilizou a análise automática das fixações fornecida pelo *Tobii Studio*.

33,37) e no número de fixações na legenda (LMS: 193,625 /RMS: 195,37). Para os surdos, a pior condição foi a RMS, em termos de conforto na recepção, a partir dos dados do rastreador ocular, tendo o número de fixações na legenda sido o maior de todas as condições (175,50) e o número de fixações na imagem sido o menor (34,37).

Em relação à duração das fixações, os ouvintes tiveram o mesmo resultado nas condições LBS e LMS (266,25 ms), não deixando claro o papel da segmentação linguística na recepção dessas duas condições. Contudo, a condição RBS foi a que apresentou o menor tempo de fixação nas legendas (256,25 ms), sugerindo uma melhor recepção dessa condição. Ao mesmo tempo, a condição RMS foi a que apresentou o maior tempo (268,75 ms). Os surdos, por sua vez, obtiveram resultados semelhantes nas condições LBS (292,5 ms), LMS (291,2 ms) e RBS (292,5 ms). Apenas a condição RMS (302,5 ms) apresentou um tempo maior de duração (MONTEIRO, 2016).

Ao contrário do que pressupõem Perego *et al.* (2010) sobre uma maior duração da fixação na imagem do que nas legendas, os surdos apresentaram um maior tempo de fixação na legenda do que na imagem na condição RBS (292,5 ms vs. 281,25 ms). Ao contrário do que seria esperado caso a velocidade tivesse influência na recepção pelos participantes, o tempo de fixação nas legendas com a condição RBS foi similar às demais condições. A partir disso, Monteiro (2016) questiona o fato de a sua pesquisa ter utilizado apenas legendas de 1 linha, prática comum na legendagem de propagandas políticas no Brasil. Para a autora, isso pode ter influenciado os resultados.

Como um todo, no que diz respeito aos dados do rastreador ocular, a condição RMS se mostrou desconfortável para os dois grupos de participantes, enquanto na RBS o conforto foi maior para os ouvintes por permitir um maior tempo de fixação na imagem do que na legenda, ao contrário do que ocorreu com os surdos. Em relação às condições LBS e LMS, os resultados destas foram bem parecidos, sendo que o tempo de fixação na imagem na condição LMS foi menor, sugerindo um desconforto maior.

Outro dado apresentado por Monteiro (2016) é que o número de fixações dos surdos foi menor do que o dos ouvintes nas legendas, algo esperado para leitores com menor proficiência. Segundo Rayner, Barbara e Pollatsek (2013), leitores mais proficientes fazem mais fixações no texto, mas mais curtas. A partir dos dados do rastreador ocular e da parte qualitativa proveniente dos relatos e questionários dos participantes, Monteiro (2016) respondeu às suas hipóteses, como se verá a seguir.

Sua hipótese 1, a de que legendas de campanhas políticas lentas bem segmentadas facilitavam a recepção de surdos e ouvintes, foi confirmada. Segundo os resultados da pesquisa,

ouvintes e surdos demonstraram apresentar conforto nessa condição, pois os dois grupos fizeram fixações mais longas na imagem do que na legenda, o que sugere que a recepção das legendas foi facilitada. Da mesma forma, os dados da pesquisa exploratória também convergiram para isso, pois os participantes relataram o conteúdo e os detalhes dos vídeos, tendo inclusive mencionado que achavam que as legendas estavam em velocidade normal (MONTEIRO, 2016).

A hipótese 2, por sua vez, a de que legendas de campanhas políticas lentas mal segmentadas dificultavam a recepção de surdos e ouvintes, também foi confirmada. Segundo os dados provenientes do rastreador ocular, tanto surdos como ouvintes tiveram maior dificuldade nesta condição, pois ambos os grupos apresentaram maior tempo de fixação na legenda do que na imagem. Em relação à parte exploratória, o relato dos ouvintes sugeriu que eles conseguiram captar os detalhes do vídeo (inclusive melhor do que os surdos). Estes, inclusive, acharam as legendas rápidas, o que pode indicar um desconforto e consequente recepção dificultada (MONTEIRO, 2016).

A hipótese 3, a de que legendas de campanhas políticas rápidas bem segmentadas facilitavam a recepção de surdos e ouvintes, foi confirmada pelos ouvintes, mas confirmada apenas parcialmente pelos surdos, pois os dados da pesquisa exploratória divergiram dos resultados do rastreador. Monteiro (2016) pondera que, na análise vídeo a vídeo, a questão do conforto/desconforto não ficou clara para os grupos de ouvintes e surdos, apesar de os ouvintes terem tido uma maior duração de fixação na imagem do que na legenda, ao contrário dos surdos, indicando um possível maior conforto por parte dos ouvintes. Em relação à pesquisa exploratória, um número maior de surdos relatou os detalhes dos vídeos se comparado com o número de ouvintes, apesar de terem tido mais reclamações sobre a imagem e as palavras desconhecidas. Esta questão de pesquisa fez a autora levantar a hipótese de que o número de linhas poderia ser prejudicial na recepção de legendas de campanhas políticas (MONTEIRO, 2016).

A hipótese 4, de que as legendas de campanhas políticas rápidas mal segmentadas dificultavam a recepção de surdos e ouvintes, foi confirmada pelos surdos; os ouvintes, por sua vez, confirmaram-na parcialmente, visto que os dados exploratórios apenas sugeriram certo desconforto. Na parte experimental da pesquisa, os dados mostraram que os ouvintes tiveram duração de fixação maior na imagem do que na legenda, ao contrário dos participantes surdos. A partir dos dados da pesquisa exploratória, foi possível perceber que os surdos tiveram mais dificuldade em relatar o conteúdo do vídeo, dado que quatro surdos não falaram muito sobre esse conteúdo. Os ouvintes também tiveram um aproveitamento de detalhes menor nas legendas



rápidas mal segmentadas do que nas outras condições experimentais. No geral, ouvintes e surdos reclamaram das legendas testadas, mencionando a presença de frases quebradas (MONTEIRO, 2016).

Os resultados de Monteiro (2016) sugeriram a influência de uma segmentação linguística adequada para garantir a boa recepção da LSE. Não houve, contudo, resultados que apontassem influência da velocidade da legenda no processamento da leitura dos participantes. No entanto, conforme apontado pela própria autora, um tratamento estatístico dos dados, conforme realizado em Vieira (2016), poderia refinar os seus achados, assim como refutar ou reforçar os resultados do seu trabalho.

Sobre a relevância do trabalho de Monteiro (2016) para a presente pesquisa, reitero o seu *design* experimental, o qual será levado em conta em vários aspectos; no que diz respeito a algumas medidas, número e perfil de participantes, assim como a divisão dos grupos experimental e controle. Outra questão apontada pela pesquisa – e importante para a realização desta – foi a não confirmação dos seus dados em relação a legendas RBS, o que, conforme apontado pela própria autora, pode ter sido motivado pelo número de linhas normalmente utilizado no gênero LSE de campanha política. Esta interpretação dos seus dados, assim como o que já foi exposto anteriormente em relação à natureza distinta do processamento de legendas de 2 linhas, fazem-nos crer que realmente tenha sido a causa de dados um pouco divergentes quando comparados com Vieira (2016), a qual utilizou o mesmo *design* e procedimentos metodológicos, assim como os mesmos participantes, em sua pesquisa.

Contudo, um ponto relevante para a interpretação dos resultados de sua pesquisa foi a falta de tratamento estatístico dos dados, algo comum em pesquisas experimentais e que trabalhem com dados oriundos de rastreamento ocular. Esta questão coloca em dúvida alguns dos seus achados, pois o efeito significativo da interação entre variáveis e performance entre grupos fica incerto. Dessa forma, os procedimentos de análise de Vieira (2016) seriam mais adequados pelo sucesso que ela teve, principalmente, com os dados quantitativos sobre fixações e o tratamento estatístico desses dados.

Souza (2020) teve como objetivo do seu trabalho de dissertação de mestrado examinar os dados de Monteiro (2016). Sendo assim, fez uso das mesmas hipóteses de Monteiro (2016) em relação às legendas bem e mal segmentadas de velocidade lenta (145 ppm) e rápida (180 ppm) processadas por surdos e ouvintes. Além disso, Souza (2020) analisou os dados provenientes do rastreamento ocular de maneira semelhante a Vieira (2016); ou seja, fazendo um acompanhamento do olhar e utilizando tratamento estatístico para a análise dos dados. Para tal, a autora examinou cinco medidas de processamento: 1) duração da primeira fixação; 2)

*gaze duration* – duração média da fixação; 3) número de fixações; 4) regressão (ocorrência); e 5) deflexão (ocorrência).

Para o tratamento estatístico, a autora utilizou Modelos de Regressão Linear Mistos para as variáveis dependentes (medidas) numéricas e Modelos de Regressão Logística Mistos para as variáveis dependentes categóricas, semelhantemente ao que fizemos aqui. Uma diferença em relação ao tratamento estatístico da autora, entretanto, foi em relação ao contraste das variáveis. Enquanto para o contraste das variáveis no modelo utilizamos um *dummy coding*, a autora utilizou um contraste de soma entre as variáveis, no qual o valor do intercepto representa a média geral e os coeficientes são positivos se os efeitos estiverem na direção esperada. Mais à frente, ao explicarmos o modelo que utilizamos para nossa análise, damos mais detalhes sobre esses contrastes e sobre os aspectos que configuram essas modelagens estatísticas.

Em relação à medida da duração da primeira fixação, Souza (2020) não encontrou significância estatística em nenhuma das suas variáveis independentes (segmentação, velocidade e grupo).

Ao observar a *gaze duration*, soma de todas as fixações feitas por palavra, a autora observou uma interação entre segmentação e velocidade. Essa interação apontou que tanto ouvintes como surdos tiveram um processamento mais rápido em legendas rápidas e bem segmentadas. Nessa medida, a hipótese de que as legendas mal segmentadas custariam um tempo maior de processamento em surdos e ouvintes foi confirmada. Além disso, a hipótese de que velocidade não teria influência no processamento dos participantes foi refutada, visto que a variável de velocidade apresentou interação com segmentação; ou seja, a boa segmentação teve uma duração menor de fixação em ambos os grupos na condição rápida.

Em relação ao número de fixações, os participantes fizeram menos fixações nas legendas rápidas e bem segmentadas, o que a autora interpretou como sinal de um processamento confortável. Mesmo havendo efeito de segmentação na velocidade lenta, o efeito de segmentação foi mais forte na condição rápida e bem segmentada. Além disso, os surdos tiveram um maior número de fixações na condição lenta e bem segmentada, diferentemente dos ouvintes, que tiveram um número maior de fixações nas legendas lentas mal segmentadas. Assim, a autora aponta que a hipótese sobre velocidade foi novamente refutada em relação a essa medida, visto que houve diferença entre os níveis dessa variável. Sobre a variável de segmentação, o efeito de segmentação foi significativo, confirmando a sua hipótese de que legendas bem segmentadas apresentam um menor número de fixações.

Em relação ao número de regressões, tanto os participantes surdos quanto ouvintes fizeram mais regressões nas legendas lentas, sendo que, nas legendas rápidas, o grupo de surdos teve uma queda acentuada na ocorrência de regressões. A segmentação, no caso, não apresentou efeito significativo nessa medida. Em relação a isso, a autora indica que o maior tempo de exposição em tela das legendas lentas pode estar relacionado a esse efeito. Além disso, Souza aponta que essas regressões, inclusive, podem indicar releituras. Ademais, os surdos realizaram menos regressões no seu experimento em relação aos ouvintes. A condição rápida e bem segmentada representou um maior conforto para os participantes, pois foi a condição com menos regressões.

Esses resultados divergiram da hipótese do seu trabalho, de que a velocidade não seria uma variável significativa no processamento de surdos e ouvintes expostos a campanhas políticas legendadas. Sobre a hipótese de que as legendas bem segmentadas apresentariam um menor número de regressões, sendo indício de um menor custo de processamento, a autora observou que a segmentação não teve efeito significativo, apesar de a condição rápida e bem segmentada ter se mostrado mais confortável em ambos os grupos. Sendo assim, essa hipótese não pode ser confirmada.

Em relação ao número de deflexões, a autora observou que os participantes surdos e ouvintes realizaram mais deflexões nas legendas lentas tanto em legendas com boa segmentação quanto com má segmentação, divergindo de Vieira (2016), que observou um maior número de deflexões na condição rápida e bem segmentada por surdos. Essa queda no número de deflexões nas legendas de velocidade rápida foi mais acentuada no grupo de surdos do que no de ouvintes. Sendo assim, a hipótese de que a variável de velocidade não influenciaria o comportamento ocular dos participantes do estudo foi refutada nessa medida, visto que as legendas lentas tiveram um maior número de deflexões. A hipótese de que a boa segmentação causaria um número maior de deflexões também foi refutada, visto que não houve diferença significativa na medida de deflexão entre legendas bem e mal segmentadas.

De uma forma geral, Souza (2020) aponta que a condição mais confortável, considerando as medidas analisadas, foi a rápida bem segmentada. Sendo assim, sua hipótese em relação à segmentação foi confirmada, visto que a boa segmentação facilitou o processamento de ouvintes e surdos. Em relação à hipótese sobre velocidade, de que esta não teria diferença significativa em relação aos níveis de 145 ppm e 180 ppm, ela foi refutada, visto que ela foi significativa em todas as medidas analisadas. Levando em conta as medidas analisadas, a velocidade rápida (180 ppm) foi mais confortável, segundo a autora.

O trabalho de Souza (2020) é importante por utilizar metodologia e procedimento de análise da movimentação ocular semelhantes aos do presente trabalho. Ademais, o tratamento estatístico através de modelos de regressão com efeitos mistos é algo que também nos propusemos a fazer. Uma consideração que a autora faz sobre a regressão em legendagem também é algo que nos chamou a atenção e irá embasar a discussão dos nossos dados. Primeiramente, que legendas mais lentas aparentemente propiciam a ocorrência de regressão em vídeos legendados, parecendo algo mais associado ao maior tempo de exposição em tela dessas legendas. Além disso, o fato de ouvintes, leitores mais proficientes, fazerem mais regressões do que surdos. Esses dois pontos sobre a ocorrência de regressão nos fazem questionar o papel dessa medida na análise do processamento de legendas, assim como o fizeram Kruger e Steyn (2014). Contudo, como ainda não é algo claro na literatura da área e é uma medida historicamente utilizada nesse tipo de investigação de processamento, decidimos mantê-la para a análise deste trabalho.

Um ponto, contudo, de que inicialmente discordamos é sobre a leitura da autora em relação à medida de número de fixações por legenda. De acordo com Rayner, Barbara e Pollatsek (2013), leitores mais proficientes fazem mais fixações no texto, mas mais curtas do que leitores menos proficientes. Assim, é esperado que condições experimentais favoráveis estimulem um maior número de fixações mais curtas; especialmente na leitura de um texto dinâmico, em que uma menor quantidade de fixações pode estar associada ao fato de o telespectador não ter processado a legenda por inteiro porque ela desapareceu antes que ele assim o fizesse. Assim, essa concepção é algo que está relacionado à leitura dessa variável em nossos dados. Essa discussão, contudo, será aprofundada durante a discussão dos nossos resultados.

Kruger e Steyn (2014) tiveram como objetivo investigar como estudantes sul-africanos em um contexto universitário de inglês como L2 são impactados pelo processamento de legendas em relação ao conteúdo de uma disciplina de Psicologia. Além disso, os autores propõem uma medida mais detalhada para se mensurar quantitativamente até que ponto os participantes leram as legendas do experimento. A essa medida deram o nome de Índice de Leitura de Textos Dinâmicos (*Reading Index for Reading Dynamic Texts – RIDT*), o qual se constitui no produto do número de fixações únicas (sem regressões ou refixações) por tamanho médio de palavra (do *corpus*, no caso, as legendas) dividido pelo tamanho médio da sacada na legenda dividido pelo tamanho médio da palavra nas legendas. O fato de não considerarem as refixações (estando entre duas fixações e com um tamanho menor que dois caracteres) e

regressões (com tamanho menor ou igual a dois caracteres), além de varreduras de retorno<sup>33</sup>, que segundo os autores, deve-se ao fato de que, nesses casos, uma informação nova não é obtida no processamento. Apesar de refixações e regressões serem um indicativo de que há uma dificuldade no processamento da palavra, ou baixa proficiência leitora, são medidas que deveriam ser menos consideradas no caso de textos dinâmicos, como o da legendagem, visto que os leitores, diferentemente de textos estáticos, não têm controle sobre a apresentação das legendas (KRUGER; STEYN, 2014, p. 110).

Em relação aos resultados da sua pesquisa, os autores não encontraram evidência de que a simples presença de legendas melhora a performance acadêmica dos estudantes participantes. Contudo, o material legendado ainda pode ser visto como fonte que enriquece o campo educacional, haja vista a “[...] alta correlação entre a leitura de legendas e desempenho” (KRUGER; STEYN, 2014, p. 118, tradução nossa). Além disso, os autores não observam diferenças significativas no processamento de legendas de 1 e 2 linhas de acordo com seu modelo de análise, o que fortalece a necessidade do nosso estudo por ser algo que é contrário ao que foi observado por outros autores (D’YDEWALLE *et al.*, 1987; D’YDEWALLE *et al.*, 1991; D’YDEWALLE; DE BRUYCKER, 2007; D’YDEWALLE; GIELEN, 1992; PRAET *et al.*, 1990).

Sobre o Índice de Leitura de Textos Dinâmicos, os autores apontam ter sido bem-sucedidos e ser um caminho possível para uma análise maior de dados no campo da pesquisa de processamento de legendas, sem a necessidade de uma inspeção manual tão laboriosa por parte dos pesquisadores. Contudo, o fato de ser um modelo que não acompanha as especificidades do comportamento ocular dos participantes, assim como não considera outras medidas relevantes para o estudo do processamento da legenda, como a regressão, e medidas que envolvam a imagem (como deflexões, por exemplo), o modelo não foi seguido pela maior parte das pesquisas que a sucederam. Apesar de ser uma possibilidade metodológica, escolhemos não a utilizar pelos motivos supracitados, visto que não se tornou uma prática canônica na área. Porém, outras perspectivas do estudo são importantes para este trabalho, como a forma de se tratar os dados advindos do processamento de legendas com rastreamento ocular. A partir da literatura e pesquisa em leitura, Kruger e Steyn (2014, p. 106-107) citam uma lista (que não é fechada) de 7 parâmetros básicos para se inspecionar a leitura de legendas:

---

<sup>33</sup> Varreduras de retorno (*return sweeps*) são sacadas longas que vão contra o movimento da leitura (direita para a esquerda) maiores que o tamanho médio de uma palavra, como já colocado anteriormente. No contexto da legendagem, também pode ser um movimento vertical de cima para baixo (de uma linha para outra da legenda) que seja maior que a distância entre as duas linhas da legenda (KRUGER; STEYN, 2014, p. 110-111).

- a) uma palavra pode ser considerada processada se houve pelo menos uma fixação nela;
- b) palavras curtas de até três caracteres à esquerda de uma fixação podem ser consideradas processadas;
- c) palavras inteiras dentro do espaço perceptual de oito caracteres à direita de uma fixação podem ser consideradas processadas mesmo que não tenham sido fixadas;
- d) palavras curtas de dois ou três caracteres que não foram fixadas, mas estão entre duas fixações mais longas, podem ser consideradas processadas a partir da noção de leitura preditiva mesmo fora do espaço perceptual;
- e) apenas palavras fixadas sucessivamente da esquerda para a direita podem ser consideradas processadas;
- f) o local de pouso da primeira fixação na área da legenda, muitas vezes, pode não estar relacionado ao processo de leitura, indo mais ao centro da legenda antes que um movimento de regressão leve o olhar para o início da legenda; nesses casos, a palavra fixada no local de pouso não deve ser considerada processada a não ser que uma fixação subsequente seja feita na palavra ou inclua a palavra em questão no espaço perceptual;
- g) no caso de dúvida, a lógica da leitura deve ser colocada em detrimento da noção de fixação aleatória.

Considerações desse tipo foram levadas em conta durante a inspeção do movimento ocular dos participantes submetidos a esta pesquisa. Na próxima seção, apresentamos o percurso metodológico do presente trabalho.

### 3 METODOLOGIA

Nesta seção, detalhamos o percurso metodológico traçado para o desenvolvimento da tese. A pesquisa proposta tem o suporte teórico-metodológico dos Estudos da Tradução, mais especificamente da Tradução Audiovisual Acessível e dos Estudos Experimentais com rastreamento ocular.

#### 3.1 Tipo de pesquisa

Trata-se de uma pesquisa experimental de natureza quantitativa. A partir de subsídios de estudos com legendagem e rastreamento ocular (D'YDEWALLE *et al.*, 1987; D'YDEWALLE; DE BRUYCKER, 2007; D'YDEWALLE; GIELEN, 1992; DE LINDE; KAY, 1999; GHIA, 2012; KRUGER; STEYN, 2014; KRUGER; SZARKOWSKA; KREJTZ, 2015; MONTEIRO, 2016; ORERO *et al.*, 2018, PEREGO *et al.*, 2010; VIEIRA, 2016; dentre outros). Pretendemos caracterizar o processamento de legendas por surdos e ouvintes utilizando, como subsídio, os dados quantitativos<sup>34</sup> gerados por rastreamento ocular, de modo a estipular o nível de conforto de participantes surdos e ouvintes quando expostos a legendas rápidas (180 ppm) e lentas (145 ppm) de 1 ou 2 linhas.

---

<sup>34</sup> Para o momento deste trabalho, iremos descrever apenas a parte quantitativa dos dados – relativo às fixações – e não a parte exploratória relativa ao relato retrospectivo e questionário pós-coleta utilizados por Vieira (2016) e Monteiro (2016), o que será feito em um momento futuro. Tomamos essa decisão com base no fato de que a parte quantitativa relativa às fixações é o cerne desta pesquisa, apesar de haver um apelo nos estudos experimentais em Tradução Audiovisual em torno de uma abordagem mista dos dados – aliando o qualitativo e o quantitativo, no que também acreditamos (ORERO *et al.*, 2018). Entretanto, observamos também que os principais resultados das pesquisas experimentais com LSE no Brasil (MONTEIRO, 2016; SOUZA, 2020; VIEIRA, 2016) e no mundo (D'YDEWALLE *et al.*, 1987, dentre outros) sobre o processamento de legendas partem da análise de medidas quantitativas enquanto indício de processamento de leitura. Além disso, Just e Carpenter (1980), os quais propuseram um modelo de leitura que é influente nos estudos de processamento até os dias de hoje e com os quais compartilhamos pressupostos teórico-metodológicos para nossa análise, apontam que o seu foco no tempo de processamento resultou em uma teoria que dá conta de características da leitura em tempo real. Com isso, seu modelo teórico também coloca menos foco na parte qualitativa do estudo do processamento, como o de reconstrução e recuperação de texto e consequente sumarização. Dessa forma, nosso estudo também coloca menos ênfase nos relatos retrospectivos e questionários pós-coleta, mesmo reconhecendo sua importância para o entendimento sobre uma compreensão mais global das legendas pelos participantes. Além disso, são elementos importantes porque fazem parte de uma tarefa à qual os participantes devem responder após assistirem aos vídeos legendados – a sumarização ou relato retrospectivos e o questionário pós-coleta; ou seja, o fato de terem que realizar essas tarefas após a exposição ao vídeo legendado no experimento faz com que prestem mais atenção ao estímulo utilizado e assistam ao vídeo como se estivessem assistindo a uma produção audiovisual em sua casa. Assim, apesar de não apresentar a análise dessa parte qualitativa da pesquisa neste trabalho, acreditamos na sua relevância para a performance dos participantes quando expostos aos vídeos legendados.

### 3.2 Contexto da pesquisa

O presente trabalho consiste em um recorte do projeto que responde pela sigla ExLeg II, que tem como objetivo o estudo da influência de parâmetros de legendagem (segmentação, velocidade, número de linhas etc.) no processamento por surdos e ouvintes de produções audiovisuais legendadas. Nesse sentido, a presente pesquisa amplia o escopo de parâmetros já estudados por pesquisas anteriores desse projeto (MONTEIRO, 2016; VIEIRA, 2016) a partir do estudo de número de linhas (além da velocidade), parâmetro este que ainda carece de estudos no campo da Tradução Audiovisual/Tradução Audiovisual Acessível.

Inicialmente, tinha como objetivo continuar o trabalho iniciado na graduação (ASSIS, 2013) e ampliado no mestrado (ARAÚJO; ASSIS, 2014; ASSIS, 2016), quando pesquisei, com auxílio da Linguística de *Corpus*, a questão da descrição e classificação de problemas de segmentação linguística na LSE de uma telenovela. Ao propor esta pesquisa de doutorado, havia pensado em continuar o estudo da segmentação linguística em LSE, mas agora focando a questão da sua influência no processamento dos espectadores surdos e ouvintes, algo semelhante ao que havia sido estudado por Vieira (2016) e Monteiro (2016), as quais participam do mesmo grupo de pesquisa que eu (LEAD). Contudo, por diferenciar o meu estudo apenas na questão do gênero analisado (LSE de telenovelas), decidi, em conjunto com minha orientadora e após conversas com colegas de laboratório, que esta seria uma lacuna pequena a ser preenchida, sendo que uma pesquisa de doutorado necessita de uma contribuição maior pelo seu próprio caráter inédito. A partir disso, veio a ideia de estudar a questão da influência de legendas (LSE) de 1 ou 2 linhas no processamento de surdos, algo que ainda não foi estudado por pesquisadores da área e que foi apontado, pelos estudos de Monteiro (2016), como um problema relevante a ser investigado sobre a recepção da LSE por surdos e ouvintes. Além do número de linhas, esta pesquisa se propôs a continuar os estudos sobre a velocidade da legenda (180 palavras por minuto e 145 palavras por minuto) e sua influência no processamento de surdos e ouvintes, algo que já havia sido feito em Monteiro (2016) e Vieira (2016), mas agora com um gênero distinto, no caso, um filme de drama.

A pesquisa como um todo está atrelada ao Grupo LEAD, que se constitui como uma linha de pesquisa do grupo Tradução e Semiótica, registrado no CNPq, e ao seu respectivo ambiente de trabalho, o LATAV (Laboratório de Tradução Audiovisual). A coleta de dados da pesquisa se daria, inicialmente, no âmbito do LERO (Laboratório Experimental de Rastreamento Ocular), situado nas dependências da Universidade Estadual do Ceará (UECE),



mais especificamente em um dos laboratórios do CCLIN (Cidade, Cultura e Linguagem), centro de pesquisa do Centro de Humanidades da UECE.

Contudo, tivemos um problema com o equipamento que havia sido utilizado até então em pesquisas anteriores (MONTEIRO, 2016; VIEIRA, 2016), o *Tobii TX300 Eye Tracker* (*hardware*), o qual parou de funcionar. Após um tempo considerável buscando conserto para a máquina, sem sucesso, fizemos contato com a profa. Dra. Elisangela Nogueira Teixeira, coordenadora do Laboratório de Psicolinguística e Ciências Cognitivas, vinculado ao Programa de Pós-Graduação em Linguística da Universidade Federal do Ceará, a qual nos cedeu (e continua cedendo), por meio da parceria de seu laboratório com os estudos desenvolvidos no LATAV, o rastreador ocular *Tobii Tx120 EyeTracker*.

### 3.3 Estímulos

Os estímulos utilizados no experimento desta pesquisa consistiram de 8 vídeos/fragmentos com LSE do filme *A rede social* (2010); este filme possui um alto número de legendas de 1 linha em sua legenda comercial em língua portuguesa, segundo a análise baseada em *corpus* empreendida por Chaves (2017)<sup>35</sup>. Esse foi um dos motivos iniciais de termos escolhido esse filme, mas não o preponderante. Na verdade, pretendíamos utilizar 4 dos 5 filmes analisados pela autora de modo que cada filme representasse uma das condições experimentais a ser testada (1 e 2 linhas; lenta e rápida), de modo semelhante à metodologia utilizada por De Linde e Kay (1999). Contudo, ao aprofundarmos nossa leitura em relação ao papel da imagem enquanto um dos canais geradores de informação durante o processamento de vídeos legendados, percebemos a complicação que seria (e quanta interferência poderia gerar) analisar filmes de gêneros distintos. Além disso, gêneros em que o diálogo e a imagem, principalmente – visto que o som não seria disponibilizado durante a execução do experimento –, possuem relevâncias distintas, acabam por induzir alocação de atenção – e processamento – de maneira diferente. Em um filme de ação ou uma animação, por exemplo, esses recursos são manipulados dentro de uma estrutura que, normalmente, privilegia a imagem em detrimento do diálogo (e, dessa forma, aquilo que é expresso pela legenda), o qual acaba por apresentar um papel secundário nos gêneros em questão.

---

<sup>35</sup> Chaves (2017) analisou a LSE de 5 filmes em seu trabalho: *Exôdo: deuses e reis*, *A rede social*, *Madagascar*, *Plano de voo* e *Click*. *A rede social*, dentre eles, foi o filme que apresentou o maior número de legendas de 1 linha na análise baseada em *corpus*.

*A rede social* (2010) é um filme americano dirigido por David Fincher e lançado em 2010. O filme é baseado no livro *Milionários acidentais* (*The accidental billionaires*) e conta a história de Mark Zuckerberg e como ele criou a rede social mais famosa do mundo, o *Facebook*. Ao longo do filme, somos apresentados ao nascimento do *Facebook* e aos desdobramentos (e intrigas) envolvidos no seu desenvolvimento. Esse desenvolvimento é retratado em forma de *flashback* durante o filme, que se mistura a cenas do presente, em que está ocorrendo o julgamento de Mark Zuckerberg em um processo movido por seu antigo colega de faculdade, Eduardo Saverin, em relação aos direitos sobre a criação do *Facebook*.

Por se tratar de um drama biográfico, o diálogo mantém um *status* relevante na compreensão do filme, favorecendo a investigação do processamento de legendas. Além disso, como grande parte de suas cenas envolve duas ou mais pessoas conversando em tela, há pouca redundância de informações entre legenda e imagem; isso também favorece (e, de certa forma, quase que obriga) o espectador a processar as legendas para entender o que acontece na cena.

Escolhido o filme, passamos às escolhas das cenas que comporiam os estímulos. A primeira questão que nos perpassou durante esse momento foi a do tamanho de cada estímulo. Nesse quesito, os estudos da área não estão em consenso sobre o tamanho ideal dos vídeos utilizados para o estudo de processamento de legendas. Perego *et al.* (2010), por exemplo, utilizaram vídeos de 15 minutos de duração, enquanto d'Ydewalle *et al.* (1987) utilizaram trechos de 10 minutos. De Linde e Kay (1999), Vieira (2016) e Monteiro (2016) trabalharam com vídeos menores, de 2 minutos e 1 minuto, respectivamente. No caso de De Linde e Kay (1999), contudo, os 30 segundos iniciais e finais do vídeo foram desconsiderados, fazendo com que 1 minuto de vídeo fosse o tempo analisado.

Sobre essa questão, Orero *et al.* (2018) afirmam que o tamanho do estímulo na pesquisa em TAV pode variar. Um período de aclimação durante o teste é aconselhável, o que pode inviabilizar o uso de vídeos mais curtos. Contudo, caso haja a escolha por vídeos mais curtos, como é o caso da nossa pesquisa, os autores apontam que um número maior de vídeos (por condição experimental) é necessário para tornar os resultados mais robustos.

De forma a aproximar o nosso estudo a uma tradição de *design* experimental da psicolinguística (testes com estímulos curtos, como sentenças, e mais de um estímulo por condição experimental testada), aliada ao postulado supracitado de Orero *et al.* (2018), decidimos utilizar vídeos curtos, de 10 a 15 segundos, mas que fossem autossuficientes em termos de compreensão<sup>36</sup>. Dessa forma, escolhemos 8 cenas que representassem as 4 condições

---

<sup>36</sup> Orero *et al.* (2018) apontam que estes devem ser os mais autênticos possíveis de modo a garantir determinada validade ecológica para o experimento. Caso sejam fragmentos, como é o nosso caso, devem ser

experimentais da pesquisa, com cada condição experimental se repetindo duas vezes. Abaixo, uma pequena descrição de cada uma das cenas.

### 3.3.1 Descrição imagética dos vídeos legendados

#### 3.3.1.1 Vídeo 1 – “Bebendo em casa”

**Figura 5 – Exemplo do Vídeo 1 com legenda de 1 linha**



Fonte: Elaborada pelo autor.

O Vídeo 1 é um dos vídeos do experimento que possuem uma maior correlação e, dessa forma, redundância, entre a imagem e a legenda, o que pode ser visto pelo título que demos a ele (“Bebendo em casa”) (Figura 5). Mark Zuckerberg, em destaque no centro do vídeo, está em casa após ter terminado um relacionamento. Na cena em questão, ele está escrevendo em um *blog* momentos antes de criar uma rede social que seria precursora do *Facebook* ao mesmo tempo que “está bêbado e não são nem 22h de uma terça”. Esse trecho do filme foi escolhido por, apesar de conter a voz do personagem em *off-screen* (fora de tela), ser visível que o personagem detém o turno de fala, com ele ao centro da imagem e a identificação do falante na LSE. Essa questão facilitou na hora de manipular as legendas de acordo com a velocidade da condição experimental, visto que não havia a necessidade de sincronismo entre lábios e legenda. Além disso, os poucos cortes de cena e o fato de que não ocorre muita ação

---

autossuficientes (*self-contained*). Além disso, caso haja uma comparação de vários fragmentos, eles devem ser parecidos em termos de complexidade, velocidade da fala, gênero etc., para que se evitem variáveis de confusão. Essas questões pautaram a escolha das cenas da presente pesquisa; contudo, mesmo tentando evitar que alguma interferência da escolha de vídeo e da composição linguística das legendas acontecesse, reconhecemos essa possibilidade, dada a natureza heterogênea de composições intersemióticas (de que imagem, som e legenda fazem parte) em produções audiovisuais, mesmo que dentro de um mesmo filme.

durante o excerto motivaram a escolha da cena. Mark continua sentado na maior parte da cena, com foco nele e sem cortes, ao mesmo tempo em que seu amigo adentra o ambiente ao fundo.

### 3.3.1.2 Vídeo 2 – “O julgamento”

**Figura 6 – Exemplo do Vídeo 2 com legenda de 1 linha**



Fonte: Elaborada pelo autor.

Apesar de menor que no Vídeo 1, a redundância entre imagem e fala (legenda) no Vídeo 2 também existe. Como também é visível a partir de como intitulamos o vídeo (“O julgamento”), a cena em questão relata um dos momentos do presente no qual o julgamento de Mark Zuckerberg movido por Eduardo Saverin acontece. A cena se inicia com um foco em Mark e em seus advogados por um tempo; logo após, corta para o *frame* exibido na Figura 6, em que a juíza (explicitada na legenda) lê as acusações contra Mark. Essa cena foi escolhida por cremos que se trata de um excerto autossuficiente, no qual vários elementos da cena (ambiente, identificação da falante como juíza etc.) auxiliam na compreensão desse trecho curto, de apenas 12 segundos. Além dos poucos cortes de cena e da pouca movimentação em tela, o fato de a juíza falar de forma pausada permitiu que atingíssemos um sincronismo considerável (não perfeito) e, dessa forma, conseguíssemos manipular as questões de velocidade da legenda (145 e 180 ppm), conforme será apresentado mais à frente.

### 3.3.1.3 Vídeo 3 – “O roubo”

**Figura 7 – Exemplo do Vídeo 3 com legenda de 2 linhas**



Fonte: Elaborada pelo autor.

O Vídeo 3, por sua vez, intitulado como “O roubo”, oferece pouca redundância entre imagem e fala. Nesta cena em questão, de cerca de 12 segundos, a câmera foca os irmãos Winklevoss, primeiramente, que praticam remo em uma piscina com ondas artificiais. Em seguida, há um corte para Navyia, amigo e sócio deles, que os informa que Mark Zuckerberg roubou o seu *website* (Figura 7). A câmera foca os gêmeos ao fundo, que param de treinar estupefatos. Ao focar novamente Navyia, ele repete a mensagem que havia dito antes: que Mark Zuckerberg roubou o seu *website*. Essa cena também foi escolhida pela autossuficiência de sua informação (ORERO *et al.*, 2018) e pelos poucos cortes de cena. Ademais, o fato de Navyia, que detém o turno de fala, expressar-se pausadamente e não haver cortes de cena durante sua fala facilitou a manipulação desse estímulo de acordo com as variáveis estudadas e com o isolamento de variáveis que pudessem influenciar o processamento dos participantes, como o corte de cenas.

### 3.3.1.4 Vídeo 4 – “O encontro”

**Figura 8 – Exemplo do Vídeo 4 com legenda de 2 linhas**

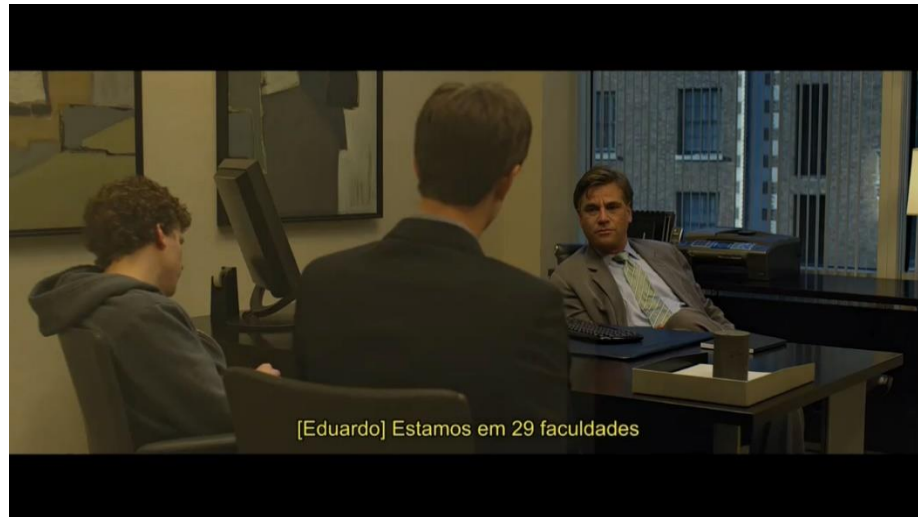


Fonte: Elaborada pelo autor.

O Vídeo 4 também estabelece pouca redundância entre legenda e o canal imagético. Na cena, Eduardo e Mark estão em foco a maior parte do tempo (Figura 8), com alguns cortes que enquadram ora o rosto de um, ora o rosto do outro. O conteúdo informacional da fala de Eduardo ao longo dos 15 segundos de cena é baixo, visto que ele apenas comunica a Mark que duas garotas – Christie e Alice – querem sair com eles essa noite. No final da cena, há um corte para um momento – provavelmente posterior à saída – em que há o enquadramento das pernas de duas pessoas entrando em um banheiro juntas, de frente para a outra, em que é possível inferir que estavam se beijando. Apesar de essa parte do excerto não ter legenda, o corte do vídeo permitindo que esse momento fosse visualizado pelos participantes deveu-se ao fato de gerar um pouco mais de contextualização à cena – visto a pouca redundância entre legenda e imagem quando Eduardo detém o turno de fala, tornando assim o trecho mais autossuficiente. A pouca movimentação em tela quando a legenda está presente (os atores estão parados e sentados), além da fala pausada de Eduardo, motivaram a escolha desse clipe.

### 3.3.1.5 Vídeo 5 – “A negociação”

**Figura 9 – Exemplo do Vídeo 5 com legenda de 1 linha**



Fonte: Elaborada pelo autor.

O Vídeo 5 estabelece pouca relação entre imagem e texto, sendo pouco redundante, dessa forma (Figura 9). No estímulo utilizado, de cerca de 13 segundos, há uma primeira cena em que Mark e Eduardo estão em um ambiente formal, como em um escritório, entreolhando-se. Após isso, há um corte de cena para o momento do *frame* em questão, em que Eduardo, de costas, está oferecendo o *Facebook* para investimento de determinado empresário, por isso intitulamos o vídeo como “A negociação”. Em sua fala, há um pouco mais de densidade lexical do que as já apresentadas aqui, visto que ele faz uso de alguns dados sobre o movimento de expansão contínua da rede social, citando que estão em “mais de 29 faculdades com mais de 75 mil usuários”. Essa cena foi escolhida pela pouca movimentação de cena e por apenas deter um turno de fala no excerto. Mesmo estando de costas, fica claro que o personagem em questão é quem está falando, porque seu corpo se movimenta, e os outros permanecem estáticos.

### 3.3.1.6 Vídeo 6 – “A festa”

**Figura 10 – Exemplo do Vídeo 6 com legenda de 2 linhas**



Fonte: Elaborada pelo autor.

Assim como o Vídeo 5, o Vídeo 7 também oferece pouca relação entre imagem e legenda. No vídeo, de cerca de 10 segundos, o personagem Sean Parker, interpretado por Justin Timberlake e que está no centro do *frame* da Figura 10, fala aos empregados do *Facebook* sobre uma festa (daí o título do vídeo) e sobre a necessidade de que todos se preparem para tal acontecimento. O vídeo se desenrola dentro desse enquadramento durante a maior parte do tempo. No final, quando não há mais fala, há um corte para um dos funcionários, que assente com a cabeça, faz sinal de positivo e continua a trabalhar no seu computador. Esse vídeo foi escolhido por estar claro quem detém o turno de fala na cena, tendo sido inclusive indicado pelos consultores presentes no nosso estudo piloto, assim como pouca movimentação de câmera. A legenda consiste em uma fala simples, de pouco conteúdo informacional, mas que ainda é autossuficiente (uma pessoa convidando outras para a festa).



### 3.3.1.7 Vídeo 7 – “Sobre Christy...”

**Figura 11 – Exemplo do Vídeo 7 com legenda de 2 linhas**



Fonte: Elaborada pelo autor.

O Vídeo 7 também estabelece pouca redundância entre a informação imagética e informação verbal; a moça sobre quem ele fala no vídeo (Christy), inclusive, nem aparece na cena. O segmento em questão, o menor do nosso experimento, de cerca de 8 segundos, começa com o foco em Eduardo, no centro da tela, como se vê na Figura 11. Ele fala sobre Christy, sua namorada, e sobre os ciúmes dele. Todo esse contexto o deixa com medo. Essa talvez seja a única complementariedade oferecida pela imagem ao que é dito em sua fala e, conseqüentemente, na legenda: ele está em um ambiente escuro e o seu semblante expressa determinada preocupação. O único corte da cena acontece no fim, quando a câmera foca o personagem que estava o tempo todo à direita do *frame* e com quem Eduardo falava, revelando se tratar de Mark Zuckerberg. Neste momento, já não há mais fala. Esse vídeo foi escolhido pela pouca movimentação de cena e pelo conteúdo informacional relativo da legenda; além disso, não há dúvidas sobre quem detém o turno de fala no vídeo em questão.

### 3.3.1.8 Vídeo 8 – “A nova sede”

**Figura 12 – Exemplo do Vídeo 8 com legenda de 2 linhas**



Fonte: Elaborada pelo autor.

O Vídeo 8, por sua vez, também oferece redundância mediana entre imagem e legenda. Ao longo do excerto, o personagem Eduardo entra na nova sede (como expresso pelo título do vídeo) do *Facebook*. Durante os 15 segundos do trecho, o personagem é acompanhado por uma câmera em movimento de *travelling*, um movimento lateral à esquerda sem cortes (Figura 12). O único corte da cena é ao final, quando, após o fim da legenda, há o enquadramento do rosto de Eduardo, contemplando a nova sede em questão. A fala do personagem é *off-screen*, em *voice over*. Contudo, está claro que é ele que está falando sobre tal segmento pela centralidade da cena e, principalmente, o fato de ele utilizar a primeira pessoa do indicativo na fala através dos verbos “fui” e “entrei”. Essa cena foi escolhida pela falta de cortes; mas o movimento lateral da câmera, mesmo que suave, fez-nos pensar, posteriormente, se causaria algum impacto na movimentação ocular dos participantes, algo que não nos ocorreu durante a manipulação e as escolhas dos estímulos para o experimento. Além disso, o fato de haver a logo do *Facebook* em uma espécie de geladeira à direita da imagem no começo do vídeo causaria uma maior contextualização para a cena e, dessa forma, ampliaria as possibilidades de compreensão.

Dessa forma, cada um dos 8 vídeos foi legendado de acordo com as 4 condições experimentais da pesquisa: lenta com 1 linha, lenta com 2 linhas, rápida com 1 linha, rápida com 2 linhas, gerando assim 32 estímulos distintos (vídeos legendados).

Ainda sobre os materiais da pesquisa, Orero *et al.* (2018) postulam que em pesquisas (quase)experimentais<sup>37</sup> em legendagem há a necessidade de que as características da legenda testada sejam reportadas (velocidade da legenda, por exemplo, assim como esta foi calculada), além de características da fala original do filme e da legenda (número de palavras, caracteres, proporção *type/token*, tamanho médio da sentença etc.), de modo que a complexidade linguística dos diálogos seja apontada. Além disso, é importante que se descreva a natureza da informação visual e sonora, assim como essas informações interagem no vídeo utilizado como estímulo. As manipulações também devem ser reportadas, caso elas ocorram.

Acima, fizemos a descrição da composição imagética dos vídeos utilizados e como esta se relaciona com o conteúdo linguístico das cenas (expresso nas legendas). Agora, de forma a adotar um procedimento de descrição dos materiais conforme apregoado por Orero *et al.* (2018), continuamos com a descrição das características linguísticas das legendas. Neste momento, atemo-nos ao produto final (vídeo legendado) levado ao experimento. Mais adiante, na subseção de “Preparação e manipulação dos estímulos”, fornecemos mais detalhes acerca do processo de preparação e manipulação dos estímulos e das legendas.

### 3.3.2 Descrição linguística das legendas

Abaixo, apresentamos o conteúdo de cada um dos 8 vídeos utilizados no experimento com detalhes acerca de uma das condições experimentais (dentre as 4) em relação ao seu tempo de início e final (através da sigla TCR<sup>38</sup>) legendados em uma das condições experimentais.

Em relação ao vídeo 1, um exemplo abaixo (Quadro 2) mostra-o na condição experimental 1L145 (com 1 linha na velocidade de 145 ppm):

---

<sup>37</sup> Orero *et al.* (2018) denominam os experimentos em Tradução Audiovisual/Legendagem de “quase-experimentos”, visto que a definição tradicional de pesquisa experimental – experimentos puros – envolve o controle muito mais rigoroso de variáveis de confusão, por exemplo, do que na pesquisa em Tradução Audiovisual. Além disso, a presença de fatores cognitivos, linguísticos e socioculturais dos participantes, questões difíceis de serem controladas e comuns na pesquisa em Tradução Audiovisual, fazem com que os autores encaixem a pesquisa sobre processamento e recepção em Tradução Audiovisual como “quase-experimental”. Por não necessitar de tanto rigor quanto “experimentos puros”, quase-experimentos podem (no sentido de “possibilidade”, não de “poder”) apresentar uma validade interna da pesquisa mais frágil, fazendo com que generalização e replicabilidade sejam mais limitadas (ORERO *et al.*, 2018).

<sup>38</sup> *Time code reader*, nome em inglês que designa a marcação do tempo em hora, minuto, segundo e quadro de uma produção audiovisual.

**Quadro 2 – Legenda do Vídeo 1 manipulado na condição 1L145**

TCR	Duração	Nº de caracteres	Velocidade	Legenda
00:08:51,883 --> 00:08:53,883	2s	28c	145ppm (14cps)	<i> <sup>39</sup> [Mark] Estou um pouco bêbado</i>
00:08:53,983 --> 00:08:55,203	1,2s	17c	145ppm (14cps)	<i>não posso mentir.</i>
00:08:55,303 --> 00:08:56,953	1,6s	23c	145ppm (14cps)	<i>E ainda não são nem 22h</i>
00:08:57,053 --> 00:08:58,413	1,3s	19c	145ppm (14cps)	<i>de uma terça-feira.</i>

Fonte: Elaborado pelo autor.

Em relação ao Vídeo 2, o Quadro 3 mostra a legenda do vídeo em questão na condição experimental 1L180 (com 1 linha na velocidade de 180 ppm):

**Quadro 3 – Legenda do Vídeo 2 manipulado na condição 1L180**

TCR	Duração	Nº de caracteres	Velocidade	Legenda
00:19:04,018 --> 00:19:05,313	1,2s	22c	180ppm (17cps)	[Juíza] Sr. Zuckenberg
00:19:05,513 --> 00:19:06,868	1,3s	23c	180ppm (17cps)	you are being accused
00:19:07,168 --> 00:19:08,408	1,2s	21c	180ppm (17cps)	of violating security
00:19:08,708 --> 00:19:09,888	1,1s	20c	180ppm (17cps)	and your rights.

Fonte: Elaborado pelo autor.

Em relação ao Vídeo 3, o Quadro 4 abaixo mostra a legenda do vídeo em questão na condição experimental 2L145 (com 2 linhas na velocidade de 145 ppm).

**Quadro 4 – Legenda do Vídeo 3 manipulado na condição 2L145**

TCR	Duração	Nº de caracteres	Velocidade	Legenda
00:39:28,659 --> 00:39:32,659	4s	56c	145ppm (14cps)	Only for you to know that Zuckerberg stole our site.
00:39:34,833 --> 00:39:38,263	3,4s	48c	145ppm (14cps)	Mark Zuckerberg... because he stole our site.

Fonte: Elaborado pelo autor.

<sup>39</sup> Os símbolos <i> e </i> colocados no início e o final de uma inserção de legenda em um programa de legenda são utilizados para transformar uma sentença em itálico. No caso dos parâmetros utilizados pela presente pesquisa, o exemplo em questão do vídeo 1 está em itálico porque o personagem responsável pela fala vem de uma fonte fora de tela (*off-screen*).

A legenda do Vídeo 4, por sua vez, está abaixo no Quadro 5 na condição experimental 2L180 (com 2 linhas na velocidade de 180 ppm).

**Quadro 5 – Legenda do Vídeo 4 manipulado na condição 2L180**

TCR	Duração	Nº de caracteres	Velocidade	Legenda
00:50:45,922 --> 00:50:48,280	2,7s	40c	180ppm (17cps)	Os nomes das garotas são Christy e Alice.
00:50:50,160 --> 00:50:52,925	3,6s	47c	180ppm (17cps)	E querem sair para beber com a gente esta noite.

Fonte: Elaborado pelo autor.

Já o Vídeo 5 tem sua legenda presente no Quadro 6 abaixo na condição experimental 1L145 (com 1 linha na velocidade de 145 ppm).

**Quadro 6 – Legenda do Vídeo 5 manipulado na condição 1L145**

TCR	Duração	Nº de caracteres	Velocidade	Legenda
01:04:47,682 --> 01:04:50,112	2,4s	34c	145ppm (14cps)	[Eduardo] Estamos em 29 faculdades
01:04:50,212 --> 01:04:52,212	2s	28c	145ppm (14cps)	com mais de 75 mil usuários.
01:04:52,312 --> 01:04:54,322	2s	28c	145ppm (14cps)	As pessoas passam mais tempo
01:04:54,422 --> 01:04:56,922	2,5	35c	145ppm (14cps)	no Facebook do que em outros sites.

Fonte: Elaborado pelo autor.

A legenda do Vídeo 6 (presente no Quadro 7) está logo abaixo na condição experimental 1L180 (com 1 linha na velocidade de 180 ppm).

**Quadro 7 – Legenda do Vídeo 6 manipulado na condição 1L180**

TCR	Duração	Nº de caracteres	Velocidade	Legenda
01:45:25,062 --> 01:45:26,122	1s	18c	180ppm (17cps)	Eu quero ver todos
01:45:26,322 --> 01:45:27,737	1,4s	24c	180ppm (17cps)	preparados para a festa.
01:45:28,037 --> 01:45:29,567	1,5s	26c	180ppm (17cps)	Vamos entrar naquele clube
01:45:29,867 --> 01:45:31,342	1s	25c	180ppm (17cps)	como se fosse um desfile.

Fonte: Elaborado pelo autor.

Já o Vídeo 7 (presente no Quadro 8) tem sua legenda abaixo na condição experimental 2L145 (com 2 linhas na velocidade de 145 ppm).

**Quadro 8 – Legenda do Vídeo 7 manipulado na condição 2L145ppm**

TCR	Duração	Nº de caracteres	Velocidade	Legenda
01:32:15,826 --> 01:32:18,686	2,8s	40c	145ppm (14cps)	A Christy é realmente uma moça psicótica.
01:32:18,786 --> 01:32:21,866	3s	43c	145ppm (14cps)	Ela é ciumenta, irracional e estou com medo.

Fonte: Elaborado pelo autor.

Por último, a legenda do Vídeo 8 (presente no Quadro 9) está abaixo na condição experimental 2L180 (com 2 linhas na velocidade de 180 ppm).

**Quadro 9 – Legenda do Vídeo 8 manipulado na condição 2L180**

TCR	Duração	Nº de caracteres	Velocidade	Legenda
01:41:34,396 --> 01:41:37,456	3s	55c	180ppm (17cps)	<i>[Eduardo] Fui até a Califórnia e entrei na nova sede.</i>
01:41:37,856 --> 01:41:40,976	3,1s	56c	180ppm (17cps)	<i>Eu não sabia se ia para uma festa ou para uma reunião.</i>

Fonte: Elaborado pelo autor.

Abaixo, no Quadro 10, há mais detalhes sobre as características linguísticas das legendas utilizadas no experimento. No quadro em questão, há detalhes acerca do número de palavras, número de caracteres de todas as palavras somadas (sem contar sinais de pontuação), o tamanho médio das palavras (obtido por meio da divisão do número de caracteres total pelo número de palavras) e a razão *type/token*, que é uma medida que considera a divisão do número de palavras distintas (*types*, que são ocorrências únicas de uma palavra no texto) pelo número de palavras corridas (*tokens*, que são todas as ocorrências lexicais no texto). Essa medida, especificamente, pode oferecer um *insight* acerca da densidade lexical do texto das legendas. Essa razão, quanto mais próxima ao número 1, representa uma densidade lexical maior do estímulo.

**Quadro 10 – Características linguísticas das legendas**

Vídeo	Legenda	Nº palavras	Nº de caracteres das palavras	Tamanho médio da palavra (sem os sinais) em caracteres	Razão <i>type/token</i>
1	[Mark] Estou um pouco bêbado, não posso mentir. E ainda não são nem 22h de uma terça-feira.	17	70	4,1c	0.94
2	[Juíza] Sr. Zuckenberg, você está sendo acusado de violar a segurança e direitos autorais.	14	73	5,2c	1
3	Só para vocês saberem que Zuckerberg roubou o nosso site. Mark Zuckerberg... pois é, ele roubou nosso site.	18	84	6c	0.77
4	Os nomes das garotas são Christy e Alice. E querem sair para beber com a gente esta noite.	18	71	3,9c	0.94
5	[Eduardo] Estamos em 29 faculdades com mais de 75 mil usuários. As pessoas passam mais tempo no Facebook do que em outros sites.	23	102	4,4c	0.91
6	Eu quero ver todos preparados para a festa. Vamos entrar naquele clube como se fosse um desfile.	17	78	4,5c	1
7	A Christy é realmente uma moça psicótica. Ela é ciumenta, irracional e estou com medo.	15	69	4,6c	0.93
8	[Eduardo] Fui até a Califórnia e entrei na nova sede. Eu não sabia se ia para uma festa ou para uma reunião.	22	83	3,6c	0.95

Fonte: Elaborado pelo autor.

Como é possível observar no Quadro 10, acima, o número de palavras das legendas dos estímulos variou entre 14 (Vídeo 2) e 23 palavras (Vídeo 5). Obtivemos, assim, uma média de 18 palavras por legenda entre os 8 vídeos a que cada participante foi exposto. Em relação ao número de caracteres, este variou entre 69 (Vídeo 7) e 102 (Vídeo 5) caracteres por legenda, com uma média de 78,7 caracteres das palavras da legenda por estímulo. No que diz respeito ao número de caracteres das palavras, este variou entre 3,6 (Vídeo 8) e 6 (Vídeo 3), apresentando uma média de 4,53 caracteres por palavra. Por fim, a razão *type/token*, utilizada com o propósito de observar a densidade lexical dos estímulos utilizados no experimento, apresentou uma variação entre 0,77 (Vídeo 3) e 1 (Vídeos 2 e 6).

Algumas considerações se fazem pertinentes a partir da observação do Quadro 10. Em relação ao número de palavras das legendas, há uma certa diferença entre os Vídeos 2 (14) e os Vídeos 5 (23) e 8 (22). Quando os Vídeos 2, 5 e 8 são comparados em relação ao tamanho

médio das palavras em caracteres, contudo, o Vídeo 2 apresenta palavras mais longas (5,2c) do que os vídeos 5 (4,4c) e 8 (3,6c). Essa diferença, no entanto, não seria significativa para o processamento das legendas dado o fato de que lemos por palavras, não por caracteres, sendo que a previsibilidade da leitura parafoveal faria essa diferença entre caracteres dos vídeos em questão pouco significativa. Além disso, o número maior de caracteres do Vídeo 2 é impulsionado pela presença do sobrenome do protagonista do filme – Zuckerberg – que contém 10 caracteres. Além do Vídeo 2 em questão, um caso semelhante ocorreu na legenda do Vídeo 3.

Ainda em relação à questão número de palavras  $\times$  tamanho médio das palavras em caracteres nos vídeos mencionados acima, é interessante ressaltar o tamanho pequeno das palavras no vídeo 8 (3,6c), apesar de ser denso lexicalmente (0,95). No caso desse vídeo, o alto número de palavras de função, como conjunções, preposições e artigos, normalmente menores em tamanho, faz com que o tamanho médio das palavras seja pequeno. Retomando a questão de tamanho médio das palavras em caracteres, ao constataremos o tamanho médio de palavras do *corpus* (4,53), observamos que ficou bem próximo ao tamanho médio da palavra em língua portuguesa, que é de 4,64, conforme Quaresma e Pinho (2007).

No que concerne essa densidade lexical representada pela razão *type/token* expressa no Quadro 10, a maior parte dos estímulos ficou entre 0,91 e 1, o que representa uma alta densidade lexical. O Vídeo 2, que apresentou a menor densidade lexical dentre os estímulos, tem uma legenda que apresenta uma alta coesão lexical por meio da estratégia de repetição, conforme é apregoado por Ghia (2012), em que as palavras “Zuckerberg”, “roubou”, “nosso” e “*site*” aparecem duas vezes cada em um trecho curto de 18 palavras. Apesar de a alta densidade lexical da maior parte dos vídeos diminuir a repetição de determinados itens lexicais, o fato de os trechos serem curtos e, em sua maioria, contendo mensagens simples nos faz crer que não foi um impedimento para um processamento mais fluido dos nossos participantes. Ademais, a coesão intersemiótica (DÍAZ CINTAS; REMAEL, 2007) propiciada pela interação da imagem e da legenda dos vídeos do experimento, com graus maiores ou menores de redundância entre os dois canais de informação, deve reforçar a compreensão do conteúdo e processamento dos excertos.

Dois vídeos que, a partir de suas características linguísticas e relação entre imagem e texto da legenda, devem ser inspecionados com mais atenção durante a análise dos dados são os Vídeos 5 e 6, respectivamente. Isto pelo fato de se tratar de vídeos com baixa redundância entre imagem e texto e alta densidade lexical (0,91 e 1, respectivamente), o que supomos que talvez pudesse apresentar alguma interferência no seu processamento. Contudo, o fato de terem



sido escolhidos mesmo assim deve-se ao fato de ser quase impossível o isolamento de todas as variáveis que compõem um produto audiovisual legendado e que talvez influenciem o processamento dos espectadores.

Todos os vídeos que estavam na condição experimental 2L continham 2 inserções de legenda, enquanto os vídeos que estavam na condição experimental 1L continham 4 inserções de legenda. Dessa forma, todos os participantes foram expostos a 16 inserções de 1 linha (4 vídeos 1L) e 8 inserções de 2 linhas (4 vídeos 2L), totalizando 24 legendas expostas a cada participante.

Na próxima seção, fornecemos mais detalhes acerca da manipulação das legendas, além de outros procedimentos metodológicos da pesquisa.

### **3.4 Procedimentos metodológicos**

Após a apresentação dos vídeos e das legendas que compõem os objetos do nosso experimento, nesta seção apresentamos as seguintes etapas da pesquisa: Preparação e manipulação dos estímulos; Seleção dos participantes; Desenho experimental; Procedimento do experimento; Procedimentos éticos; e Análise dos dados.

#### **3.4.1 Preparação e manipulação dos estímulos**

A preparação dos vídeos e legendas utilizados como estímulo nos nossos experimentos foi a parte mais problemática da pesquisa. Inicialmente, utilizaríamos quatro filmes distintos de modo a representar as 4 condições experimentais da pesquisa. Esses filmes seriam: *Exôdo: deuses e reis*, *A rede social*, *Madagascar* e *Plano de voo*, os quais foram analisados por Chaves (2017) e haviam apresentado um alto número de legendas de 1 linha em suas legendas para ouvintes. Contudo, a ideia (acertada) de que gêneros distintos pudessem interferir nos resultados a partir das características textuais e visuais particulares a cada um fez-nos escolher apenas um deles: *A rede social* (2010), conforme já mencionado anteriormente, sendo ele o filme cuja legenda profissional para ouvintes apresentou a maior porcentagem de legendas de 1 linha segundo os resultados de Chaves (2017). Esse filme, por se tratar de um drama biográfico, apresentaria *a priori* uma interferência menor das imagens no processamento das legendas, dado o fato de ser um filme com pouca ação em tela e muito diálogo. Dessa forma, passamos à escolha dos excertos e posterior manipulação das legendas que comporiam o filme.

A indefinição acerca da duração dos estímulos, assim como quais características imagéticas e narrativas os vídeos escolhidos teriam, permeou um tempo considerável da pesquisa. Isso ocorreu devido à tentativa de isolar as variáveis que, até então, a literatura da área de Legendagem, como um todo, e, mais especificamente, a sua interface com estudos com rastreamento ocular haviam categorizado como sendo influenciadores do processamento de legendas em menor ou maior grau. Dentre esses parâmetros, a segmentação linguística, o corte de cena (segmentação visual), a redução linguística, assim como os parâmetros técnicos (sinais gráficos, tamanho e cor da fonte etc.), linguísticos (coerência e coesão, por exemplo) e paralinguísticos (identificação dos falantes e tradução de música e ruídos), comuns à prática da legendagem, deveriam ser seguidos durante a manipulação.

Após a definição de que os vídeos seriam curtos, entre 8 e 15 segundos, autossuficientes (sem a necessidade de um contexto maior para que fossem entendidos), que os estímulos de 2 linhas e de 1 linha teriam 2 e 4 inserções de legenda, respectivamente, e que as velocidades de 145 ppm teria 14 caracteres por segundo e a de 180 ppm teria 17 caracteres por segundo, selecionamos cerca de 47 cenas do filme. Todos esses trechos foram legendados de acordo com os parâmetros preconizados na área no *software* gratuito de legendagem *Subtitle Workshop* 6.0b.

### 3.4.1.1 *Subtitle Workshop*

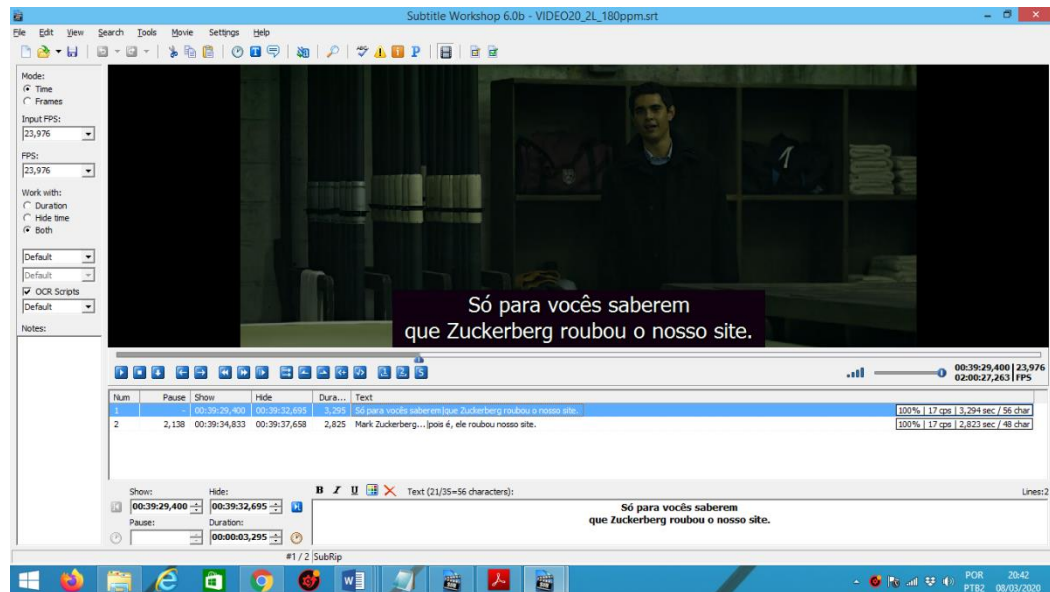
O *Subtitle Workshop* é o programa de legendagem gratuito mais conhecido que existe, tendo sido responsável em grande parte pela democratização e popularização do ato de legendar através de *fansubbers* (legendistas fãs) (ASSIS, 2016). Segundo o seu próprio manual, trata-se da “[...] ferramenta mais completa, eficiente e prática para a adição de legendas de vídeo<sup>40</sup>”. A prática foi adotada inicialmente por fãs que ansiavam por suas séries e *animes* (desenhos animados japoneses) antes de elas serem disponibilizadas oficialmente, o que demorava bastante tempo em décadas anteriores. Devido a sua praticidade, interface amigável e conjunto de ferramentas que facilitavam a vida do tradutor-legendista, hoje em dia também é adotado por alguns legendistas profissionais.

Abaixo, na Figura 13, podemos ver o *Subtitle Workshop* aberto durante a manipulação das legendas.

---

<sup>40</sup> Disponível em: <http://bemvin.org/manual-do-subtitle-workshop--portugus-br->. Acesso em 8 mar. 2020.

**Figura 13 – Interface do *Subtitle Workshop* durante manipulação das legendas**

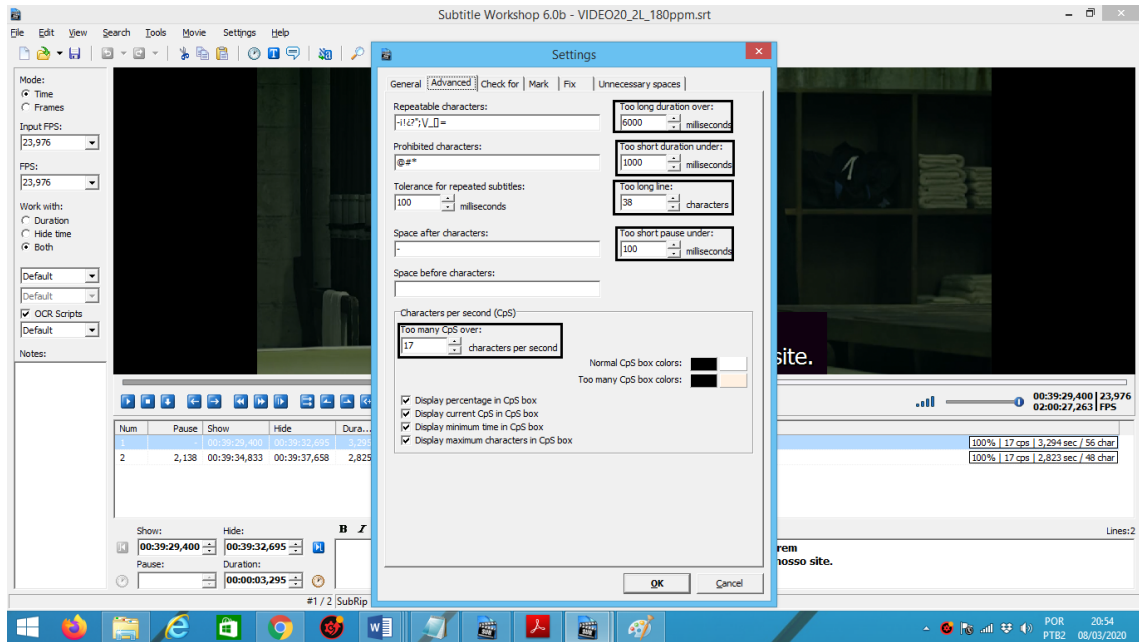


Fonte: Elaborada pelo autor.

Algumas ferramentas disponibilizadas pelo *Subtitle Workshop* em sua versão 6.0b facilitaram bastante a tarefa de legendagem empreendida no nosso trabalho. A partir da guia *TOOLS* (ferramentas) > *INFORMATIONS AND ERRORS* (informações e erros) > *SETTINGS* (configurações) > *ADVANCED* (avançado), conseguimos estabelecer alguns parâmetros técnicos que pautaram o nosso trabalho, como o número máximo de caracteres por linha (37), o tempo mínimo e máximo de exposição de legenda (1 segundo e 6 segundos<sup>41</sup>, respectivamente) e a pausa mínima entre legendas (100 milissegundos), evitando que a legenda apenas “piscasse” em tela (DÍAZ CINTAS; REMAEL, 2007). Na Figura 14, uma imagem da configuração em questão.

<sup>41</sup> Apesar de o tempo máximo de exposição de legenda de 4 segundos ser o tempo adotado pelo Grupo LEAD, ao qual estou vinculado, e ser comum no Brasil (ARAÚJO, 2008), tomei como tempo máximo a regra de 6 segundos proposta por d'Ydewalle *et al.* (1987). Essa decisão foi motivada pelo fato de que legendas com mais palavras poderiam ser adequadas na velocidade de 145 ppm sem necessitar de redução linguística, como era nossa intenção, para que uma mesma legenda manipulada em suas 4 condições experimentais conservasse o mesmo texto, sem alterações lexicais que pudessem influenciar o processamento.

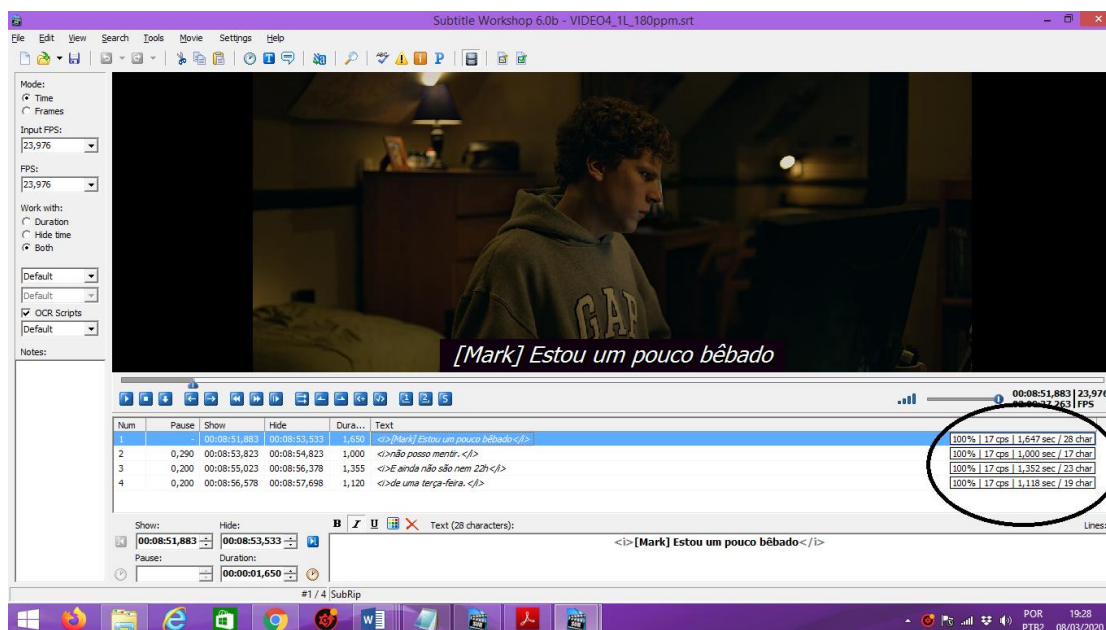
Figura 14 – Interface do *Subtitle Workshop* na guia SETTINGS



Fonte: Elaborada pelo autor.

A partir do momento em que esses parâmetros técnicos são delimitados, o programa avisa ao legendista, seja através de marcações coloridas ou “caixinhas” de informação, caso aqueles parâmetros sejam ultrapassados. Sobre essas “caixinhas”, podemos ver a Figura 15, abaixo, em que, à direita da imagem, há informações sobre o número de caracteres por segundo da legenda (no caso, 17 cps), além de informações a respeito da porcentagem para se chegar ao número de caracteres por segundo desejado (no caso da imagem, já atingido com 100%) e o tempo, a partir dos caracteres da legenda, necessário para que se alcançasse a velocidade de legenda de 17 cps.

**Figura 15 – Interface do *Subtitle Workshop* aberto durante manipulação das legendas**



Fonte: Elaborada pelo autor.

Dessa forma, como ainda havia dúvidas em relação a algumas características dos vídeos a serem utilizados, fomos legendando várias cenas nas quais seria possível manipular as 4 condições experimentais do estudo (1 e 2 linhas e velocidade rápida e lenta) de modo que apenas os parâmetros em questão fossem modificados. Dessa forma, teríamos que achar cenas em que, além das questões de autossuficiência e pouca movimentação de cena mencionadas: 1) não houvesse sobreposição de falas entre personagens ou diálogo, de modo que a condição experimental de 1 linha fosse possível; nestes dois casos, sobreposição de fala e diálogo, seria quase impossível transformar as legendas de 2 linhas em 1 linha sem maiores prejuízos aos outros parâmetros da legendagem adotados na produção das legendas; 2) tivesse voz fora de tela (*off-screen*) ou que a velocidade de fala não fosse tão rápida, de modo que, ao se diminuir o tempo de exposição de uma legenda para que se adequasse à velocidade de 145 ppm, não houvesse um prejuízo considerável para a sincronia entre movimento dos lábios e legenda; e 3) não houvesse corte de cena durante a exposição das legendas.

Assim, legendamos cerca de 47 excertos para que, posteriormente, fossem selecionados 8 que se adequassem melhor ao que foi descrito no parágrafo acima. Além disso, outras características pautaram a manipulação das legendas: 1) para que um determinado bloco de fala fosse transformado em uma legenda, ela deveria ter pelo menos 38 caracteres, ou seja, ter um número máximo de caracteres por linha adotado no presente trabalho (37); dessa forma, seria justificada a divisão desse bloco de fala em duas legendas de 1 linha, no caso das condições experimentais que pressupunham a condição 1L, constituindo duas linhas de no mínimo 19

caracteres; 2) como o texto da legenda deveria se adequar a ambas as velocidades<sup>42</sup> (145 ppm e 180 ppm), uma legenda não poderia ter menos de 17 caracteres por linha, visto que, sendo o tempo mínimo de exposição adotado no trabalho de 1 segundo, ao se utilizar uma legenda com menos de 17 caracteres nas condições experimentais cuja velocidade era de 180 ppm (exatamente 17 caracteres por segundo), a legenda ficaria menos do que 1 segundo na tela; 3) as palavras utilizadas nas legendas deveriam conter um número mínimo de pronomes, dada a baixa referência propiciada pelos trechos utilizados serem curtos, a não ser que houvesse a referência dentro da própria legenda ou houvesse redundância com a imagem; e 4) palavras consideradas de frequência baixa foram evitadas.

Diferentemente de d'Ydewalle e De Bruycker (2007), os quais não isolaram o número de linhas nos estímulos do seu experimento, tivemos como objetivo fazer com que os estímulos não misturassem legendas com números de linhas distintos de modo a investigar se a compreensão dos vídeos legendados em questão seria influenciada pelos participantes no que diz respeito a esse parâmetro. Essa diretriz também dificultou a seleção de cenas para o experimento, dado que quase impossibilita o uso de diálogos.

Uma observação em relação aos sinais gráficos na manipulação das legendas foi a adoção de uma prática compartilhada por mim e outros pesquisadores do LEAD (ARAÚJO, 2008; ARAÚJO; ASSIS, 2014) sobre o uso de vírgulas. Foram retiradas as vírgulas que estavam no final de uma legenda e antecedendo outra inserção de legenda, visto que a passagem de um bloco de legenda para o outro já representa uma pausa, organização do pensamento.

Após a manipulação das legendas a partir do que foi dito ao longo desta seção, foram escolhidos 8 vídeos com a ajuda de um consultor surdo e um ouvinte, este intérprete de Libras e CODA<sup>43</sup>, que participaram de um piloto da pesquisa, o qual será descrito adiante.

Como cada vídeo de 1 linha continha 4 inserções de legenda e cada vídeo de 2 linhas continha 2 inserções de legenda, número este utilizado como um padrão a partir do tempo de 8 a 15 segundos de vídeo adotado, a manipulação das legendas e o *design* utilizado no experimento resultaram em 32 vídeos legendados contendo um total de 96 inserções de legenda, dentre os quais 64 legendas de 1 linha e 32 de 2 linhas e 48 de 145 ppm e 48 de 180 ppm. Os arquivos das legendas foram nomeados contendo informações acerca das condições

---

<sup>42</sup> O conteúdo linguístico da legenda dos vídeos não foi alterado para que se adequasse às condições experimentais do estudo. Ou seja, o texto da legenda dos vídeos foi o mesmo nas condições experimentais 1L145, 1L180, 2L145 e 2L180. Dessa forma, conseguimos adequar o número de caracteres da legenda dos vídeos às diferentes velocidades manipulando o tempo de entrada e de saída da legenda e levando em consideração que a sincronia não fosse prejudicada.

<sup>43</sup> Filhos de pais surdos, do inglês *children of deaf adults*.

experimentais; um exemplo é a legenda do Vídeo 5 na condição experimental de 1 linha e de 145 ppm, que ficou “Vídeo5\_1L145ppm”.

Abaixo, no Quadro 11, um exemplo da legenda do Vídeo 5 nas 4 condições experimentais.

### Quadro 11 – Legenda do Vídeo 5 manipulada nas 4 condições experimentais

<p><b>1. Legenda de 1 linha e de 145 ppm (lenta)</b>  <b>Inserção 1</b> – [Eduardo] Estamos em 29 faculdades  <b>Inserção 2</b> – com mais de 75 mil usuários.  <b>Inserção 3</b> – As pessoas passam mais tempo  <b>Inserção 4</b> – no Facebook do que em outros sites.  <b>Total:</b> 125 caracteres; <b>tempo:</b> 8,9s</p>
<p><b>2. Legenda de 1 linha e de 180 ppm (rápida)</b>  <b>Inserção 1</b> – [Eduardo] Estamos em 29 faculdades – 34 caracteres  <b>Inserção 2</b> – com mais de 75 mil usuários. – 28 caracteres  <b>Inserção 3</b> – As pessoas passam mais tempo – 28 caracteres  <b>Inserção 4</b> – - no Facebook do que em outros sites – 35 caracteres  <b>Total:</b> 125 caracteres; <b>tempo:</b> 7,3s</p>
<p><b>3. Legenda de 2 linha e de 145 ppm (lenta)</b>  <b>Inserção 1</b> – [Eduardo] Estamos em 29 faculdades  com mais de 75 mil usuários. – 62 caracteres  <b>Inserção 2</b> – As pessoas passam mais tempo  no Facebook do que em outros sites. – 63 caracteres  <b>Total:</b> 125 caracteres; <b>tempo:</b> 8,9s</p>
<p><b>4. Legenda de 2 linhas e de 180 ppm (rápida)</b>  <b>Inserção 1</b> – [Eduardo] Estamos em 29 faculdades  com mais de 75 mil usuários. – 62 caracteres  <b>Inserção 2</b> – As pessoas passam mais tempo  no Facebook do que em outros sites. – 63 caracteres  <b>Total:</b> 125 caracteres; <b>tempo:</b> 7,3s</p>

Fonte: Elaborado pelo autor.

Como o *Subtitle Workshop* gera arquivos de legenda, mas não grava as legendas no vídeo, após a manipulação gravamos e editamos as legendas nos seus vídeos correspondentes com o auxílio do programa *AVIRecomp* 1.5.6. Nesse programa, utilizamos as legendas na extensão *SubStationAlpha* (.ssa), extensão que permite a edição de legendas. Nele, definimos

outras especificações técnicas, como a fonte Arial, a cor amarela, o tamanho 32 e o contorno preto. Como estávamos elaborando as legendas no arquivo de vídeo do filme inteiro – dada a indefinição acerca dos vídeos a serem escolhidos –, tivemos que cortar os vídeos utilizados depois com o programa de edição *Videopad Video Editor*.

### 3.4.2 Seleção dos participantes

Segundo Orero *et al.* (2018), considerar o poder estatístico e tamanho do efeito são questões essenciais para que comparações estatísticas sejam validadas. E isso mesmo em um contexto de modelos de efeitos mistos, que permitem o controle de diferenças individuais e, conseqüentemente, tornam possível surgir resultados válidos a partir de amostragens menores, um cálculo do número de participantes para que se alcance poder estatístico é necessário.

Nesse contexto, grupos de menos de 25 participantes não fornecem tanto poder estatístico. Isso ocorre porque, dada a variabilidade possível nesse número de participantes, o tamanho da amostragem deve ser ajustado para cima. Contudo, dada a perda comum de dados de participantes em pesquisas experimentais, por diversos fatores, coletar 20% a 30% a mais que o necessário, quantidade que é normalmente perdida nesse tipo de pesquisa, é aconselhável segundo Hennink *et al.* (2016 *apud* ORERO *et al.*, 2018). Entre os motivos para essa perda de dados, os autores citam: fadiga, perda de motivação/engajamento, dentre outros. Essas questões devem ser reportadas pelo pesquisador, assim como se deu o tratamento de *outliers*<sup>44</sup>.

Mesmo com esse número mínimo de participantes (cerca de 30), os autores reconhecem a fragilidade do recrutamento de pessoas em vulnerabilidade como pessoas surdas, o que pode diminuir consideravelmente a amostragem, algo que deve ser reportado na pesquisa. Entretanto, “[...] um número crítico de participantes com deficiência deve sempre ser alcançado em qualquer experimento em Tradução Audiovisual para ser considerado válido e confiável, e dessa forma publicável<sup>45</sup>” (ORERO *et al.*, 2018, p. 112).

Levando em consideração a dificuldade de se recrutar pessoas surdas para esse tipo de pesquisa, estipulamos inicialmente 48 participantes para os grupos que constituíram o experimento: 24 participantes surdos compondo o grupo experimental, e 24 participantes

---

<sup>44</sup> Em estatística, *outliers* são dados que fogem da normalidade e, por isso, podem causar alteração na análise estatística. Contudo, a noção do que é um *outlier* ou não no estudo da movimentação ocular em textos dinâmicos como a legendagem não é explícito na literatura da área. Dessa forma, descrevemos ao longo de 4.1 o processo que nos levou a determinar o ponto de corte do que seria um *outlier* ou não em todas as medidas analisadas.

<sup>45</sup> “[...] a critical number of participants with impairments should always be reached in any AVT research experiment to be considered valid and reliable, and therefore publishable”.



ouvintes compondo o grupo controle. Os ensurdecidos não foram incluídos na pesquisa por representarem outro padrão de processamento diferente dos surdos (KRUGER; SZARKOWSKA; KREJTZ, 2015). Contudo, devido ao contexto da pandemia de Covid-19 no ano de 2020, em que coletamos os dados, reduzimos o número de participantes para 40 (20 ouvintes e 20 surdos).

O grupo composto por surdos foi de usuários de Libras e leitores de textos em língua portuguesa. Como tínhamos a intenção de recrutar surdos graduandos ou graduados já acostumados com a leitura de textos em língua portuguesa e sendo foco do estudo no processamento de legendas, não na leitura de textos em si, não administramos teste de proficiência em língua portuguesa. O grupo de surdos, especificamente, foi composto de pessoas que tinham surdez pré-linguística e profunda.

Um questionário pré-coleta foi realizado com os participantes ouvintes (Apêndice A) e com os participantes surdos (Apêndice B) de modo que o seu perfil fosse traçado. Todos tinham visão normal ou corrigida, de forma que os resultados do seu comportamento ocular não fossem comprometidos.

#### **Quadro 12 – Dados demográficos dos participantes do grupo controle de ouvintes**

<b>Dados demográficos</b>	<b>Grupo dos ouvintes</b>
Número de participantes	20
Sexo	10 masc. / 10 fem.
Idade	29,1 (média)
Escolaridade	14,6 (média)

Fonte: Elaborado pelo autor.

Com os dados do grupo de ouvintes, podemos ver, no Quadro 12, o perfil dos participantes em relação a sexo, idade e escolaridade. A idade dos ouvintes variou entre 21 e 40 anos, com escolaridade dividida entre superior completo (13) e ensino médio (7), sendo 10 homens e 10 mulheres. Para o grau de escolaridade, definimos os valores em números: 12 anos para o ensino médio concluído e 16 anos para a graduação concluída. Dessa forma, a média do grupo de ouvintes para os anos no grau de escolaridade foi 14,6.

**Quadro 13 – Dados demográficos dos participantes do grupo controle de surdos**

Dados demográficos	Grupo dos Surdos
Número de participantes	20
Sexo	12 masc. / 8 fem.
Idade	33 (média)
Escolaridade	14,8 (média)

Fonte: Elaborado pelo autor.

Em relação ao grupo de surdos, o Quadro 13 traz as informações sobre o perfil dos participantes em relação a sexo, idade e escolaridade. A idade dos surdos variou entre 19 e 57 anos, com escolaridade dividida entre superior completo (14) e ensino médio (6), sendo 12 homens e 8 mulheres. A média do grupo de surdos para os anos no grau de escolaridade foi 14,8, bem próxima à média máxima possível de 16 e um pouco maior que a de ouvintes.

### 3.4.3 Desenho experimental

Retomando as perguntas de pesquisa que pautaram a investigação do custo de processamento de legendas por surdos e ouvintes, temos as seguintes:

- a) Até que ponto legendas de 1 linha ou 2 linhas influenciam o processamento de legendas por surdos e ouvintes?
- b) Até que ponto legendas rápidas (180 palavras por minuto) ou lentas (145 palavras por minuto) influenciam o processamento de surdos e ouvintes?

Para essas perguntas, baseadas na literatura da área, formulamos as seguintes hipóteses:

- a) Legendas de 1 linha induzem um processamento menos linear e, dessa forma, mais custoso do que legendas de 2 linhas;
- b) Legendas rápidas (180 palavras por minuto) e lentas (145 palavras por minuto) não diferem significativamente no processamento de surdos e ouvintes.

De acordo com as questões e hipóteses apresentadas acima, utilizamos quatro condições experimentais na pesquisa proposta: vídeos com i) legendas de 1 linha e de 145 ppm (1L145); ii) legendas de 1 linha e de 180 ppm (1L180); iii) legendas de 2 linhas e de 145 ppm (2L145); e iv) legendas de 2 linhas e de 180 ppm (2L180). Cada condição experimental foi vista duas vezes em dois estímulos distintos por cada participante, fazendo com que cada participante fosse exposto aos 8 vídeos legendados em condições experimentais distintas.

Esse tipo de desenho experimental constitui o quadrado latino (MARQUER, 2003), desenho este que já vem sendo utilizado pelas pesquisas no âmbito do Projeto ExLeg (MONTEIRO, 2016; VIEIRA, 2016) e também foi escolhido para esta tese. Esse desenho é propício aos objetivos da presente pesquisa, pois nele todos os participantes são expostos a todas as condições experimentais; da mesma forma, todas as condições são testadas por participantes diferentes, sendo, assim, possível observar o processamento dos vídeos legendados nas 4 condições experimentais pelos participantes dos dois grupos. Isso, em teoria, faz com que a influência de determinadas características linguísticas da legenda ou características do vídeo sejam sondadas, visto que não há uma ligação única entre um estímulo e uma condição experimental. O Quadro 14 abaixo traz a distribuição dos participantes, em pares, em todas as condições experimentais.

**Quadro 14 – Quadrado Latino: condições, vídeos e pares de participantes**

Condição	Vídeo	Par 1	Par 2	Par 3	Par 4
1L145	1	O1 S1 O17 S17	O5 S5	O9 S9	O13 S13
	2	O2 S2 O18 S18	O6 S6	O10 S10	O14 S14
	3	O3 S3 O19 S19	O7 S7	O11 S11	O15 S15
	4	O4 S4 O20 S20	O8 S8	O12 S12	O16 S16
	5	O5 S5	O9 S9	O13 S13	O1 S1 O17 S17
	6	O6 S6	O10 S10	O14 S14	O2 S2 O18 S18
	7	O7 S7	O11 S11	O15 S15	O3 S3 O19 S19
	8	O8 S8	O12 S12	O16 S16	O4 S4 O20 S20
1L180	1	O4 S4 O20 S20	O8 S8	O12 S12	O16 S16
	2	O5 S5	O9 S9	O13 S13	O1 S1 O17 S17
	3	O6 S6	O10 S10	O14 S14	O2 S2 O18 S18
	4	O7 S7	O11 S11	O15 S15	O3 S3 O19 S19
	5	O8 S8	O12 S12	O16 S16	O4 S4 O20 S20
	6	O9 S9	O13 S13	O1 S1 O17 S17	O5 S5
	7	O10 S10	O14 S14	O2 S2 O18 S18	O6 S6
	8	O11 S11	O15 S15	O3 S3 O19 S19	O7 S7
2L145	1	O3 S3 O19 S19	O7 S7	O11 S11	O15 S15
	2	O4 S4 O20 S20	O8 S8	O12 S12	O16 S16
	3	O5 S5	O9 S9	O13 S14	O1 S1 O17 S17
	4	O6 S6	O10 S10	O14 S14	O2 S2 O18 S18
	5	O7 S7	O11 S11	O15 S15	O3 S3 O19 S19
	6	O8 S8	O12 S12	O16 S16	O4 S4 O20 S20
	7	O9 S9	O13 S14	O1 S1 O17 S17	O5 S5
	8	O10 S10	O14 S14	O2 S2 O18 S18	O6 S6
2L180	1	O2 S2 O18 S18	O6 S6	O10 S10	O14 S14
	2	O3 S3 O19 S19	O7 S7	O11 S11	O15 S15
	3	O4 S4 O20 S20	O8 S8	O12 S12	O16 S16
	4	O5 S5	O9 S9	O13 S13	O1 S1 O17 S17
	5	O6 S6	O10 S10	O14 S14	O2 S2 O18 S18
	6	O7 S7	O11 S11	O15 S15	O3 S3 O19 S19
	7	O8 S8	O12 S12	O16 S16	O4 S4 O20 S20
	8	O9 S9	O13 S13	O1 S1 O17 S17	O5 S5

Fonte: Elaborado pelo autor.

No Quadro 14, acima, os participantes ouvintes são identificados pela sigla O e os surdos, pela sigla S, além dos números estipulados a partir de sua ordem de participação. De acordo com Vieira (2016), o desenho do quadrado latino é dividido em duas formas de estudo do comportamento ocular: i) pelas linhas horizontais, é possível comparar 4 pares de participantes em cada condição; e ii) pelas linhas verticais, é possível observar o desempenho dos 40 participantes da pesquisa nas quatro condições experimentais.

Inicialmente, esse *design* previa 48 participantes (24 ouvintes e 24 surdos). Contudo, o fato de termos coletado os dados de 40 participantes (20 ouvintes e 20 surdos), já justificado anteriormente, não prejudicou o *design* do experimento porque os 8 estímulos experimentais (vídeos) representavam cada condição experimental duas vezes. Dessa forma, o

*design* ainda ficou equilibrado em relação ao número de informações por participante (8 vídeos assistidos por cada) e por estímulo (cada vídeo assistido 40 vezes) e por cada uma das quatro condições experimentais (cada uma assistida 80 vezes).

Uma pesquisa-piloto foi realizada<sup>46</sup> para validar o desenho experimental proposto. Apesar de o *design* experimental ser bastante semelhante ao proposto por Vieira (2016) e Monteiro (2016), validar o uso de vídeos curtos, como proposto pela presente pesquisa, assim como os procedimentos do experimento, era necessário para que a pesquisa fosse colocada em prática. Dessa forma, realizamos uma pesquisa-piloto com dois participantes, sendo que um era surdo com bom nível de compreensão do português escrito, enquanto o outro era ouvinte CODA e intérprete de Libras. Como os participantes também eram consultores da pesquisa, o piloto também serviu para validar a escolha das próprias cenas e as escolhas linguísticas e estruturas sintáticas presentes nas legendas de forma que não soassem estranhas ao público surdo.

O uso de vídeos curtos foi validado no estudo-piloto. Foi possível observar o processamento dos dois participantes a partir da sua movimentação ocular, além da compreensão geral dos vídeos de acordo com o que foi obtido em resposta ao questionário pós-coleta e relato retrospectivo, principalmente por parte do surdo, mesmo que esses dois instrumentos não estejam presentes na escrita final deste trabalho. Além disso, o participante surdo, também consultor do projeto, sugeriu a exclusão de dois vídeos do experimento por falta de clareza em relação a quem detinha o turno de fala expresso na legenda. Como havíamos construído inicialmente mais de 40 estímulos, convidamos o consultor surdo a nos auxiliar com a escolha dos outros dois vídeos, o que foi feito de forma bem-sucedida.

Ademais, modificamos a execução da tarefa do experimento: ao invés de verem os 8 vídeos seguidos um do outro e, após isso, responderem ao questionário pós-coleta, o participante surdo sugeriu que a coleta fosse feita em dois blocos de 4 vídeos com uma pausa após os blocos para responder ao questionário e fazer o relato retrospectivo dos vídeos. A ideia de expor os 8 vídeos sem pausa foi motivada pela noção de que isso diminuiria o esforço e estresse dos participantes. A ideia inicial era que, caso fizéssemos a divisão de 1 vídeo para cada tarefa (assistir ao vídeo no rastreador ocular e fazer o relato livre e questionário técnico pós-coleta em seguida), seria necessário recalibrar<sup>47</sup> novamente o olhar dos participantes cada vez que ele voltasse ao rastreador ocular. Fazer isso 8 vezes, mesmo que utilizando vídeos

---

<sup>46</sup> Os dados provenientes do estudo-piloto foram importantes para a validação da metodologia aqui empregada, mas não foram considerados na escrita final desta pesquisa.

<sup>47</sup> O ato de o rastreador ocular reconhecer o olhar dos participantes para que assim consiga obter os dados relativos à fixação.

curtos, faria com que passássemos muito tempo para coletar cada participante, causando estresse e desconforto na realização da tarefa. Dessa forma, a opção de utilizar 4 vídeos divididos em dois blocos nos pareceu razoável a partir da sugestão do consultor.

Os participantes também sinalizaram uma construção sintática obscura em um dos vídeos (Vídeo 1), o que poderia dificultar a compreensão principalmente dos participantes surdos, que têm o português como LE. O texto de uma das legendas, que anteriormente estava como “E não são ainda nem 22h” foi modificado por “E ainda não são nem 22h”, uma ordem sintática mais direta.

Uma sugestão de ambos os participantes que não havíamos considerado anteriormente foi o uso de títulos para os vídeos. Assim, antes da exibição, os participantes poderiam contextualizar um pouco mais os vídeos e estabelecer hipóteses e associações sobre o conteúdo. Tentamos estabelecer nomes que se relacionassem ao conteúdo dos estímulos, os quais ficaram da seguinte forma: “Bebendo em casa” (Vídeo 1); “O julgamento” (Vídeo 2); “O roubo” (Vídeo 3); “O encontro” (Vídeo 4); “A negociação” (Vídeo 5); “A festa” (Vídeo 6); “Sobre Christy...” (Vídeo 7) e “A nova sede” (Vídeo 8).

O fato de termos dois níveis de grupo, um controle (ouvintes) e um experimental (surdos), assim como dois níveis para cada uma das outras variáveis independentes testadas (1 e 2 linhas, no caso de número de linhas, e 145 e 180 ppm, no caso de velocidade), faz com que o *design* fatorial desse estudo seja 2 x 2 x 2.

#### 3.4.4 Variáveis e medidas

Por um lado, as variáveis independentes da pesquisa são duas: a velocidade das legendas (145 ppm e 180 ppm) e o número de linhas de legenda (1 ou 2), sendo estes os aspectos a serem manipulados. Por outro lado, as variáveis dependentes (medidas geradas pelo rastreador ocular) são 9: 1) duração média das fixações; 2) número de fixações por legenda; 3) duração média de regressão; 4) ocorrência de regressão; 5) duração média de releitura; 6) ocorrência de releitura; 7) duração média de encapsulamento; 8) ocorrência de deflexão; e 9) duração média de latência.

1) duração média das fixações: essa medida diz respeito ao tempo médio das fixações de leitura na área da legenda. Um tempo de duração média alto está relacionado a um custo de processamento.

2) número de fixações por legenda: essa medida está relacionada ao número de fixações realizadas pelos participantes na área da legenda e da imagem (quando a legenda está

em exibição). Um baixo número de fixações pode estar relacionado a um custo de processamento no caso de um texto dinâmico como a legendagem. Isso, contudo, deve ser qualificado mais à frente na discussão dos dados junto a outras variáveis como releitura e regressão.

3) duração média de regressão e 4) ocorrência de regressão: medidas que se relacionam aos movimentos de fixação da direita para a esquerda realizados após o início da leitura das legendas e antes do término dessa leitura. Normalmente, as regressões estão relacionadas a uma dificuldade de processamento, como uma busca por uma informação que não foi processada de forma satisfatória.

5) duração média de releitura e 6) ocorrência de releitura: medidas relacionadas às fixações que ocorrem após a finalização da leitura das legendas. Podem estar relacionadas a um custo de processamento. Contudo, esses dados precisam ser qualificados, pois podem ser motivados pelo fato de a legenda ter ficado mais tempo do que o necessário em tela e, dessa forma, o participante pode ter finalizado a leitura da legenda, ido à imagem e voltado à área da legenda por ela ainda estar em exibição.

7) duração média de encapsulamento: medida relacionada ao tempo médio das últimas fixações em uma inserção de legenda. Estão relacionadas a um custo de processamento que envolve busca por referentes que não foram designados e por relações entre as orações, sendo uma tentativa de resolver algo que não ficou claro dentro da sentença. Um tempo alto de encapsulamento significa um custo de processamento.

8) ocorrência de deflexão: essa medida diz respeito às idas à área da imagem após a primeira fixação na legenda durante o processamento da legenda<sup>48</sup>. Uma grande quantidade de deflexões inicialmente pode estar relacionada a um custo de processamento.

9) duração média de latência: essa medida está relacionada ao intervalo entre a última fixação realizada na área da imagem e a primeira fixação realizada na área da legenda assim que ela aparece. Um tempo alto de latência pode significar um custo de processamento, visto que há menos tempo disponível para a leitura das legendas.

Essas medidas foram escolhidas a partir da leitura de trabalhos sobre legendagem e rastreamento ocular presentes na seção de Fundamentação Teórica que, ao nosso ver, apresentaram dados mais significativos para o estudo do processamento de legendas e, mais especificamente, o estudo da influência do número de linhas e velocidade no processamento.

---

<sup>48</sup> Como já colocado anteriormente, trata-se de uma medida controversa. Inicialmente, adotamos a visão de que se trata de uma medida que denota uma ruptura no processamento leitor, como é a opinião da maior parte da literatura da área, apesar de ser uma posição distinta da adotada por Vieira (2016).

### 3.4.5 Instrumentos

O instrumento para obtenção de dados da pesquisa foi o rastreador ocular, responsável pela geração de dados sobre as fixações e sobre o comportamento ocular no geral dos participantes.

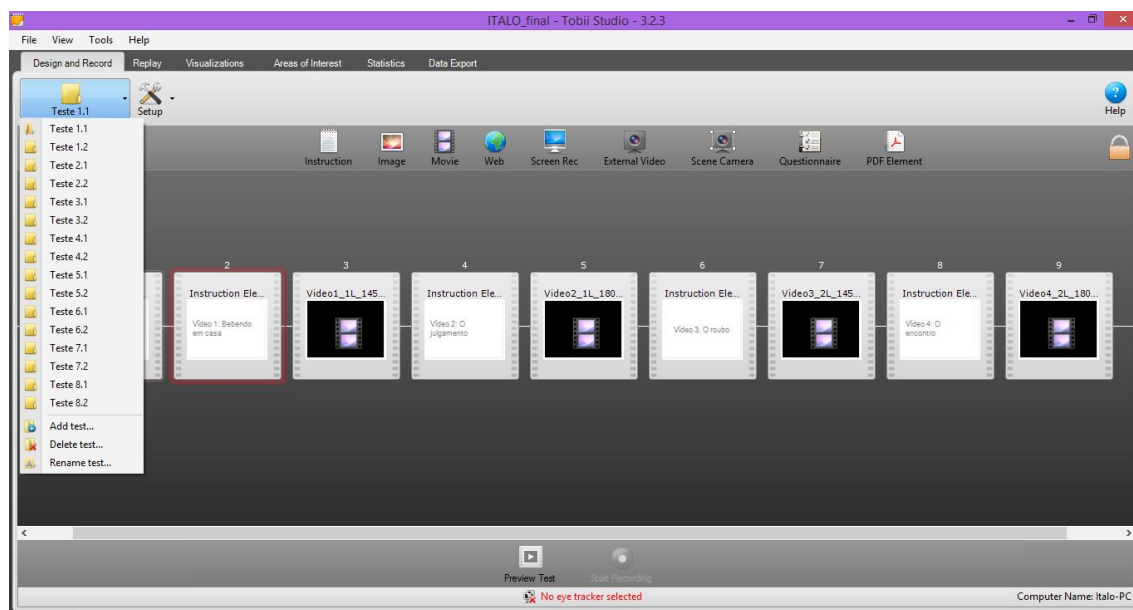
#### 3.4.5.1 O rastreador ocular

O rastreador utilizado na pesquisa foi o *Tobii T120 Eye Tracker (hardware)*, localizado no Laboratório de Ciências Cognitivas e Psicolinguística no Centro de Humanidades-I da Universidade Federal do Ceará. Trata-se de um monitor de 17 polegadas que trabalha com uma resolução máxima de  $1280 \times 1024$  pixels. Esse monitor conta com câmera embutida, que produz vídeos de resolução  $640 \times 480$  numa taxa de 30 *frames* (quadros) por segundo (30 fps), além de alto-falantes. A precisão do aparelho para o rastreamento ocular é de 120 hz, ou 120 milissegundos, ou seja, ele consegue gerar dados a respeito de 120 movimentações por segundo para cada olho a uma distância de 65 cm.

Os dados do *Tobii T120 Eye Tracker* foram analisados com o auxílio do *software Tobii Studio* em sua versão 3.2.3. Foi no *Tobii Studio* que montamos o experimento da pesquisa, distribuído em 8 testes que, seguindo a ordem apresentada no quadrado latino anteriormente, possibilita o revezamento dos participantes e das 4 condições experimentais entre os 8 vídeos. Como citado anteriormente, os 8 vídeos foram divididos em dois blocos de 4 vídeos para realização da tarefa. Dessa forma, os testes (fileira de tarefa) foram subdivididos em 1.1 e 1.2, por exemplo, como no caso do teste 1, e assim sucessivamente até os testes 8.1 e 8.2. Tudo isso foi feito na aba *Design and Record*, no canto superior esquerdo da Figura 16. Através da linha horizontal de itens, que pode ser vista na parte central superior da tela na figura abaixo, adicionamos os vídeos (*movie*), as instruções (*instruction*) do experimento e os títulos dos vídeos; estes também foram produzidos com o item *instruction* por questões de praticidade. Abaixo, na Figura 16, a interface do *Tobii Studio* aberta para a realização da Tarefa 1.1.



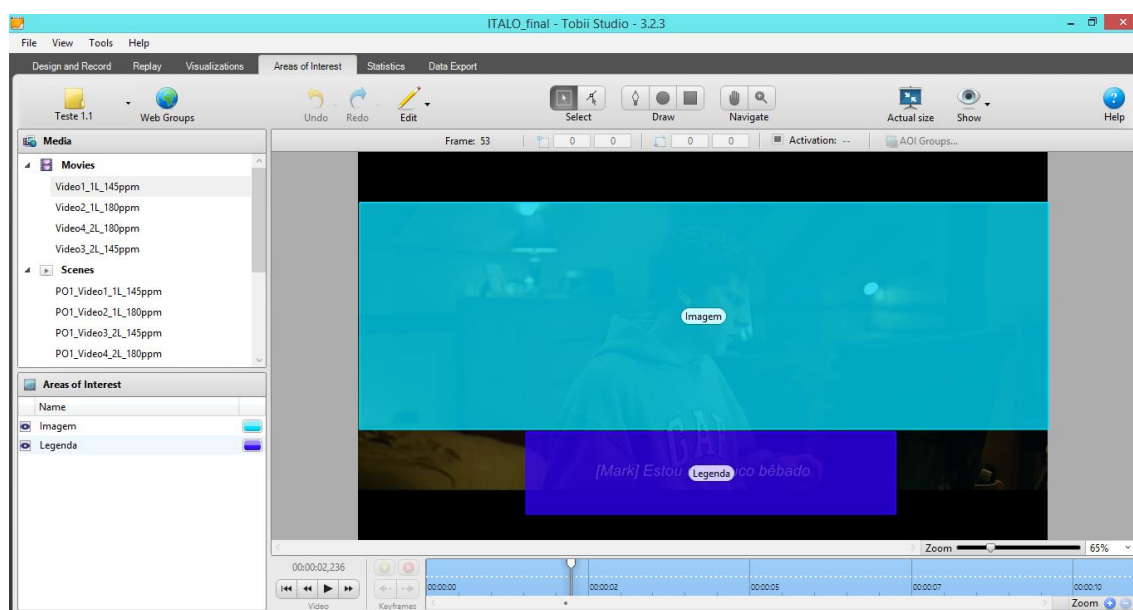
**Figura 16 – Interface do *Tobii Studio* aberta para início da realização do experimento**



Fonte: Elaborada pelo autor.

Um procedimento que foi feito no *Tobii Studio* após a coleta de dados dos participantes foi o de marcação de Áreas de Interesse nos estímulos visualizados por cada participante. A marcação de Áreas de Interesse, no nosso caso, área da legenda e área da imagem, delimitam os dados fornecidos sobre as fixações dos participantes na exportação de dados números através de planilhas. É possível que se estabeleçam as Áreas de Interesse que estejam ativas durante todo o tempo de estímulo, fazendo com que o programa gere dados estatísticos automáticos sobre essas fixações. Visto que apenas nos interessam as fixações na área da legenda, contudo, e de forma a reduzir a carga de trabalho na inspeção da movimentação ocular dos participantes, tivemos o cuidado extra de ativar (e desativar) (*Show* no canto superior direito da Figura 17) as Áreas de Interesse apenas durante a exibição das legendas, o que demandou mais trabalho durante essa etapa. Abaixo, a ferramenta Área de Interesse (*Area of Interest*) aberta no *Tobii Studio*.

**Figura 17 – Interface do *Tobii Studio* com Áreas de Interesse delimitadas em um dos estímulos**

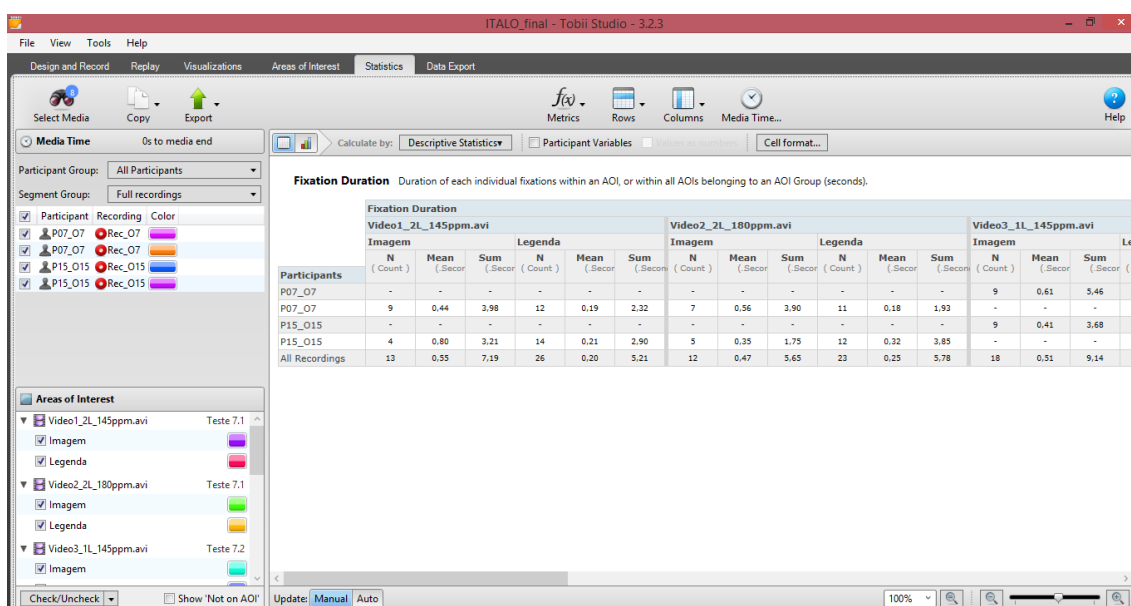


Fonte: Elaborada pelo autor.

O rastreador ocular gera dados quantitativos que podem ser visualizados através de ferramentas do *Tobii Studio*. Esses dados podem ser obtidos de duas formas: 1) de forma automática através da ferramenta *statistics*; e 2) por meio do acompanhamento do olhar com auxílio da ferramenta *GazePlot*.

Por meio da ferramenta *Statistics*, podemos visualizar os dados automáticos relativos às fixações dos participantes. Como as fixações realizadas no experimento precisavam ser refinadas a partir das categorias previamente citadas, o uso do *Statistics* foi feito no trabalho para efeitos de comparação e revisão com os dados refinados na inspeção do olhar. Todavia, como as Áreas de Interesse foram ativadas apenas nos momentos em que há legenda nos estímulos, os dados gerados por essa ferramenta são mais confiáveis do que se tivessem sido gerados até quando não houvesse exibição de legenda. Abaixo, na Figura 18, podemos ver a análise dos dados de dois participantes, O7 e 15 (legenda para identificar os Participantes Ouvintes 7 e 15, respectivamente), em relação ao número e duração média de fixação no experimento.

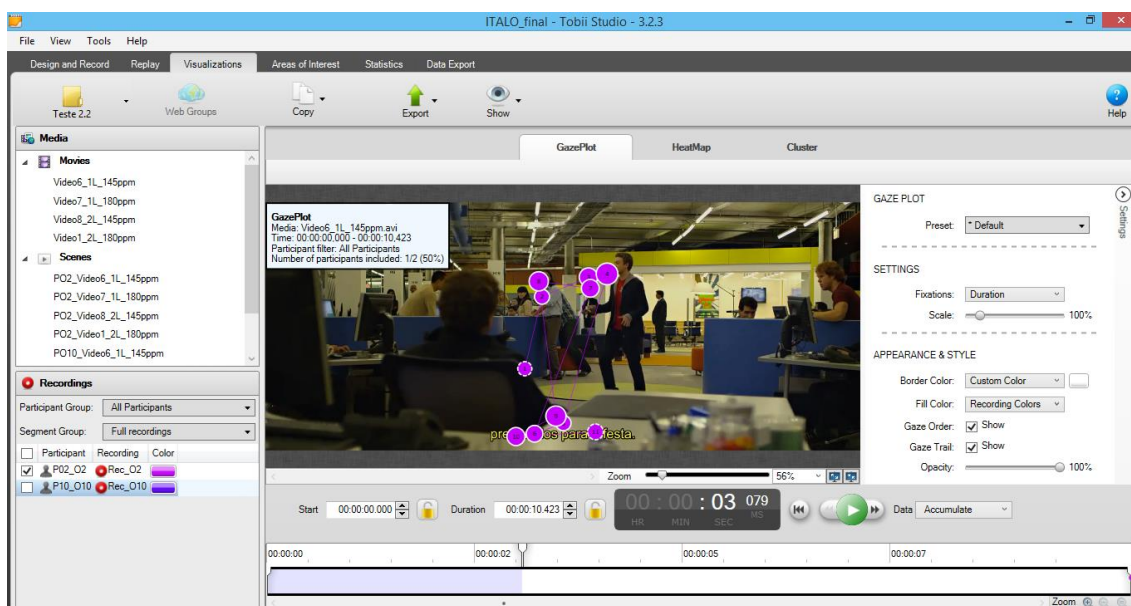
**Figura 18 – Interface da ferramenta *Statistics* com dados sobre número e duração média de fixações**



Fonte: Elaborada pelo autor.

A ferramenta *GazePlot*, por sua vez, presente na aba *Visualizations*, mostra a sequência e posição de fixação nos estímulos através de pequenos pontos, os quais são numerados de acordo com a sequência de fixações, e o seu tamanho indica a duração das fixações. Essa ferramenta pode auxiliar a visualização das fixações de um participante por vez ou de vários ao mesmo tempo em um mesmo estímulo. Abaixo, na Figura 19, a ferramenta *GazePlot* indica a sequência e duração de fixações de um dos participantes no Vídeo 2.

**Figura 19 – Interface da ferramenta *GazePlot***



Fonte: Elaborada pelo autor.

A inspeção do olhar dos participantes por essa ferramenta permite o estudo de medidas como regressões e releituras, por exemplo, assim como a identificação de fixações que realmente processaram palavras ao invés de distrações e pousos na área da legenda.

#### 3.4.6 Procedimentos éticos

O projeto ExLeg, de processo número 30569414.9.0000.5534, foi submetido, através da Plataforma Brasil, ao comitê de ética da Universidade Estadual do Ceará e teve parecer favorável. Como esta pesquisa acontece no âmbito do projeto, essa também é a autorização do comitê de ética para a realização deste estudo.

O Termo de Consentimento Livre e Esclarecido (TCLE), presente nos Apêndices C e D (para ouvintes e surdos, respectivamente), informa aos participantes os pormenores da pesquisa de modo a esclarecê-los sobre as implicações do estudo e seus possíveis benefícios/desconfortos. Além disso, estão presentes os contatos tanto do comitê quanto do pesquisador, o objetivo da pesquisa e a garantia de ressarcimento por quaisquer danos. No TCLE também está explícito que participantes poderiam deixar de participar da pesquisa e os seus dados seriam descartados a qualquer momento caso fosse do seu interesse. Os participantes poderiam deixar de participar da pesquisa a qualquer momento.

#### 3.4.7 Procedimentos de coleta do comportamento ocular

A coleta de dados dos 40 participantes ocorreu entre os meses de fevereiro e agosto de 2020. Durante o período de isolamento social mais rígido devido à Covid-19, paralisamos a coleta de dados, retornando apenas quando os números relativos à contaminação estavam baixos em Fortaleza e seguindo todos os protocolos de segurança. As sessões de coleta foram individuais e duraram uma média de 15 a 20 minutos por participante. O experimento foi iniciado com a leitura do TCLE e, a partir da concordância do participante, a posterior assinatura do termo. Para os participantes surdos, todas as instruções e orientações eram feitas em Libras. Após essa etapa, foi aplicado um questionário pré-coleta para o estabelecimento do perfil dos participantes (Apêndices A e B). Como as instruções dentro da tarefa no *Tobii Studio* eram bem sucintas, a fim de não confundir os participantes, foi dado a eles um resumo do passo a passo do experimento de forma oral. No geral, eles foram informados que veriam trechos de um filme legendados e de curta duração e que esses vídeos estavam organizados em 2 blocos de 4; ao final de cada bloco, eles responderiam a perguntas sobre o conteúdo dos vídeos. A eles

também enfatizamos que não se estressassem caso não conseguissem lembrar de determinados aspectos dos vídeos legendados, visto que era um teste de processamento, não de memória. Além disso, ressaltamos que o que estava sendo testado eram os vídeos, não eles.

Depois disso, para o início do experimento no rastreador ocular, os participantes se sentavam a uma distância de mais ou menos 65 cm do monitor do *Tobii T120 Eye Tracker*, distância esta necessária para a calibragem dos olhos do participante, de modo que o rastreador ocular conseguisse captar os seus movimentos durante a reprodução do estímulo. Para a calibragem, é realizada uma tarefa no monitor do *Tobii Studio* antes da exibição dos estímulos do experimento, durante a qual os participantes seguem com o olhar a movimentação de uma bola vermelha. Após essa tarefa, o programa apresenta um quadro com círculos em vários pontos da tela, os quais são preenchidos com uma cor verde. Quanto maior a precisão da calibragem, mais precisa é a captação dos dados. Contudo, a porcentagem de precisão dessa captação só pode ser vista após o experimento. Dessa forma, como utilizamos uma porcentagem mínima de 77% de precisão, um participante teve os seus dados descartados ao obter 65% (ao final, coletamos os dados de 41 participantes, visto que tivemos que descartar os dados desse participante em questão).

#### 3.4.8 Procedimentos de análise

A análise foi realizada a partir da inspeção do movimento do olhar feito com o auxílio da ferramenta *GazePlot*, dos dados numéricos exportados pelo programa em planilhas no formato .xlsx e dos dados gerados pela ferramenta *Statistics*, os quais foram complementares à observação feita com o *GazePlot*. A partir desses dados, montamos um diário de bordo em uma planilha de formato .xlsx em que discriminamos as fixações de acordo com a medida que ela representava (uma regressão, por exemplo). A minúcia nessa etapa de análise e a organização do banco de dados para posterior análise estatística demandaram um tempo considerável da nossa análise. Contudo, acreditamos que esse acompanhamento do olhar, conforme feito por Vieira (2016) e por Souza (2020), fortalece a robustez dos dados encontrados, porque garante que as fixações analisadas são realmente indícios de um processo de leitura.

Para a análise estatística dos dados, utilizamos o programa *RStudio*. Esse programa utiliza o R, principal linguagem de programação estatística no mundo, para realização de tarefas de cálculo de estatística descritiva e inferencial. Além disso, o programa é utilizado para visualização e exportação de gráficos.

Por meio do *RStudio*, aplicamos 4 testes estatísticos de acordo com os nossos dados e as necessidades da pesquisa: 1) Análise de Variância<sup>49</sup> (ANOVA); 2) Teste T Unicaudal<sup>50</sup>; 3) Modelo de Regressão Linear Misto; e 4) Modelo de Regressão Logística Misto<sup>51</sup>.

Na próxima seção, apresentamos os resultados sobre a movimentação ocular dos participantes da pesquisa e discutimos esses dados à luz da teoria em legendagem/processamento de legendas e as hipóteses do trabalho.

---

<sup>49</sup> A ANOVA testa se as médias de duas ou mais amostras são iguais de acordo com as variáveis independentes (fatores) testadas. Esse teste é utilizado com variável numérica contínua, como a duração média de fixação.

<sup>50</sup> O Teste T se assemelha à ANOVA em relação ao tipo de dado testado (variável numérica contínua). Além disso, uma das possibilidades do Teste T também é o de testar uma hipótese de acordo com as médias de duas amostras (apenas duas, no caso desse teste). Contudo, uma especificidade do Teste T que não é possível de se realizar com a ANOVA é a de direcionar a expectativa da hipótese para mais ou para menos, ao invés de apenas examinar se duas amostras são diferentes. Ou seja, como a nossa hipótese relativa ao número de linhas previa que as legendas de 1 linha seriam mais custosas do que legendas de 2 linhas, a utilização do Teste T Unicaudal permitiu observar os dados de acordo com essa expectativa, como no caso da medida de duração média de fixação que será apresentada na próxima seção.

<sup>51</sup> Os Modelos de Regressão Linear e Logístico permitem ao pesquisador resultados mais robustos do que a ANOVA e o Teste T. Eles não apenas vislumbram uma diferença entre médias de amostras, mas estimam um valor para a variável resposta (medida) de acordo com as variáveis independentes significativas no modelo que podem ser extrapoladas para além das amostras do estudo. Ou seja, eles estimam valores que consideram uma população mais ampla do que aquelas amostras coletadas, por isso são mais robustos. Enquanto o Modelo de Regressão Linear trabalha com uma variável resposta numérica contínua, o Modelo de Regressão Logística trabalha com uma variável categórica binária. O caráter “misto” desses modelos diz respeito à variação natural entre participantes e itens experimentais (vídeos legendados, no nosso caso), os quais também são importantes na estimativa desses valores. Mais detalhes em relação a esses modelos são dados na próxima seção.

## 4 RESULTADOS E DISCUSSÃO DOS DADOS

Nesta seção, apresentamos os resultados da análise advindos do estudo da movimentação ocular de 20 participantes surdos e 20 ouvintes (N = 40) do experimento. A análise se pautou em dados quantitativos referentes às fixações dos participantes do experimento. Dividimos esta seção em dois momentos: em 4.1, apresentamos toda a estatística descritiva e inferencial das medidas analisadas no trabalho (duração média de fixação, número de fixações, duração média e ocorrência de regressão, duração média e ocorrência de releitura, duração média de encapsulamento, ocorrência de deflexão e duração média de latência). Essa subseção é extensa, dada a minúcia com que tratamos os dados quantitativos, explorando tabelas e gráficos de visualização de dados para o tratamento estatístico desses dados. Neste primeiro momento, a correlação com a teoria é pouca, visto que é uma etapa mais focada na compreensão do percurso que levou ao uso de cada um dos testes, assim como do que consideramos *outliers* ou não. Em 4.2, por ocasião da discussão dos dados, fazemos um resumo dos achados apresentados em 4.1 de acordo com cada medida analisada, correlacionando com estudos anteriores que utilizaram objeto de estudo e metodologia similares aos do presente trabalho (D'YDEWALLE *et al.*, 1987; D'YDEWALLE; DE BRUYCKER, 2007; DE LINDE; KAY, 1999; MONTEIRO, 2016; SOUZA, 2020; SZARKOWSKA; GERBER-MORÓN, 2018; VIEIRA, 2016, dentre outros), assim como autores basilares no estudo do processamento linguístico através da movimentação ocular (JUST; CARPENTER, 1980; RAYNER, 1998; RAYNER; LIVERSEDGE, 2004; RAYNER; BARBARA; POLLATSEK, 2013; REICHLE; RAYNER; POLLATSEK, 2003).

### 4.1 Resultados quantitativos sobre a movimentação ocular dos participantes

Antes de apresentar os dados quantitativos sobre a movimentação ocular dos participantes *per se*, é interessante explicar como a primeira parte da análise, a inspeção da movimentação ocular do olhar dos participantes, ocorreu, por meio da ferramenta *GazePlot* do *Tobii Studio*. Apesar de termos delimitado e ativado as áreas de interesse correspondentes à área da imagem e da legenda no *software Tobii Studio* versão 3.3.2, os dados sobre fixação que se encontram nas planilhas exportadas pelo programa (utilizadas para análise estatística posteriormente) necessitavam ser tratados, pois nem toda fixação na área da legenda é realmente um processo de leitura. Em consequência disso, observamos no *GazePlot* as fixações que consistiam em pouso na área de interesse e não iniciavam um movimento natural de leitura

de esquerda para direita<sup>52</sup> (KRUGER; STEYN, 2014). Depois disso, fomos limpando as linhas de informação no arquivo .xlsx aberto no *Excel*. Abaixo, na Figura 20, observamos a ferramenta *GazePlot* sendo utilizada para inspeção da movimentação do olhar dos participantes.

**Figura 20 – *GazePlot* durante inspeção da movimentação do olhar**



Fonte: Elaborada pelo autor

Nesse exemplo, temos a movimentação do olhar do participante ouvinte identificado como O6 durante a exibição da primeira inserção de legenda do vídeo 5 na condição experimental 2L180. Esses círculos amarelos representam o percurso do olhar do participante identificados por número e ligados por linhas (sacadas) que representem sua sequência. Como dito anteriormente, uma das opções de visualização nesta ferramenta é a de que a duração das fixações seja representada pelo tamanho desses círculos, conforme podemos observar na Figura 20. Nela, podemos perceber que as fixações de 1 a 9 estavam na área da imagem antes de a legenda aparecer, concentrando-se na parte central da tela e no rosto do personagem à direita. No momento em que a legenda aparece, o participante demora 302 milissegundos para realizar a primeira fixação na área da legenda, a fixação 10. Essa medida, denominada como tempo de latência, é mais uma medida de alocação de atenção do que processamento em si, entretanto, diz respeito ao tempo disponível pelo espectador para processar a legenda. Ou seja, quanto maior o tempo de latência, menos tempo disponível para

<sup>52</sup> No caso da língua portuguesa e das outras línguas ocidentais em geral.



esse processamento. Essa foi uma das medidas observadas por d'Ydewalle *et al.* (1987) e que mostrou que legendas de 2 linhas normalmente demandam um tempo maior de latência; no entanto, proporcionalmente, 2 legendas de 1 linha demandam mais tempo. Essa diferença em termos de latência não foi significativa entre adultos e crianças (D'YDEWALLE; DE BRUYCKER, 2007).

Ainda na Figura 20, como dito no parágrafo acima, a Fixação 10 consiste em uma fixação de pouso e, dessa forma, foi retirada dos dados para análise estatística. Por ser de pouso, não é o início de uma leitura, pelo fato de a fixação número 11 estar à esquerda dela. No exemplo, o início do movimento de leitura natural da esquerda para a direita é composto de 5 fixações na primeira linha, da 11 à 15. Similarmente, a fixação 16, primeira fixação na segunda linha da legenda, também consiste em uma fixação de pouso. O movimento de leitura da segunda linha da legenda se inicia na fixação 17 e vai até a 20. Após essa fixação, o participante em questão fez outras fixações enquanto movimentos regressivos nesta legenda, os quais não foram retirados dos dados da análise, mas serão reportados com mais detalhes mais à frente.

Esse processo de inspeção do olhar foi feito com os dados dos 40 participantes – 20 ouvintes e 20 surdos – do experimento. Como cada participante viu 4 vídeos de 1 linha (4 inserções de legenda por vídeo) e 4 vídeos de 2 linhas (2 inserções de legenda por vídeo), totalizando 24 legendas processadas por participante, foram analisadas as fixações de 960 legendas, sendo 640 inserções de 1 linha e 320 inserções de 2 linhas. Da mesma forma, como a metade das 960 legendas era composta de legendas de 145 ppm e a outra metade, de legendas de 180 ppm, foram analisadas 480 legendas de cada uma dessas velocidades.

Para fins de organização, os resultados nesta subseção estão dispostos por variável dependente (também conhecida como “medida” ou “variável resposta”), sendo estas: 1) duração média de fixação; 2) número de fixações; 3) duração média de regressão; 4) ocorrência de regressão; 5) duração média de releitura; 6) ocorrência de releitura; 7) duração média de encapsulamento; 8) ocorrência de deflexão; e 9) duração média de latência. Dentre estas, podemos observar variáveis dependentes numéricas contínuas<sup>53</sup>, como as de duração média de fixação, número de fixação, duração média de regressão, duração média de releitura e duração média de latência. Também é possível observar variáveis categóricas<sup>54</sup> – ou nominais – como

---

<sup>53</sup> Uma variável numérica contínua é um número que, dentro de uma determinada grandeza, pode assumir qualquer valor em um intervalo de números. No nosso caso, a variável numérica contínua está relacionada ao valor em milissegundos das fixações.

<sup>54</sup> Uma variável categórica ou nominal contém um número finito de possibilidades. No contexto deste trabalho, tomando como exemplo a ocorrência de regressão, essa regressão pode ocorrer (sim) ou pode não ocorrer (não). Neste caso, assim como nos outros sobre ocorrência de releitura e deflexão, interpretamos a “não ocorrência”

ocorrência de regressão, ocorrência de releitura e ocorrência de deflexão. Visto que algumas medidas, como regressão, por exemplo, são analisadas tanto em relação a sua média quanto a sua ocorrência, elas estão dispostas uma após a outra na mesma subseção.

De acordo com cada medida e como ela está relacionada às variáveis independentes do estudo (Grupo x Número de linhas x Velocidade<sup>55</sup>), apresento primeiramente os dados de forma descritiva através de tabelas e gráficos (histogramas<sup>56</sup>, *boxplots*<sup>57</sup> e gráficos de barras) para então apresentar os dados advindos dos testes estatísticos realizados.

#### 4.1.1 Duração média das fixações

A duração média das fixações é uma medida muito importante para o exame do processamento leitor segundo o Princípio da Ligação Olho-Mente (JUST; CARPENTER, 1980; RAYNER, 1998), sendo neste estudo a medida central para examinarmos a influência do número de linhas e da velocidade nos dois grupos de participantes (surdos e ouvintes).

Como dito anteriormente, apesar de ser possível extrair automaticamente a duração de média de fixação no *Tobii Studio* ao determinarmos a área de interesse no *software*, a análise automática não apresenta, contudo, o que acontece durante o processamento; assim, algumas medidas, como regressões, releitura, latência, só podem ser analisadas se acompanharmos o comportamento do olhar. Nesse sentido, Rayner e Liversedge (2004) ressaltam a importância de acompanhar todo o movimento ocular durante a leitura, a fim de entender essas diferenças de tempo.

Destarte, as 4.038 observações relativas às fixações foram obtidas por meio da inspeção do movimento do olhar. Além disso, dada a nossa experiência com pesquisas anteriores sobre legendagem e movimentação ocular, decidimos extrair apenas fixações que

---

como “0” e a ocorrência como “1” de acordo com a exigência do teste estatístico de regressão logística, o qual utilizamos para esse tipo de variável.

<sup>55</sup> Quando nos referirmos aos resultados das variáveis independentes do estudo, vamos utilizar a notação com letra maiúscula (“Velocidade”, por exemplo), visto que foi assim que as variáveis foram codificadas no *RStudio* e, dessa maneira, é assim que estão presentes nos gráficos gerados pelo programa. Quando estivermos falando dessas variáveis de uma maneira mais ampla, para além de como foram manipuladas neste trabalho, vamos utilizar a notação com letra minúscula (“número de linhas”, por exemplo).

<sup>56</sup> Um histograma é um gráfico de barra ou coluna que representa a distribuição de frequência de determinada medida. No contexto desse estudo, utilizamos para dividir em classes de milissegundos as durações das fixações dos participantes. Por exemplo, na medida de duração média de fixação, cada coluna do histograma representa o número de fixações que os participantes realizaram naquele intervalo de milissegundos (250 a 300 milissegundos, por exemplo).

<sup>57</sup> Um *boxplot*, ou diagrama de caixa, é uma representação gráfica que apresenta a posição, dispersão, simetria, caudas e *outliers* dos dados. No contexto desse trabalho, ele foi utilizado como recurso para identificação dos valores discrepantes (*outliers*), que foram representados graficamente como pontos que ficam além dos limites inferior e superior da caixa.

acontecessem no intervalo de 80 e 600 milissegundos para essa medida específica (apesar de reconhecermos que em algumas pesquisas com rastreamento ocular com textos impressos há um limite de até 800 milissegundos para fixações). Contudo, sabendo que a maior parte dos testes estatísticos utilizados aqui tem como pressuposição a distribuição normal dos dados e se mantivéssemos o limite de 800 milissegundos, ao examinarmos os dados, teríamos uma quantidade grande de *outliers*, decidimos por realizar esse filtro previamente<sup>58</sup>. No entanto, em medidas que apresentaram uma quantidade pequena de dados, utilizamos o corte de 800 milissegundos, como será explicitado mais adiante.

Dessa maneira, revendo as expectativas sobre essa medida em relação às hipóteses do trabalho, esperávamos que: 1) os participantes surdos tivessem uma duração média maior de fixação; 2) as legendas de 1 linha tivessem uma duração média maior de fixação; e 3) as legendas de 145 e 180 palavras por minuto não diferissem significativamente em relação a essa média.

Antes da etapa de análise estatística, é interessante observarmos, através de números de estatística descritiva, como essa variável resposta se apresentou em nossos dados a partir das variáveis independentes (Grupo, Número de Linhas e Velocidade). Assim, temos nas Tabelas 2, 3 e 4 o tempo médio de duração de fixação em milissegundos, além de outras informações, como o desvio padrão, erro padrão e mediana.

**Tabela 2 – Duração média das fixações por Grupo**

Grupo	Duração de fixação (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>Ouvinte</b>	263,17	98,93	2,13	250
<b>Surdo</b>	299,07	110,29	2,53	267

Fonte: Elaborada pelo autor.

**Tabela 3 – Duração média das fixações por Número de Linhas**

Número de Linhas	Duração de fixação (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>1</b>	283,24	107,13	2,42	266
<b>2</b>	277,07	104,76	2,28	250

Fonte: Elaborada pelo autor.

<sup>58</sup> Ou seja, fizemos uma limpeza prévia das fixações para análise da medida de duração média de fixações, de acordo com os valores mencionados; mesmo assim, realizamos outra checagem posterior em relação aos *outliers* através da visualização por *boxplot*, como será mostrado adiante.

**Tabela 4 – Duração média das fixações por Velocidade**

Velocidade	Duração de fixação (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>145</b>	280,17	104,94	2,27	266
<b>180</b>	279,90	107,07	2,45	250

Fonte: Elaborada pelo autor.

Em uma primeira vista nos dados da estatística descritiva, podemos perceber que a diferença entre as médias de Grupo é grande (263,17 ms para ouvintes e 299,07 ms para surdos), apontando para a expectativa de que surdos, enquanto leitores de português como L2, têm um maior custo de processamento na leitura de legendas. Além disso, o desvio padrão maior no grupo de surdos aponta para uma heterogeneidade maior das observações extraídas, algo que também era previsto, dada a composição desse grupo de participantes (DE LINDE; KAY, 1999). Em relação ao Número de Linhas, há uma diferença também na direção que esperávamos: tempo maior em legendas de 1 linha do que nas legendas de 2 linhas (283,24 ms vs. 277,07 ms, respectivamente), mas de forma bem menos acentuada do que em Grupo. Já em Velocidade, a diferença entre legendas de 145 e 180 ppm é praticamente inexistente (280,17 ms vs. 279,90 ms, respectivamente), apontando também para a confirmação de nossa hipótese alternativa<sup>59</sup> de que não haveria diferença significativa entre esses dois níveis de velocidade.

Em relação às condições experimentais do estudo (1L145, 1L180, 2L145 e 2L180), a duração média de fixação por surdos e ouvintes pode ser vista abaixo, na Tabela 5.

**Tabela 5 – Duração média de fixação entre condições experimentais**

Grupo	Número de Linhas	Velocidade	Duração Média de Fixação (ms)
<b>Surdo</b>	1	145	302,71
	1	180	298,60
	2	145	290,27
	2	180	305,66
<b>Ouvinte</b>	1	145	270,60
	1	180	263,44
	2	145	261,16
	2	180	257,91

Fonte: Elaborada pelo autor.

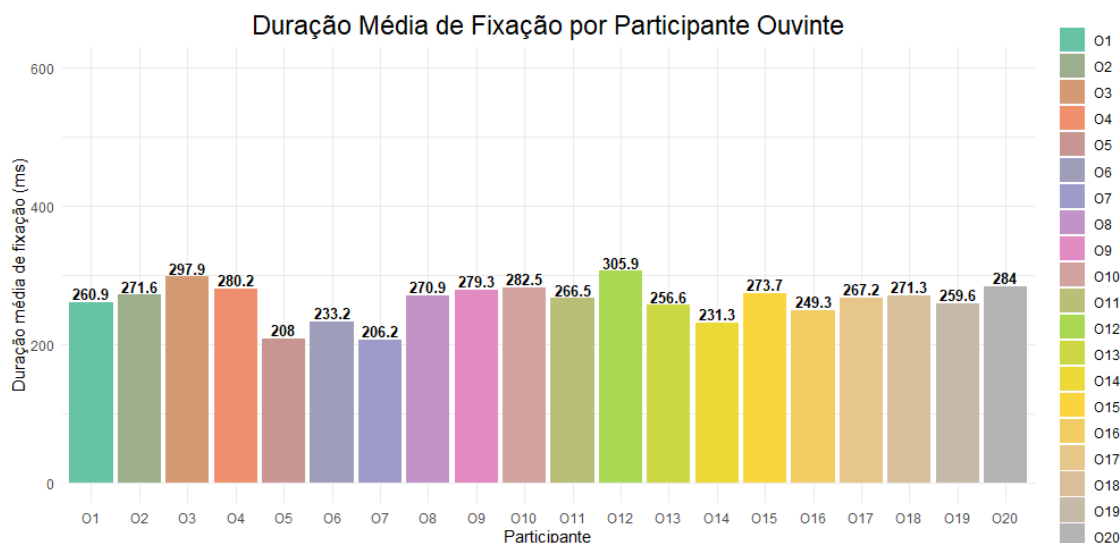
<sup>59</sup> Em estatística, a hipótese alternativa, ou  $H^1$ , é a hipótese formulada pelo estudo a partir da teoria e pesquisas prévias; ou seja, a hipótese alternativa quer provar que a hipótese nula, ou  $H^0$ , está equivocada. No caso deste exemplo, a hipótese nula é de que há diferença significativa entre as velocidades de 145 ppm e 180 ppm no processamento de legendas por surdos e ouvintes. Caso não haja diferença significativa, confirmamos, assim, a hipótese do trabalho, ou seja, a hipótese alternativa. Para os efeitos deste trabalho, utilizamos, na maior parte das ocorrências, o termo mais geral “hipótese” no lugar de “hipótese alternativa” (hipótese do trabalho).

De acordo com os dados dispostos acima, os ouvintes tiveram um processamento menos custoso nas condições experimentais de 2L: 2L145 (261,16 ms) e, mais ainda, 2L180 (257,91 ms). Ainda sobre os ouvintes, enquanto a diferença de 1L180 (263,44 ms) esteve próxima das condições experimentais 2L145 e 2L180, 1L145 (270,6 ms) apresentou um distanciamento maior, sendo a condição experimental mais custosa para os participantes ouvintes.

Já em relação aos surdos, por sua vez, os números de duração média de fixação nas condições experimentais com 2 linhas (2L145 e 2L180) foram ligeiramente diferentes. Ao passo que a condição experimental 2L145 (290,27 ms) teve o menor tempo médio e, dessa forma, o menor custo de processamento para os participantes, a 2L180 (305,66 ms) foi a mais desconfortável entre todas as condições experimentais, sugerindo uma possível interação entre Velocidade e Número de Linhas entre surdos: 2 linhas são confortáveis, mas apenas quando em uma velocidade baixa (145 ppm). Além disso, a condição experimental 1L180 (298,6 ms) teve o segundo menor tempo médio de fixação, seguido por 1L145 (302,71 ms).

Quando consideramos apenas o Número de Linhas e Grupo, sem Velocidade, os ouvintes tiveram um processamento mais rápido em legendas de 2 linhas (259,65 ms) do que 1 linha (267,08 ms). Em surdos, essa tendência também se repetiu, mas a diferença foi menor: legendas de 1 linha tiveram 300,79 ms, enquanto legendas de 2 linhas tiveram 297,41 ms.

Após realizarem a tarefa de assistir aos 8 vídeos manipulados em cada condição experimental (1L145, 1L180, 2L145, 2L180), sendo que cada participante assistiu a dois itens por condição, o grupo de participantes ouvintes apresentou o seguinte comportamento em relação à duração média de fixação, conforme é possível observar no Gráfico 1 abaixo.

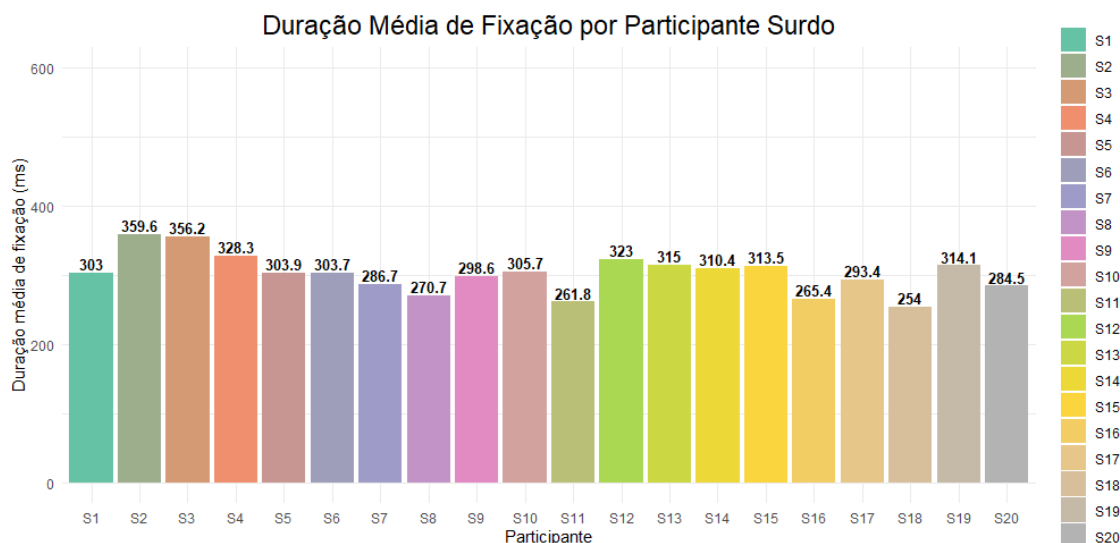
**Gráfico 1 – Duração média de fixação por participante ouvinte**

Fonte: Elaborado pelo autor.

Os participantes O5 e O7<sup>60</sup> tiveram médias bem abaixo do restante dos participantes (208 e 206,2 milissegundos, respectivamente), número abaixo até do tempo médio de fixação para uma leitura silenciosa em textos impressos, que é de 225 ms (RAYNER, 1998). No lado inverso da moeda, os participantes O3 e O12 apresentaram as maiores médias de tempo de fixação no experimento (297,9 e 305,9 milissegundos, respectivamente), inclusive sendo um tempo médio de fixação próximo ao da média de duração de fixação dos participantes surdos lendo português como L2 nesse estudo.

Os participantes surdos, por sua vez, tiveram suas médias de duração de fixação distribuídas da seguinte forma:

<sup>60</sup> No caso da apresentação das médias individuais por participante do experimento nas medidas ao longo desta seção, isso se dá mais no intuito de correlacionar possíveis dificuldades ou facilidades no processamento a partir das medidas analisadas. É importante enfatizar, contudo, que a análise estatística colocada em prática aqui vê o grupo de participantes, tanto surdo como ouvinte, como um todo, e não individualmente.

**Gráfico 2 – Duração média de fixação por participante surdo**

Fonte: Elaborado pelo autor.

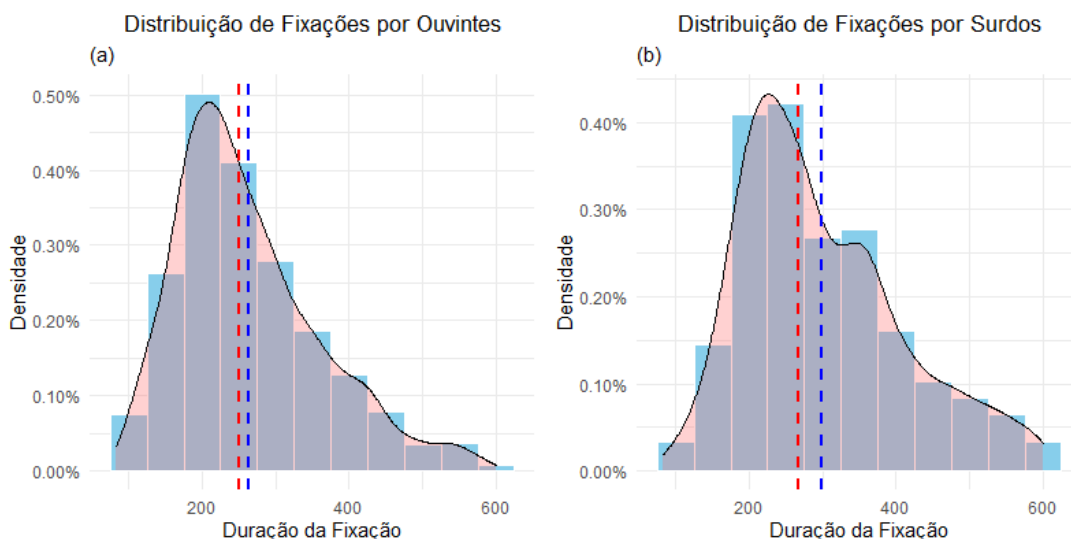
Os participantes S11 e S18 tiveram as menores médias do grupo (261,8 e 254 milissegundos, respectivamente), médias estas, inclusive, semelhantes às de muitos participantes do grupo de ouvintes, sugerindo que são leitores proficientes em língua portuguesa e, provavelmente, acostumados a assistir a vídeos legendados. Os participantes S2 e S3, por sua vez, tiveram as maiores médias de fixação no grupo (359,6 e 356,2 milissegundos, respectivamente), médias altas até mesmo para leitores supostamente proficientes em língua estrangeira. São médias maiores até que a média de reconhecimento de cena, conforme Rayner (1998).

De uma forma geral, a duração média de fixação maior dos participantes surdos já era esperada, pois, como dito anteriormente, o fato de os surdos terem a língua portuguesa como L2 – além de todo o contexto de como o surdo tem acesso à língua portuguesa, seja em ambientes formais de aprendizagem ou na sociedade como um todo – nos traz a expectativa de que o processamento desse grupo de participantes seja mais custoso em relação aos ouvintes. Contudo, a característica heterogênea da comunidade surda (DE LINDE; KAY, 1999) em termos de proficiência em língua oral e de acesso aos meios de comunicação por meio de legendas faz com que participantes como o S11 e o S18 tenham desempenho superior ao de vários ouvintes que participaram da pesquisa.

Além do tempo médio de fixação supracitado, é interessante observar os dados de desvio padrão (Tabela 2), que indicam uma dispersão maior no conjunto de dados dos participantes surdos (110,2), assim como no erro padrão do grupo (2,53), confirmando a expectativa em torno da heterogeneidade do grupo de surdos. Abaixo, no Gráfico 3, temos os

histogramas que nos dão mais detalhes sobre a distribuição dos dados sobre as fixações nas legendas nos dois grupos.

**Gráfico 3 – Distribuição das fixações de ouvintes e surdos**



Fonte: Elaborado pelo autor.

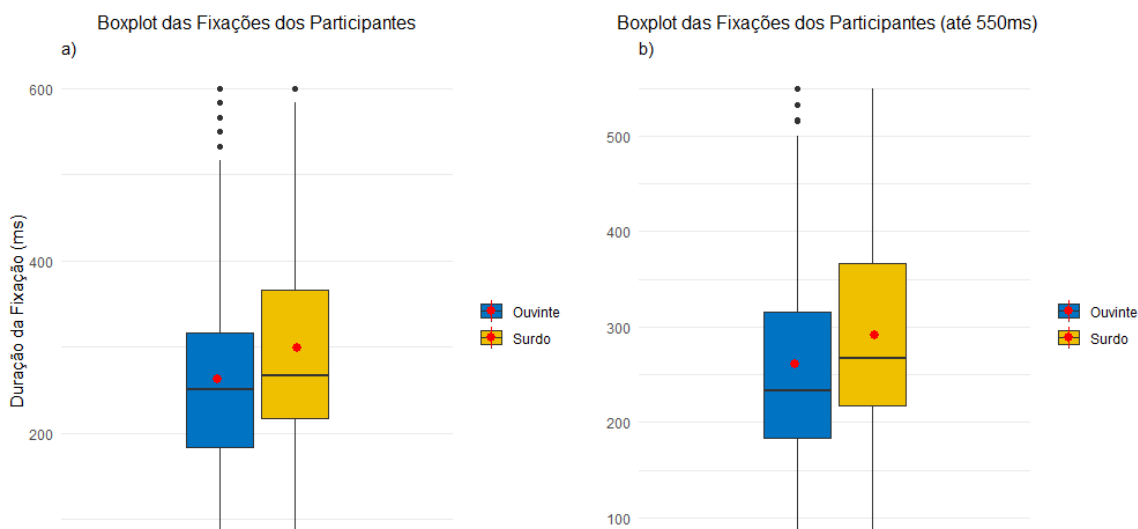
Os dois histogramas no Gráfico 3 mostram uma distribuição não normal<sup>61</sup> na frequência de fixações dos participantes em relação a sua duração, com a distribuição de ambos apresentando uma cauda à direita; ou seja, apesar de a maior parte dos dados estar na casa dos 200 milissegundos, existem fixações mais longas que puxam a média para a direita. Isso é possível visualizar ao notarmos que a linha tracejada azul (representando a média dos grupos) encontra-se à direita da vermelha (representando a mediana). O fato de a mediana do grupo de surdos (b) se afastar mais da média do que do grupo de ouvintes (a) mostra que os dados dos surdos são mais acentuados em direção à não normalidade.

Quando visualizamos o conjunto de todos os dados através de *boxplot* (Gráfico 4), percebemos uma quantidade de *outliers* na faixa dos 550 a 600 milissegundos.

<sup>61</sup> A normalidade da distribuição de uma variável é uma necessidade para a análise estatística. Uma normalidade perfeita dos dados pode ser vista em um histograma quando o valor da média e da mediana coincidem; ou seja, fosse este o caso dos nossos dados, as linhas tracejadas vermelha e azul dos histogramas no Gráfico 3 deveriam se tocar. Isso faria com que as caudas da direita e da esquerda de um histograma tivessem o mesmo tamanho. Contudo, uma distribuição normal perfeita é praticamente impossível quando se analisam dados de fixação do olhar. Rayner, Barbara e Pollatsek (2013, p. 99), inclusive, conforme apontado anteriormente, apresentam um histograma de distribuição de fixações segundo seu tempo de duração bem parecido com os presentes no Gráfico 3. Dessa forma, o que se espera é que, na verdade, esses dados não se afastem tanto assim da normalidade. Para a análise estatística, de forma sucinta, uma distribuição normal – ou próxima dela – torna mais fácil que os resíduos – valores dos dados que se afastam da estimativa – não se afastem tanto assim da distribuição esperada dos dados (GODOY, 2019).



**Gráfico 4 – Boxplot da duração média das fixações**



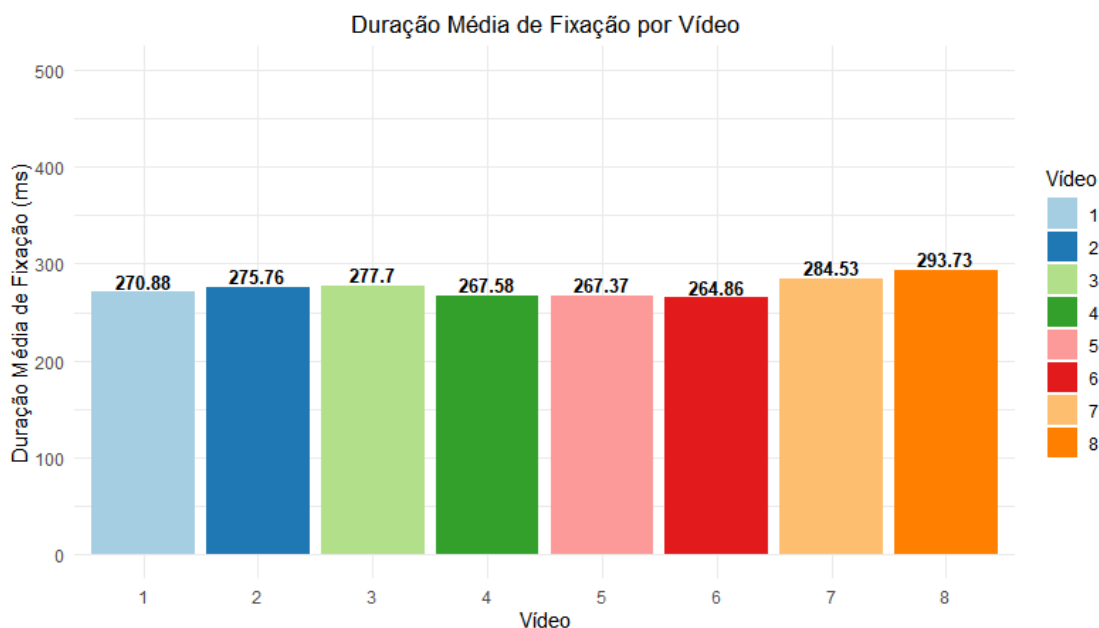
Fonte: Elaborado pelo Autor

Ao olharmos os *boxplots* divididos por grupos à esquerda (antes da filtragem) e à direita (depois da filtragem), percebemos que houve uma diminuição dos valores fora do limite superior (os pontinhos em preto) do gráfico, mas não seu desaparecimento por completo. Optamos por não diminuir ainda mais a linha de corte para que não perdêssemos mais dados do grupo de participantes surdos, principalmente, e por acreditarmos que o valor 550 ms é um valor de corte adequado. Dessa forma, restaram 3.973 observações.

Inicialmente, apesar de termos baseado a escolha dos vídeos utilizados – dentre outros fatores – na noção de que a questão imagética fosse de certa forma regular em todos os vídeos, sabíamos que utilizar 8 cenas distintas poderia resultar em relações causais espúrias entre variável dependente e independente. A partir disso, mesmo utilizando vídeo como efeito aleatório nos modelos de regressão ajustados ao longo do trabalho, como será delineado mais adiante, também tratamos o vídeo enquanto possível variável de confusão<sup>62</sup> para, dessa forma, investigarmos a possível influência desses itens experimentais ao longo dos testes estatísticos.

No Gráfico 5, vemos os tempos médios de fixação por vídeo (item experimental) utilizado no trabalho.

<sup>62</sup> Uma variável de confusão, em estatística, é um fator que modifica a associação entre causa e efeito. No caso deste estudo, a escolha dos vídeos pode ser uma variável de confusão, daí a importância de se testar a variância nos vídeos utilizados na pesquisa em legendagem e tentar neutralizar seu efeito nos testes estatísticos.

**Gráfico 5 – Duração Média de Fixação por Vídeo**

Fonte: Elaborado pelo autor.

A variância entre os vídeos 1, 2, 3, 4, 5 e 6 não é tão grande, variando em cerca de 10,33 milissegundos entre o Vídeo 3 (277,7 ms) e o Vídeo 6 (264,86 ms). Em relação às médias dos Vídeos 7 (284,53 ms) e 8 (293,73 ms), entretanto, observamos um distanciamento maior, sugerindo uma possível interferência de vídeo enquanto variável de confusão no experimento. Apesar de aplicarmos efeitos mistos no modelo de regressão linear utilizado no estudo, ou seja, de levarmos em consideração a variância entre itens – que, no nosso caso, estamos considerando o vídeo e sua legenda como um todo e não apenas palavras/sentenças, como é concebido o efeito misto de item na pesquisa em psicolinguística com textos estáticos –, decidimos checar se a variação de vídeo seria significativa em todos os nossos testes estatísticos com variáveis numéricas contínuas, porque não pudemos prever esse efeito mesmo tendo cuidado na seleção das cenas e estas serem de um mesmo filme.

Antes de apresentar os testes estatísticos, contudo, mostraremos novamente os tempos médios de fixação por variável independente após a filtragem dos dados para eliminação de *outliers*, conforme pode ser visualizado nas Tabelas 6, 7 e 8 abaixo.

**Tabela 6 – Duração média de fixação por grupo após filtragem**

Grupo	Duração de fixação (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>Ouvinte</b>	260,67	95,29	2,06	233
<b>Surdo</b>	291,73	101,71	2,36	267

Fonte: Elaborada pelo autor.

**Tabela 7 – Duração média de fixação por número de linha após filtragem**

Número de Linhas	Duração de fixação (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>1</b>	278,13	100,65	2,30	250
<b>2</b>	272,35	98,42	2,16	250

Fonte: Elaborada pelo autor.

**Tabela 8 – Duração média de fixação por velocidade após filtragem**

Velocidade	Duração de fixação (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>145</b>	276,35	99,98	2,17	250
<b>180</b>	273,76	99,03	2,29	250

Fonte: Elaborada pelo autor.

A partir dos dados das Tabelas acima, rodamos, inicialmente, uma análise de Variância (ANOVA) para a variável resposta de duração média de fixação, tendo como variáveis independentes Grupo, Número de Linhas, Velocidade e Vídeo, visto que queríamos investigar o efeito de Vídeo junto às outras variáveis sobre o tempo de fixação dos participantes. Esse teste retornou uma alta significância para Vídeo, conforme havíamos previsto ( $f(1, 3.962) = 4,902, p < 0,0001$ ). Ao aplicarmos um Teste de *Tukey* para investigar os pares múltiplos que seriam significativos, o Vídeo 8 apresentou diferença significativa em relação ao vídeo 6 ( $p = 0,0001$ ); o Vídeo 7, por sua vez, foi marginalmente significativo em relação ao Vídeo 6 ( $p = 0,0873$ ). Sendo assim, decidimos rodar a ANOVA tanto com todos os vídeos, quanto sem os Vídeos 7 e 8, visto que o 7 quase foi significativo e, como veremos adiante no modelo de regressão linear, ele foi significativo em relação ao Vídeo 1 do experimento.

Dessa forma, sem excluir os Vídeos 7 e 8, e utilizando Grupo, Número de Linhas e Velocidade enquanto variáveis, rodamos uma ANOVA e encontramos uma significância marginal na variável de Número de Linhas ( $f(1, 3.962) = 3,447, p = 0,0634$ ) e significância na variável de Grupo ( $f(1, 3.962) = 98,706, p < 0,0001$ ), ou seja, em relação ao número de linhas, os resultados mostraram uma possibilidade de essa variável influenciar o processamento leitor

dos participantes. Em relação à velocidade, conforme previsto, não encontramos significância na variável de velocidade ( $f(1, 3.962) = 0,536, p = 0,4627$ ).

Também rodamos uma ANOVA sem os Vídeos 7 e 8. Os resultados desse teste foram semelhantes aos supracitados: uma quase significância na variável de Número de Linhas ( $f(1, 3.024) = 2,896, p = 0,0889$ ), uma significância de Grupo ( $f(1, 3.024) = 80.955, p < 0, 0001$ ) e uma não significância de Velocidade ( $f(1, 3.024) = 0,371, p = 0,5426$ ).

Ao checarmos uma possível interação entre as variáveis de Grupo, Número de Linhas e Velocidade, uma ANOVA considerando todos os vídeos sugeriu uma significância *de Número de Linhas:Velocidade:Grupo* ( $f(1, 3.965) = 3,637, p = 0,0566$ ). Essa mesma interação foi confirmada pela ANOVA realizada sem os Vídeos 7 e 8 ( $f(1, 3.020) = 3,920, p = 0,0478$ ). Contudo, ao realizarmos um Teste de *Tukey* para investigar as comparações entre pares significativas, percebemos que a variável Grupo (par surdo/ouvinte) estava presente em todas as comparações significativas. Assim, não foi possível fazer a análise entre as três variáveis independentes (Grupo, Número de Linhas e Velocidade).

Por último, realizamos um Teste T para checar a significância marginal apresentada por Número de Linhas nas ANOVAs supracitadas. Apesar de funcionar dentro da mesma lógica de uma ANOVA, o Teste T rodado na interface do R, programa de análise de estatística utilizado nesta pesquisa, tem uma vantagem de poder se especificar a direção na qual esperamos que a hipótese alternativa se comporte (teste unicaudal). Isso é relevante no caso da variável Número de Linhas, especialmente, porque a hipótese desta pesquisa é que o processamento de legendas de 1 linha seja mais custoso como um todo e, dentro do parâmetro de duração média da fixação, isso significa um tempo maior de duração. Ou seja, a nossa hipótese não é apenas que as legendas de 1 e 2 linhas sejam diferentes, mas que a legenda de 1 linha apresente um tempo maior de duração média de fixação.

Assim, através do argumento *alternative = "greater"*, avisamos ao R que a expectativa é que o processamento de legendas de 1 linha (primeiro nível de referência lido pelo programa em ordem numérica crescente) seja maior e, neste caso, distribua os 5% de intervalo de confiança em uma direção só, e não 2,5% para cima e para baixo, como é padrão em testes que querem estimar apenas se a diferença atinge um nível de significância ou não.

Assim, ao rodarmos o Teste T unicaudal com todos os vídeos, observarmos que a diferença entre a duração média de fixação para 1 linha ( $M^{63} = 278,13, SD^{64} = 101$ ) e 2 linhas ( $M = 272,35, SD = 98,4$ ), ( $t(1) = 1,8266, p = 0,0339$ ) atingiu significância estatística. Da mesma

---

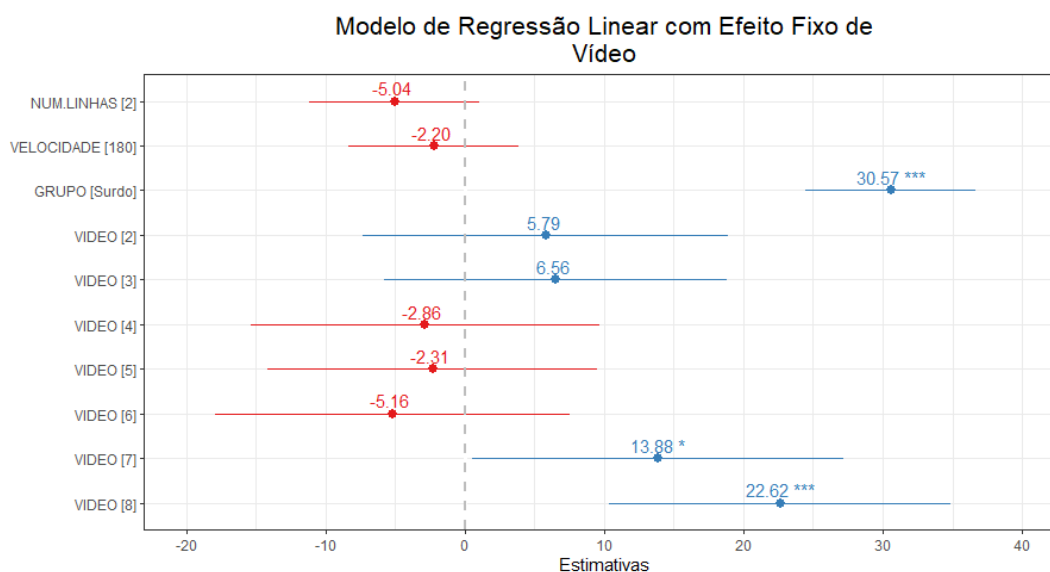
<sup>63</sup> “Mean” (“média” em inglês).

<sup>64</sup> “Standard Deviation” (“Desvio Padrão” em inglês).

forma, ao retirarmos os Vídeos 7 e 8, o resultado foi praticamente o mesmo, com o tempo médio de duração de fixação em legendas de 2 linhas ( $M = 267,55$ ,  $SD = 95,5$ ),  $t(1) = 1,7692$ ,  $p = 0,0384$ ) significativamente menor do que de 1 linha ( $M = 273,78$ ,  $SD = 98$ ).

Concomitantemente às ANOVAs e ao Teste T, rodamos um modelo de regressão linear (sem efeitos mistos) o qual foi ajustado com os efeitos fixos previstos no trabalho (Grupo, Número de Linhas e Velocidade) e a adição da possível variável de confusão (Vídeo). O resultado desse modelo retornou um efeito significativo para Vídeo, especificamente os Vídeos 7 ( $\beta = 13,88$ ,  $p = 0,04$ ) e 8 ( $\beta = 22,62$ ,  $p = 0,0002$ ). Nesse modelo, as variáveis de número de linhas ( $\beta = -5,04$ ,  $p = 0,1055$ ) e velocidade ( $\beta = -2,20$ ,  $p = 0,4804$ ) não apresentaram efeito significativo, mas não se trata do modelo mais ajustado possível, o qual será apresentado mais adiante. Essas informações podem ser vistas no modelo plotado abaixo (Gráfico 6).

**Gráfico 6<sup>65</sup> – Modelo de Regressão Linear com Efeito Fixo de Vídeo**



Fonte: Elaborado pelo autor.

Juntamente a esse, um Teste de *Tukey* foi rodado para uma comparação entre pares em relação ao efeito fixo de Vídeo aplicado no modelo. Os resultados indicaram que a comparação múltipla entre estimativas dos vídeos foi significativa em pares envolvendo o

<sup>65</sup> A interpretação desse tipo de gráfico, o qual será utilizado várias vezes ao longo desta seção, é dessa forma: as linhas horizontais coloridas ao lado dos efeitos mostram se, dentro daquele modelo, a estimativa do efeito foi positiva (linha azul) ou negativa (linha vermelha) de acordo com a medida (variável dependente) ajustada. No nosso caso, especificamente, mostra se a estimativa do tempo médio da duração de fixação caiu (linha vermelha) ou aumentou (linha azul) a partir das variáveis. Esse efeito, contudo, só é significativo caso a linha colorida não cruze a linha tracejada vertical que está em 0. Caso isso ocorra, o efeito não pode ser considerado significativo, pois não há a certeza se aquele efeito é positivo ou negativo de acordo com seu intervalo de confiança.

Vídeo 8, que teve média de duração de fixação maior em relação aos Vídeos 1 ( $\beta = 22,61$ ,  $p = 0,0007$ ), ao 4 ( $\beta = 25,46$ ,  $p = 0,0006$ ), ao 5 ( $\beta = 24,91$ ,  $p = 0,0002$ ) e ao 6 ( $\beta = 27,75$ ,  $p = 0,0001$ ), enquanto o Vídeo 7 apresentou uma diferença quase significativa em relação ao vídeo 6 ( $\beta = 19,02$ ,  $p = 0,0873$ ). Assim como fizemos com as ANOVAs acima, decidimos estimar os valores da variável resposta – a partir das variáveis independentes do trabalho – por meio de um modelo de regressão linear misto sem os Vídeos 7 e 8, conforme pode ser visto abaixo.

Ao tentarmos ajustar um modelo com convergência com o máximo de interceptos<sup>66</sup> e *slopes*<sup>67</sup> aleatórios possíveis<sup>68</sup>, foi alcançado um modelo de regressão linear misto com intercepto de participante e *slope* aleatório de velocidade por intercepto de participante e intercepto de vídeo (item), tendo as variáveis categóricas de Grupo, Número de Linhas e Velocidade como efeitos fixos e a duração média de fixação como variável numérica contínua. Ao rodarmos o modelo com todas as variáveis categóricas e percebermos que a variável de Velocidade não se mostrou significativa ( $p = 0,6632$ ), rodamos o modelo apenas com as variáveis de Grupo e de Número de Linhas. Abaixo, na Tabela 9, há uma descrição dos valores associados aos dois modelos (modelo com todos os preditores<sup>69</sup> e modelo sem Velocidade).

<sup>66</sup> De acordo com Godoy (2019), o valor de intercepto (*intercept*) em um modelo linear é o valor em que a reta ajustada pelo modelo cruza o eixo y quando o valor de x (o efeito fixo) é igual a zero no caso de variáveis independentes numéricas. No caso do nosso trabalho, em que utilizamos variáveis independentes categóricas, o valor do intercepto diz respeito ao valor de y (variável dependente ou, no nosso caso, tempo médio de fixação) quando ela está no primeiro nível de referência da variável independente (eixo x). Esse valor de referência normalmente é tomado por ordem alfabética ou, caso seja representado por um número, como em Número de Linhas, em ordem crescente. Ou seja, tomando por base um modelo de regressão linear com Grupo, Número de Linhas e Velocidade como variáveis independentes, o valor do intercepto será o tempo médio de duração de fixação nas seguintes circunstâncias: participante Ouvinte, 1 linha, Velocidade de 145 ppm (baixa), pois todos esses níveis são os primeiros na ordem alfabética ou numérica.

<sup>67</sup> O *slope* (“declive” em inglês), por sua vez, ainda de acordo com Godoy (2019), representa o quanto y aumenta ou diminui a cada unidade de x. No caso do nosso trabalho, que utiliza variáveis independentes categóricas, representa quanto y (o tempo médio de duração de fixação, utilizando como exemplo esta medida) aumenta ou diminui a cada mudança de nível; ou seja, quanto o tempo médio de duração de fixação diminui quando passamos do nível de referência de 1 linha para 2 linhas e assim por diante levando em consideração as outras variáveis independentes.

<sup>68</sup> De acordo com Godoy (2019, p. 55), é razoável que cada participante e cada item experimental influenciem as respostas do experimento de forma particular. Ou seja, cada participante e cada item experimental (os vídeos, no nosso caso) “contribui com uma variabilidade aleatória, não controlada pelo experimentador, mas que deve ser prevista no modelo”. Dessa forma, um modelo de regressão linear (ou logístico) com efeitos aleatórios (mistos) deve considerar interceptos diferentes (**interceptos aleatórios**) para cada participante e para cada item experimental. Além disso, o modelo deve considerar as diferenças dos declives (**slopes aleatórios**) associados a esses efeitos aleatórios a partir da variável independente (efeito fixo). Exemplo: um **modelo de regressão linear misto com intercepto de participante e slope aleatório de velocidade por intercepto de participante** significa que estamos modelando uma estimativa (duração média de fixação) a partir de como a resposta de cada participante varia quando passa do nível de 145 ppm (velocidade baixa) para 180 ppm (velocidade alta). O ideal é que tentemos ajustar (convergir) um modelo misto utilizando todos os interceptos e *slopes* aleatórios. Contudo, por diversas razões que fogem ao escopo deste trabalho, isso nem sempre é possível; dessa forma, utilizamos o máximo de interceptos e *slopes* aleatórios possíveis para que o modelo consiga convergir (seja ajustado sem falhas, em outras palavras).

<sup>69</sup> Utilizamos “preditor” como um sinônimo de “variável independente”.

**Tabela 9 – Comparação entre Modelos Aninhados**

<i>Coefficient</i>	<b>Modelo com todos os preditores</b>			<b>Modelo sem Velocidade</b>		
	<i>Estimates</i>	<i>Conf.int (95%)</i>	<i>P-value</i>	<i>Estimates</i>	<i>Conf.int (95%)</i>	<i>P-value</i>
<b>Intercepto</b>	260,92	248,39 – 273,46	<0,001	260,10	248,14 – 272,06	<0,001
<b>Linhas (2)</b>	-7,55	-14,38 – -0,72	0,030	-7,51	-14,34 – -0,68	0,031
<b>Grupo (Surdo)</b>	33,99	17,94 – 50,03	<0,001	33,96	17,91 – 50,00	<0,001
<b>Velocidade (180)</b>	-1,85	-10,09 – 6,40	0,661			

Fonte: Elaborada pelo autor.

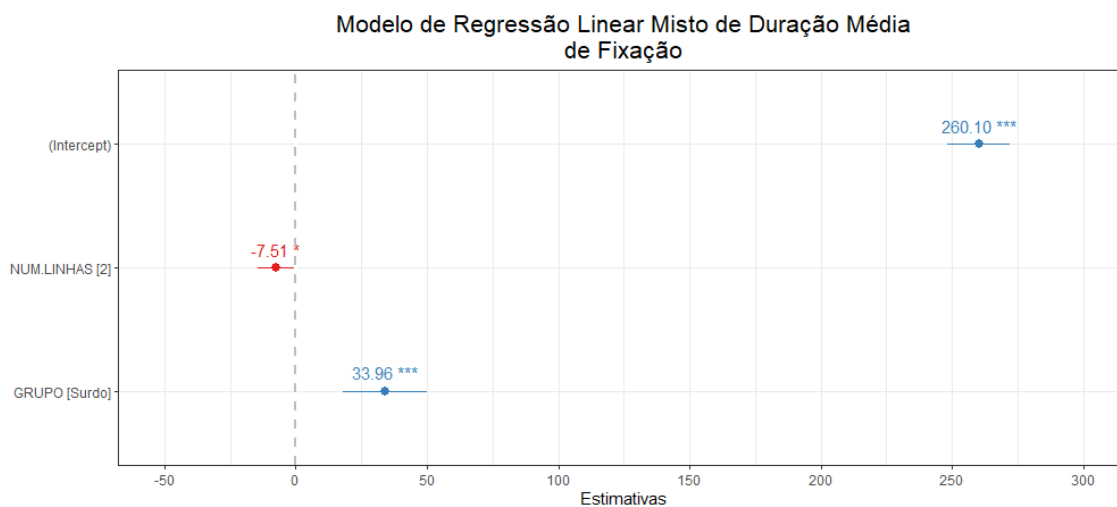
A Tabela 9 mostra que o modelo com todos os preditores (variáveis independentes) e o modelo sem o preditor Velocidade apresentaram valores de  $R^2$  marginal<sup>70</sup> e condicional<sup>71</sup> semelhantes<sup>72</sup>, ou seja, o poder explicativo do modelo considerando apenas os efeitos fixos e aleatórios combinados, respectivamente. Dessa forma, uma ANOVA comparando os dois modelos foi feita. Essa comparação de modelos aninhados mostrou que não há diferença significativa entre os modelos ( $X^2 = 0,1982$ ,  $p = 0,65$ ), sendo assim, aquele com menos preditores é o melhor, pois é o mais simples. Da mesma forma, uma comparação com modelos aninhados – utilizando i) um modelo ajustado com os efeitos fixos de Grupo e Número de Linhas e efeitos mistos e ii) um modelo apenas com intercepto – indicou que o modelo com as variáveis independentes em questão contribui para a explicação da duração média de fixação significativamente ( $X^2 = 19,073$ ,  $p < 0,0001$ ). No Gráfico do modelo ajustado abaixo, observamos a significância do efeito fixo de Número de Linhas ( $\beta = -7,51$ ) e Grupo ( $\beta = 33,96$ ).

<sup>70</sup> Poder explicativo do modelo para aquele fenômeno (medida) sem os efeitos mistos/aleatórios.

<sup>71</sup> Poder explicativo do modelo para aquele fenômeno (medida) com os efeitos mistos/aleatórios.

<sup>72</sup> Modelo com todos os preditores ( $R^2$  marginal/condicional = 0,032/0,097); Modelo sem Velocidade ( $R^2$  marginal/condicional = 0,032/0,094).

### Gráfico 7 – Modelo de Regressão Linear Misto de Duração Média de Fixação



Fonte: Elaborado pelo autor.

O que os coeficientes apresentados no Gráfico 7 dizem, a partir do nosso modelo ajustado, é que os participantes ouvintes e surdos, em média, têm um decréscimo de 7,51 milissegundos quando passam do processamento de uma legenda de 1 linha para 2 linhas levando em consideração as duas velocidades utilizadas nas legendas do experimento (145 e 180 ppm). Além disso, mostra um efeito forte de Grupo, visto que a diferença do valor previsto pelo modelo na fixação média de participantes surdos é de 33,96 milissegundos a mais do que os ouvintes.

**Tabela 10 – Valores previstos pelo modelo para duração média de fixação**

Grupo	Número de Linhas	Duração Média de Fixação (ms)
<b>Ouvinte</b>	1	260,1
	2	252,59
<b>Surdo</b>	1	294,05
	2	286,54

Fonte: Elaborada pelo autor.

O fato de o efeito de Grupo ser bem significativo ( $\beta = 33,96$ ) pode ser explicado pelo fato de um dos grupos ter a língua portuguesa como L1 (primeira língua), no caso os ouvintes, e outro grupo ler em L2 (segunda língua). A esse aspecto, Rayner, Barbara e Pollatsek (2013) explicam que aprendizagens de língua estrangeira ou L2 levam mais tempo para processar textos em L2. O processamento de legendas de 1 e 2 linhas, segundo as estimativas do modelo, também está de acordo com o que há de revisão teórica sobre o assunto. No geral,



legendas de 2 linhas demandam menos custo de processamento no que diz respeito à duração das fixações (D'YDEWALLE; DE BRUYCKER, 2007). Contudo, a variável de Número de Linhas, apesar de significativa, tem efeito apenas moderado na duração da fixação (uma redução de 7,51 milissegundos, de acordo com o modelo) quando comparado com o efeito de Grupo (um acréscimo de 33,96 milissegundos para os participantes surdos, segundo as estimativas). Isso também é notório quando observamos os valores de  $R^2$  marginal e condicional do modelo (0,032 e 0,094). Esses valores apontam que o modelo sem os efeitos mistos explica 3,2% ( $R^2$  marginal) do fenômeno da duração média de fixação. Por outro lado, com a adição de efeitos mistos, o modelo explica 9,4% do fenômeno, um poder explicativo relativamente baixo, ou seja, apesar de Grupo e Número de Linhas influírem na duração média de fixação de surdos e ouvintes conforme apontado por nosso modelo, eles são responsáveis apenas por uma parte pequena do fenômeno.

Tendo em vista as particularidades e as necessidades do trabalho, pensamos que o uso do Modelo de Regressão Linear Misto, a ANOVA e o Teste T Unicaudal são suficientes para afirmar a significância do efeito fixo de Grupo e de Número de Linhas no trabalho, apontando, nesta medida, para a confirmação das nossas hipóteses alternativas sobre: i) as legendas de 1 linha terem um maior custo de processamento do que as legendas de 2 linhas; e ii) as legendas de 145 ppm e 180 ppm não diferirem significativamente. O uso dos três testes combinados com um teste *post-hoc*<sup>73</sup> como o Teste de *Tukey* foi motivado, como já explicitado anteriormente, pelo receio de que uma singularidade dos vídeos utilizados na pesquisa pudesse influenciar um falso positivo (Erro do Tipo II<sup>74</sup>) ou um falso negativo (Erro do tipo I<sup>75</sup>). Ainda é escasso na literatura da área no que diz respeito a como determinados tipos de cena, tendo em vista aspectos cinematográficos de enquadramento, movimentação de câmera, luz etc., para além do gênero fílmico, podem influenciar o processamento de texto e imagem em filmes.

---

<sup>73</sup> Um teste *post-hoc* é um teste realizado após os dados terem sido vistos. Ele é utilizado para determinar onde as diferenças apontadas pelo nível de significância residem.

<sup>74</sup> Erro do Tipo I é quando confirmamos a hipótese alternativa, mas a hipótese nula é a verdadeira. Exemplificando com o nosso trabalho, seria chegar à conclusão de que as legendas de 2 linhas são mais confortáveis para participantes e ouvintes quando, na verdade, não existe diferença ou então as legendas de 1 linha são mais confortáveis.

<sup>75</sup> Erro do Tipo II é quando rejeitamos a hipótese alternativa, mas ela é a verdadeira. Exemplificando com o nosso trabalho, seria chegar à conclusão de que as legendas de 2 linhas não são mais confortáveis para participantes e ouvintes, ou seja, não existe diferença entre legendas de 2 linhas e 1 linha ou então as legendas de 1 linha são mais confortáveis, quando, na verdade, as legendas de 2 linhas realmente demandam um custo menor de processamento.

#### 4.1.2 Número de fixações

O número de fixações por legenda também é uma variável dependente relevante da presente pesquisa. Para leitores de língua materna, é esperado, com base em Rayner, Barbara e Pollatsek (2013), que, além de durações de fixação mais curtas, eles realizem mais fixações proporcionalmente. Isso também foi observado nos estudos de Vieira e Araújo (2017) e Monteiro e Dantas (2017). Dessa forma, hipotetizamos que, naturalmente, os participantes ouvintes teriam em média um número de fixações por linha de legenda menor do que os participantes surdos, por estarem tendo acesso às legendas em língua materna. Além disso, a nossa hipótese alternativa prevê um número maior de fixações em legendas de 2 linhas proporcionalmente, tendo em vista que as legendas de 2 linhas apresentam um material linguístico maior. Em relação à variável de Velocidade, acreditamos que haveria um número ligeiramente maior de fixações nas legendas mais lentas (145 ppm), visto que elas passavam mais tempo em tela, sem que, contudo, fosse uma diferença significativa.

No que tange à relação entre o número de fixações por legenda no experimento e as variáveis independentes da pesquisa (Grupo, Número de Linhas e Velocidade), as Tabelas 11, 12 e 13 abaixo trazem os valores da média de fixações por legenda, além de outras estatísticas descritivas como desvio padrão, erro padrão e mediana.

**Tabela 11 – Número de fixações por legenda por Grupo**

Grupo	Média de Fixação por Legenda	Desvio Padrão	Erro Padrão	Mediana
<b>Ouvinte</b>	5,85	2,72	0,05	6
<b>Surdo</b>	5,41	2,85	0,06	5

Fonte: Elaborada pelo autor

**Tabela 12 – Número de fixações por legenda por Número de Linha**

Número de Linhas	Média de Fixação por Legenda	Desvio Padrão	Erro Padrão	Mediana
<b>1</b>	3,69	1,45	0,03	3
<b>2</b>	7,46	2,49	0,05	7

Fonte: Elaborada pelo autor.

**Tabela 13 – Número de fixações por legenda por Velocidade**

Velocidade	Média de Fixação por Legenda	Desvio Padrão	Erro Padrão	Mediana
<b>145</b>	6,02	3,04	0,06	6
<b>180</b>	5,23	2,42	0,05	5

Fonte: Elaborada pelo autor.

De acordo com as médias acima, percebemos que, assim como era previsto, os ouvintes, leitores de português como L1, fizeram mais fixações (5,85) por legenda do que os participantes surdos (5,41); estes, apesar de proficientes, são leitores de português como L2. Além disso, uma proporção ligeiramente maior de fixações em legendas de 2 linhas foi observada (3,73 por cada linha de legenda; 7,46 levando em conta as 2 linhas) em relação às legendas de 1 linha (3,69), apesar de ser uma diferença pequena. Sobre a variável de Velocidade, observamos que as legendas de 145 ppm (6,02) apresentaram mais fixações do que as de 180 ppm (5,23), sugerindo uma relação entre o número de fixações e o tempo de exposição dessas duas velocidades de legenda.

Sobre os números de fixação por legenda nas condições experimentais do estudo (1L145, 1L180, 2L145 e 2L180), a Tabela 14 abaixo apresenta as médias dessa variável nessas condições.

**Tabela 14 – Número de fixações por condição experimental**

Grupo	Número de Linhas	Velocidade	Média de Fixação Por Legenda
<b>Surdo</b>	1	145	3,85
	1	180	3,30
	2	145	7,74 (3,87) <sup>76</sup>
	2	180	6,53 (3,26)
<b>Ouvinte</b>	1	145	3,87
	1	180	3,68
	2	145	8,15 (4,07)
	2	180	7,20 (3,60)

Fonte: Elaborada pelo autor.

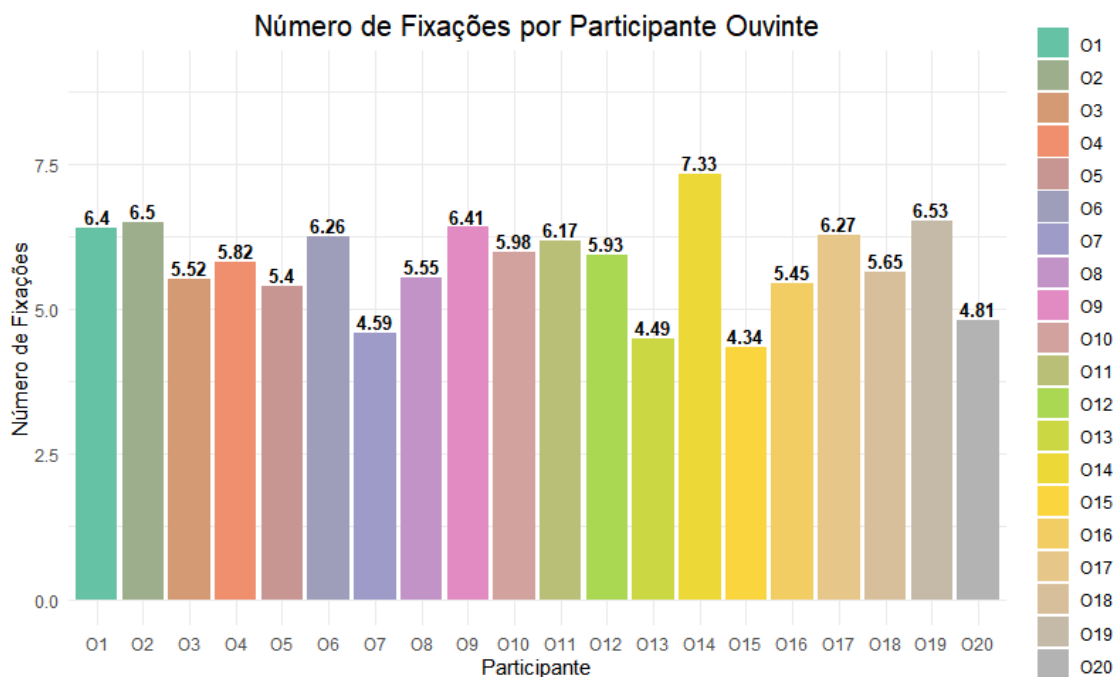
De acordo com os dados acima, percebemos que, entre os participantes surdos, a condição de 2L145 foi a que, proporcionalmente, teve o maior número de fixações por legenda (3,87 por linha; 7,74 ao todo), seguido bem próximo por 1L145 (3,85). Visto que a condição experimental de 1L180 (3,3) apresentou um número ligeiramente maior de fixações por legenda do que 2L180 (3,26 por linha; 6,53 ao todo), percebemos que em 145 ppm os surdos tiveram um desempenho ligeiramente melhor em 2 linhas e em 180 ppm tiveram um desempenho ligeiramente melhor em 1 linha.

Os participantes ouvintes, por sua vez, apresentaram uma tendência parecida. Enquanto na velocidade de 145 ppm a condição experimental de 2 linhas foi a que proporcionou um número maior de fixações por legenda (4,07 por linha; 8,15 ao todo) em relação a 1L145 (3,87), na velocidade de 180 ppm a condição experimental de 1 linha (1L180) apresentou um número de fixações por legenda (3,68) maior do que 2L180 (3,6).

No Gráfico 8, abaixo, podemos ver o número de fixações por participante do grupo de ouvintes.

<sup>76</sup> Os valores entre parênteses na Tabela 14 representam a média do número de fixações por linha nas legendas de 2 linhas.

Gráfico 8 – Número de fixações por participante ouvinte



Fonte: Elaborado pelo autor.

Dentre os 20 participantes ouvintes, o O19 (6,53) e o O14 (7,33) apresentaram os maiores valores de número de fixações, sendo o O14, inclusive, o único a apresentar um número na casa de 7 fixações por legenda. No caso da relação com o tempo médio de duração de fixação apresentado, os mesmos participantes foram o 6º (259,6 milissegundos) e 3º (231,3 milissegundos) com o menor tempo médio apresentado no experimento, o que permite correlacionar, em certa medida, um número alto de fixações por legenda com um bom processamento do texto legendado.

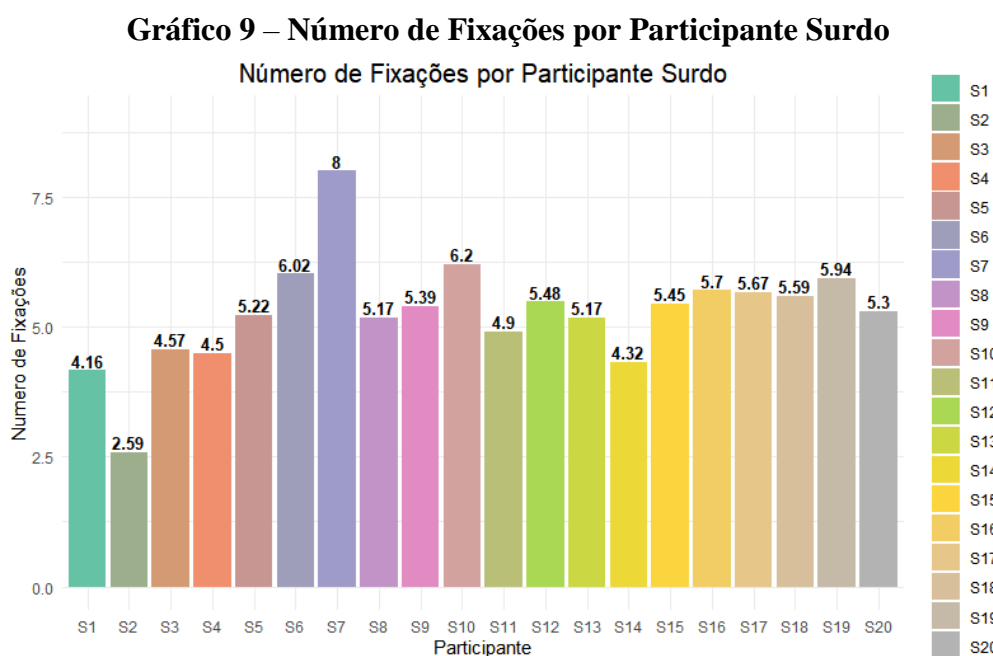
Contudo, essa concepção, proveniente, principalmente, do processamento de textos escritos, não se aplica de maneira diametral quando se trata de textos multimodais como no caso do produto legendado. Os dois menores tempos de duração média de fixação no grupo dos ouvintes, o O7 (206,2 milissegundos) e O5 (208 milissegundos), não tiveram um número de fixações por legenda tão alto assim: o O5 teve 5,4 fixações por legenda (bem próximo à média dos ouvintes, que foi de 5,85) e o O7, por sua vez, apresentou o terceiro menor número de fixações do grupo com 4,59 fixações por legenda.

Ainda nesse sentido, O13 e O15, que apresentaram, respectivamente, o 2º menor e o menor número de fixações por legenda (4,49 e 4,34, na sequência), tiveram um tempo médio de duração de fixação relativamente próximos à média de duração de fixação do grupo de

ouvintes (256,6 e 273,7, respectivamente, sendo o tempo de 263,1 milissegundos a média do grupo), o que não sinaliza um processamento tão custoso assim.

Isso pode significar padrões de processamento de vídeos legendados dentre os participantes do experimento. É possível que os participantes que tenham tido um tempo de processamento menor e, concomitantemente, fizeram menos fixações por legenda i) sejam proficientes e tenham um alto poder de previsibilidade ou ii) tenham utilizado a imagem de forma que não foi necessário fazer tantas fixações na legenda. Ainda é possível que tenham sido participantes que não fizeram regressões ou releituras, que também foram consideradas na medida de número de fixações.

No Gráfico 9, temos as médias de total de fixações por participante surdo.



Fonte: Elaborado pelo autor.

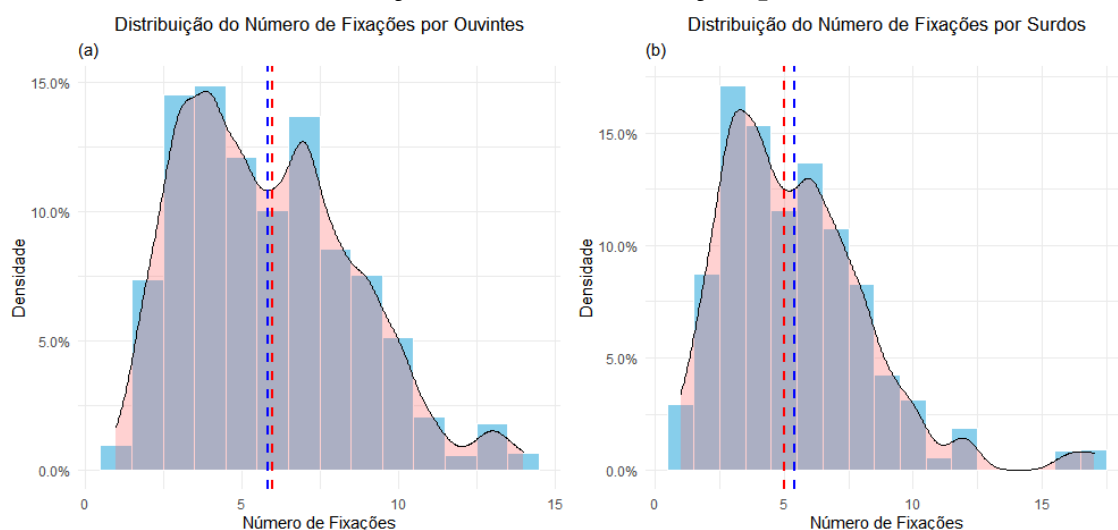
O desempenho do grupo de surdos no que diz respeito a essa variável foi bem parecido com o de ouvintes, com a maior parte dos participantes tendo mais de 5 fixações em média por legenda. No Gráfico 9, podemos constatar que o participante S7 se destacou em relação aos demais com 8 fixações em média por legenda. Comparado com o seu tempo médio de fixação de 286,7 ms, que foi o 6º menor dentre os surdos, podemos inferir que S7 teve um processamento confortável.

S11, por sua vez, apresentou o menor tempo de duração média de fixação entre os surdos (254 milissegundos). Ele não esteve entre os participantes com a maior média de fixações por legenda (5,59), mas foi ainda acima da média de fixações do grupo de surdos

(5,23). Ainda em relação aos participantes com o menor número de fixações, S2 apresentou a maior média de duração de fixação do grupo (359,6 milissegundos) e a última posição na média de fixações por legenda (2,59). Neste caso, a relação esperada entre baixo número de fixações e um alto tempo de duração média de fixação se confirmou.

No que diz respeito à distribuição do número de fixações por grupo, os dois histogramas no Gráfico 10 mostram que as distribuições estiveram bem próximas da normalidade, com as linhas tracejadas azul (média) e vermelha (mediana) praticamente se tocando.

**Gráfico 10 – Distribuição do número de fixações por ouvintes e surdos**

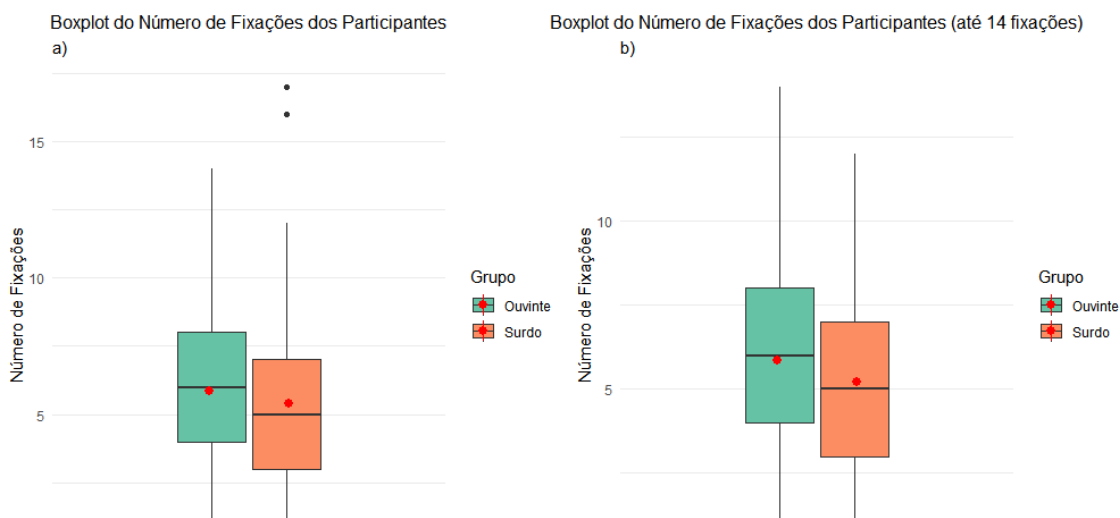


Fonte: Elaborado pelo autor.

No histograma relativo ao grupo de surdos, à direita, percebe-se que há valores altos à direita puxando a média. Há, inclusive, um vazio onde estariam as barras correspondentes ao número de 13, 14, e 15 fixações. Nesse sentido, enquanto o grupo de ouvintes teve um valor máximo de 14 fixações em uma legenda, o grupo de surdos apresenta os valores de 16 e 17 fixações por legenda, indicando uma forte tendência de se tratar de um *outlier*. Além disso, percebemos que o número de fixação por legenda que mais aconteceu entre ouvintes foi 4, enquanto entre os surdos esse número foi o de 3 fixações por legenda.

Por meio do Gráfico 11, abaixo, pudemos constatar que, realmente, esses dados consistiam em *outliers*. Por causa disso, limitamos o número de 14 fixações por legenda para a análise estatística.

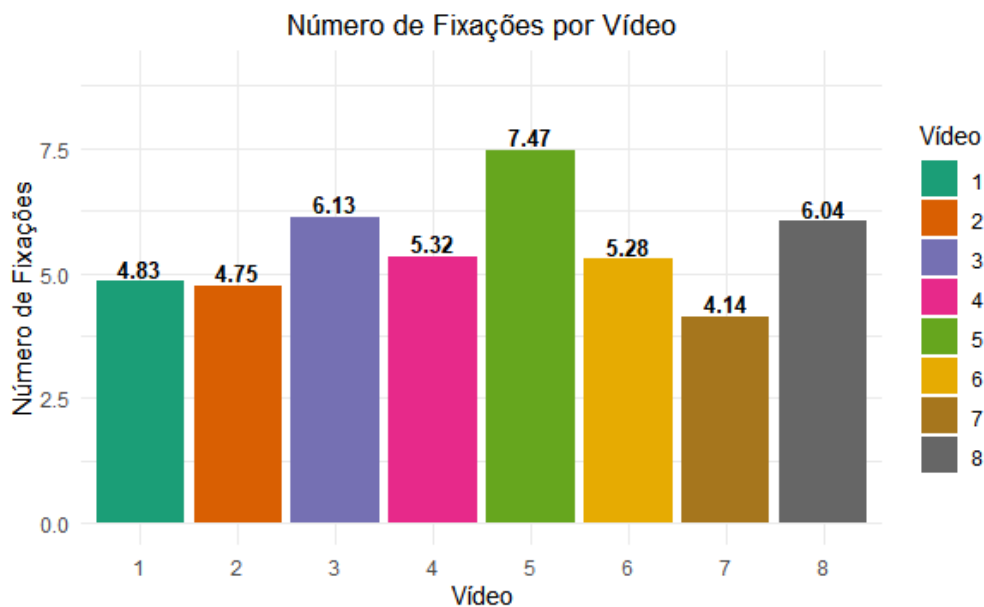
**Gráfico 11 – Boxplot do número de fixações antes e depois da filtragem**



Fonte: Elaborada pelo autor.

Ainda no campo descritivo, o Gráfico 12 mostra como o número de fixações por legenda se distribuiu pelos estímulos do experimento.

**Gráfico 12 – Número de fixações por vídeo**



Fonte: Elaborado pelo autor.

O Gráfico 12 revela algo já esperado. Dada a impossibilidade de se uniformizar o número de caracteres e ainda manipular os cortes de cena, além de outras questões que pautaram a preparação dos estímulos do experimento, um número de fixações distinto por vídeo era de



certa forma esperado. Assim, a legenda do Vídeo 5 ter mais fixações em média (7,47) está em consonância com o seu número total de palavras (23) e de caracteres por palavra (102), ambos os maiores dentre os estímulos.

Um ponto interessante a ser levado em conta, contudo, são os casos dos Vídeos 7 e 8, anteriormente vistos como problemáticos em relação à duração média de fixação. Por um lado, o Vídeo 8, apesar de ter tido o terceiro maior número de fixações por legenda, teve a maior duração média de fixação, conforme já mencionado, dando indício ainda mais forte de que a movimentação da cena no estímulo foi um fator relevante para o comportamento discrepante na variável dependente analisada na seção anterior. O Vídeo 7, por outro lado, apresentou o menor número de fixações por legenda (4,14), algo que pode estar relacionado ao baixo número de palavras em sua legenda (15) e caracteres (69).

A partir dos dados sobre o número de fixações nas legendas do experimento por surdos e ouvintes, rodamos uma Análise de Variância (ANOVA) que apontou diferenças estatisticamente significativas nas variáveis de Grupo ( $f(1, 3.997) = 115,229, p < 0,0001$ ), Número de Linhas ( $f(1, 3.997) = 3824,247, p < 0,0001$ ), Velocidade ( $f(1, 3.997) = 103,181, p < 0,0001$ ). Além da significância nos efeitos principais, a ANOVA também identificou uma interação entre *Grupo:Número de Linhas* ( $f(1, 3.997) = 31,098, p < 0,0001$ ), *Número de Linhas:Velocidade* ( $f(1, 3.997) = 14,028, p < 0,0001$ ) e *Grupo:Número de Linhas:Velocidade* ( $f(1, 3.997) = 9,160, p < 0,002$ ). Não detectou, contudo, diferença significativa a partir da interação *Grupo:Velocidade* ( $f(1, 3.997) = 0,003, p = 0,95$ ).

Além da ANOVA, um modelo de regressão linear misto foi rodado tendo o número total de fixações como variável dependente numérica discreta (**filtrado em 14 fixações**) e os efeitos fixos de Grupo, Número de Linhas e Velocidade. Além disso, o modelo foi ajustado tendo como efeitos aleatórios i) o intercepto de participante e *slope* aleatório de Número de Linhas por participante e ii) o intercepto de vídeo (enquanto item) e *slopes* aleatório de Grupo e Velocidade por vídeo. Esse modelo foi o que convergiu com o maior número de interceptos/*slopes* aleatórios. Ao rodarmos o modelo uma vez com os efeitos fixos enquanto aditivos e uma vez com interação entre os efeitos, percebemos uma significância na interação entre os três níveis de preditores (variáveis independentes) do modelo ( $X^2 = 60,436, p < 0,0001$ ). Além disso, uma comparação deste entre modelos aninhados (o modelo com interação supracitado e um modelo apenas com intercepto) revelou significância estatística ( $X^2 = 169,44, p < 0,0001$ ). Esse procedimento apontou os efeitos fixos utilizados no modelo como preditores para a medida número de fixações por legenda. Além disso, o modelo ainda se mostrou com um alto poder explicativo, considerando seus efeitos aleatórios conforme o **R<sup>2</sup> condicional** de

0,715; ou seja, as variáveis de Grupo, Número de Linhas e Velocidade explicam 71,5% do número de fixações por legenda.

Abaixo, na Tabela 15, podemos ver as estimativas do modelo ajustado.

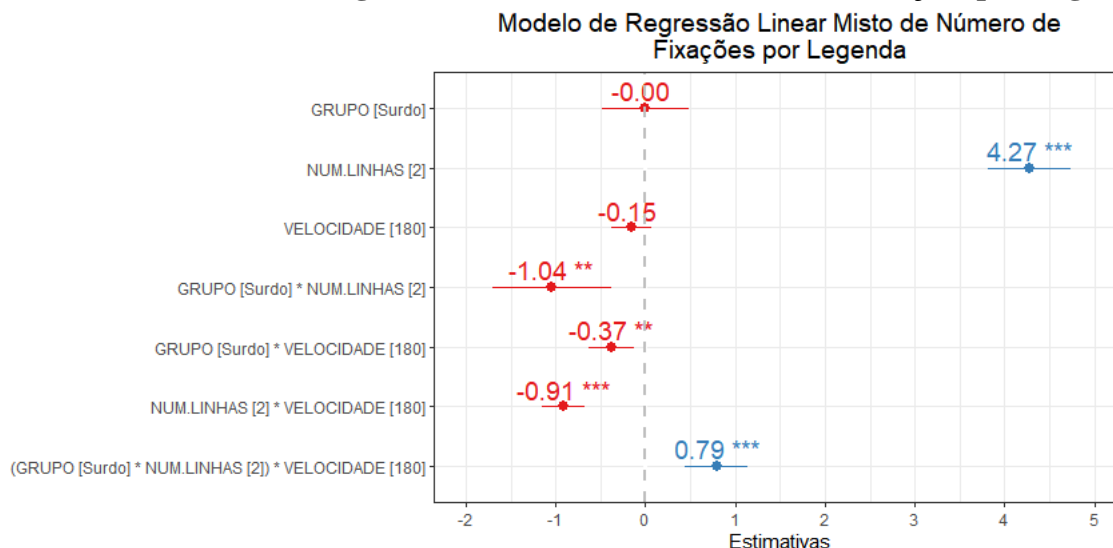
**Tabela 15 – Modelo de Regressão Linear Misto de Número de Fixações por Legenda**

Coefficient	Estimates	Conf.int (95%)	P-value
<b>Intercepto</b>	3,65	2,98 – 4,32	<0,001
<b>Grupo(Surdo)</b>	-0,00	-0,48 – 0,48	0,996
<b>Número de linhas(2)</b>	4,27	3,81 – 4,74	<0,001
<b>Velocidade(180)</b>	-0,15	-0,37 – 0,08	0,194
<b>Grupo(Surdo)*Número de Linhas(2)</b>	-1,04	-1,70 – -0,38	0,002
<b>Grupo(Surdo)*Velocidade(180)</b>	-0,37	-0,63 – -0,12	0,004
<b>Número de Linhas(2)*Velocidade(180)</b>	-0,92	-1,16 – -0,67	<0,0
<b>Velocidade(180)</b>	0,80	0,44 – 1,15	<0001
<b>R2 Marginal/R2 Condicional</b>	0,484/0,715		

Fonte: Elaborada pelo autor.

O modelo apresentado na Tabela 15 quando plotado pode ser visualizado da seguinte forma:

**Gráfico 13 – Modelo de Regressão Linear Misto de Número de Fixações por Legenda**



Fonte: Elaborado pelo autor.

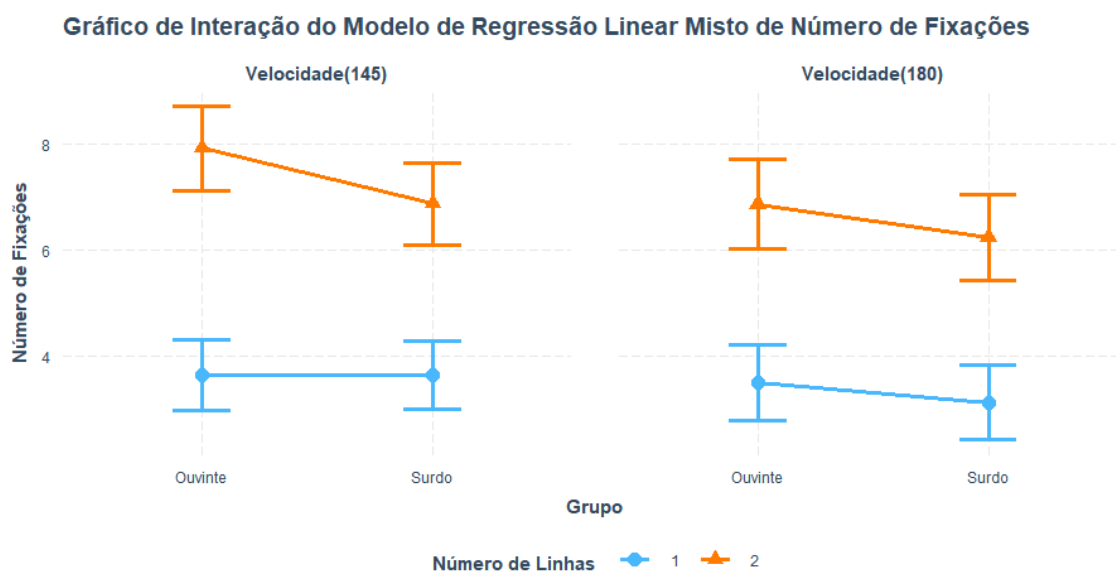
A Tabela 15 e o Gráfico 13 acima esclarecem que, isoladamente, os efeitos fixos de Grupo ( $\beta = 0,00$ ,  $p = 0,10$ ) e de Velocidade ( $\beta = -0,15$ ,  $p = 0,194$ ) não atingiram o nível de significância estatística, diferentemente do efeito de Número de Linhas na velocidade de

145 ppm por participantes ouvintes ( $\beta = 4,27$ ,  $p < 0,0001$ ), sendo este, inclusive, mais do que o dobro do efeito de 1 linha para os participantes ouvintes na mesma velocidade, conforme marcado no valor do intercepto ( $\beta = 3,65$ ). Contudo, esse é um modelo com várias interações entre os três preditores aplicados e, dessa maneira, a visualização desses efeitos se torna complicada. Por isso, iremos nos ater a explicar com mais detalhes como o número de fixações do grupo de surdos foi influenciado por essas interações.

Apesar de o modelo apontar a não significância da variável Grupo, podemos ver que, quando os surdos processaram legendas de 2 linhas na velocidade de 145 ppm, o seu número de fixações caiu em comparação ao de ouvintes ( $\beta = 4,27$ ,  $p = 0,002$ ). Além disso, apesar de a variável Velocidade não ter se mostrado significativa isoladamente, ela foi significativa quando em interação com o grupo de surdos ( $\beta = -0,37$ ,  $p = 0,004$ ). No que diz respeito à interação Número de Linhas (2) e Velocidade (180), ela foi significativa, tendo esse coeficiente negativo ( $\beta = -0,92$ ,  $p < 0,001$ ) para os dois grupos, mas com uma pequena correção na interação *Grupo(Surdo):Número de Linhas(2):Velocidade(180)* ( $\beta = -0,80$ ,  $p < 0,001$ ).

Essa interação pode ser visualizada no Gráfico 14 abaixo:

**Gráfico 14 – Modelo de Regressão Linear Misto de Número de Fixações por Legenda com Interação**



Fonte: Elaborado pelo autor.

No Gráfico 14, observa-se que o número de fixações por legenda permanece igual ou com pouca variação entre velocidades (145 e 180) nos grupos de ouvintes e surdos

processando legendas de 1 linha. Contudo, legendas de 2 linhas fazem com que haja uma variação maior entre grupos nas velocidades de legenda, especialmente na legenda de 145 ppm.

Tendo em vista os efeitos simples e as interações sugeridas pelo modelo, a Tabela 16 abaixo mostra as estimativas de número de fixações (média por legenda) de acordo com as variáveis independentes aplicadas no experimento.

**Tabela 16 – Valores Previstos pelo Modelo de Regressão Linear Misto de Número de Fixações**

Grupo	Linha	Velocidade	Média de Fixação
<b>Surdo</b>	1	145	3,65
	1	180	3,13
	2	145	6,88
	2	180	6,24
<b>Ouvinte</b>	1	145	3,65
	1	180	3,50
	2	145	7,92
	2	180	6,85

Fonte: Elaborada pelo autor.

A Tabela 16 aponta que, em praticamente todas as condições experimentais, as estimativas dão conta de mais fixações dos ouvintes na legenda, com exceção das legendas de 145 ppm de 1 linha. Além disso, podemos dizer que as legendas de 145 ppm proporcionaram mais fixações ao longo das condições experimentais, como era esperado, visto que passam mais tempo em tela. Em relação ao grupo de surdos, é possível dizer que a estimativa do modelo sugere que a condição experimental com mais fixações proporcionalmente é a de 1 linha e 145 ppm (3,65 fixações por linha de legenda) enquanto que para o grupo de ouvintes a legenda de 2 linhas e 145 ppm é estimada como a que oferece mais fixações por linha de legenda (3,96 fixações por linha de legenda), possivelmente por causa do tempo maior de exposição em tela e consequentes regressões e releituras realizadas pelos participantes. Um número maior de fixações por ouvintes também já era esperado, visto que usuários de L1 (ouvintes) fazem mais fixações – e mais curtas – do que usuários de L2 (surdos) (RAYNER; BARBARA; POLLATSEK, 2013). Contudo, há que se levar em conta também que as legendas mais lentas (145 ppm) e de 2 linhas passam mais tempo em tela, o que pode relacionar essa medida com as medidas de releitura e de regressão, que serão apresentadas a seguir; ou seja, quanto mais tempo em tela (condição experimental 2L145), mais os leitores proficientes, como os ouvintes

participantes do estudo, podem realizar fixação na legenda, o que também será levado em conta na seção de discussão desses resultados.

#### 4.1.3 Regressão

Como vimos anteriormente, a variável resposta de regressão diz respeito aos movimentos oculares à esquerda durante a leitura da legenda e está relacionada a possíveis dificuldades na leitura. Ao contrário da leitura de textos escritos, contudo, o processamento de legendas não permite que o leitor tenha tanto controle sobre seu ritmo de leitura, sendo este controlado pelo filme/legendista em torno dos parâmetros utilizados na confecção da legenda. Com base nesse aspecto, os resultados aqui vistos precisam ser circunstanciados em relação a essa característica dinâmica do texto legendado.

Essa medida foi analisada em duas perspectivas: i) duração média da regressão (em milissegundos); e ii) proporção de casos de regressão. Como cada uma dessas perspectivas pressupõe testes estatísticos e maneiras distintas de apresentar os dados, iremos mostrá-los de maneira seccionada, iniciando com a duração.

Após a análise de inspeção da movimentação ocular dos participantes, identificamos 371 casos de regressão. Diferentemente da medida de duração da fixação, optamos por não limitar, inicialmente, a duração da regressão tida como válida nessa etapa, mas fazê-lo apenas durante a busca por *outliers*, posteriormente. Isso ocorreu pela quantidade menor de observações dessa medida, visto que a ocorrência da regressão não é normalmente controlada pelo pesquisador, mas fruto do contato do leitor com o texto e possíveis dificuldades durante a leitura. Ainda sem a filtragem de *outliers*, a estatística descritiva dividida por variável independente pode ser vista nas Tabelas 17, 18 e 19 abaixo.

**Tabela 17 – Duração média de regressão por Grupo**

Grupo	Duração de Regressão (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>Ouvinte</b>	282,82	134,50	9,44	250
<b>Surdo</b>	290,60	125,59	9,68	250

Fonte: Elaborada pelo autor.

**Tabela 18 – Duração média de regressão por Número de Linhas**

Número de Linhas	Duração de Regressão (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>1</b>	300,77	141,92	10,72	250
<b>2</b>	273,47	118,12	8,43	233

Fonte: Elaborada pelo autor.

**Tabela 19 – Duração média de regressão por Velocidade**

Velocidade	Duração de Regressão (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>145</b>	279,05	121,18	8,34	233
<b>180</b>	295,96	141,51	11,18	250

Fonte: Elaborada pelo autor.

De uma forma geral, é possível perceber as que as médias por Grupo e Número de Linhas (maior em surdos e em legendas de 1 linha) parecem apontar para uma confirmação de nossa hipótese alternativa, de que legendas de 2 promovem um melhor processamento do vídeo legendado. Sobre a hipótese para a variável de Velocidade, contudo, uma média substancialmente maior nas legendas de 180 ppm sugere que, nessa medida, há uma diferença entre os níveis dessa variável.

**Tabela 20 – Duração Média de Regressão por Condição Experimental**

Grupo	Condição Experimental	Duração de Regressão (ms)
<b>Surdo</b>	1L145	284,22
	1L180	323,10
	2L145	271,57
	2L180	297,04
<b>Ouvinte</b>	1L145	296,16
	1L180	308,31
	2L145	267,54
	2L180	261,71

Fonte: Elaborada pelo autor.

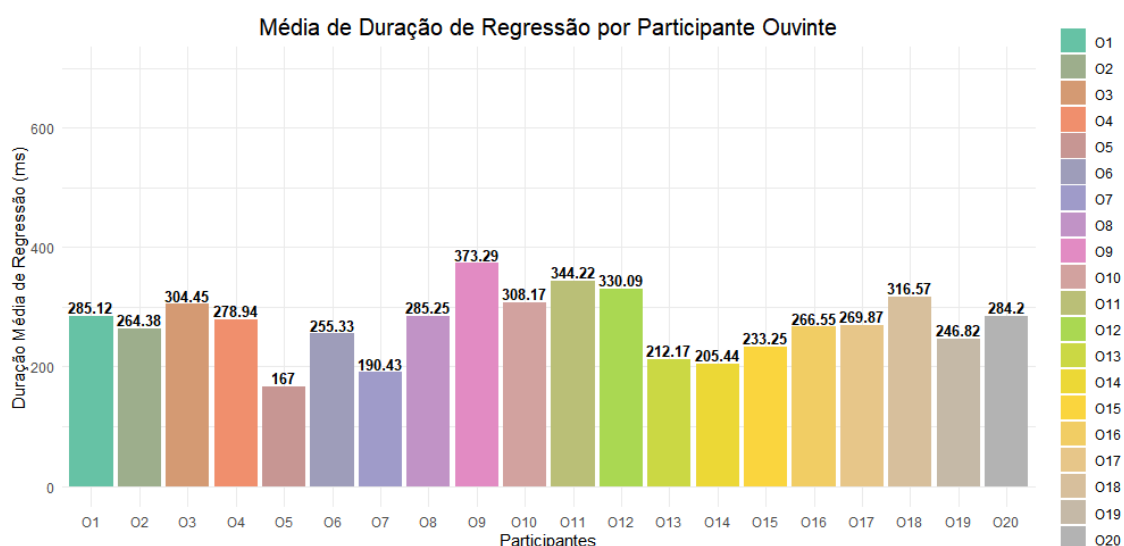
De acordo com a Tabela 20, a condição experimental em que os surdos tiveram o menor custo de processamento em termos de regressão foi a 2L145 (271,57 ms), seguida da 1L145 (284,22 ms). Podemos ver, então, que as duas condições experimentais em que os

participantes surdos tiveram um desempenho melhor em termos de regressão foram com a velocidade de 145 ppm. Sendo assim, obviamente, as condições de maior custo de processamento para regressão foram com a velocidade de 180 ppm, principalmente 1L180, que apresentou a maior média de tempo de regressão entre os participantes desse grupo, 323,1 ms, enquanto 2L180 teve como média o tempo de 297,04 ms.

Os ouvintes, por sua vez, apresentaram um resultado um pouco distinto dos surdos. Os tempos médios de regressão mais baixos e, por essa razão, menos custosos, foram os de 2 linhas (2L145 e 2L180). Enquanto 2L145 teve como tempo médio de regressão 267,54 ms, 2L180 teve média de 261,71 ms, sendo o menor tempo médio entre todas as condições experimentais para os ouvintes. As condições experimentais de 1L, dessa forma, apresentaram o maior tempo médio para essa medida com 296,16 ms (1L145) e 308,31 ms (1L180). A condição experimental de 1L180, aliás, como a de maior tempo médio no estudo foi algo comum nos dois grupos.

No que concerne às médias por participante, o Gráfico 15 abaixo traz as médias dos participantes ouvintes.

**Gráfico 15 – Média de duração de regressão do grupo de ouvintes**



Fonte: Elaborada pelo autor.

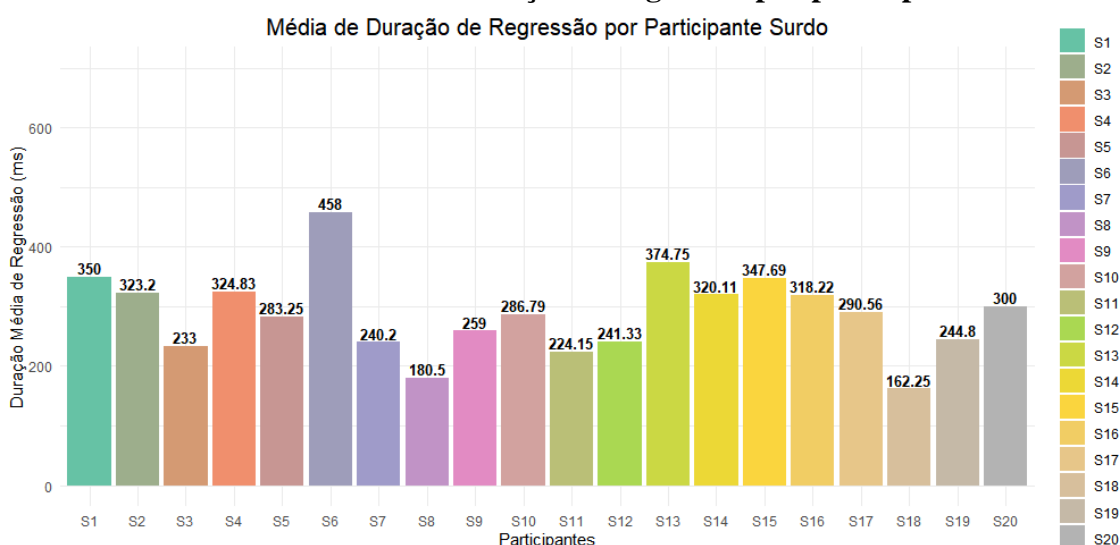
Como os participantes realizaram regressões em proporções distintas, há uma variação ainda maior quando examinamos a duração da regressão com as medidas de duração média de fixação e total de fixações por legenda vistas anteriormente. Observamos uma grande

variação entre a menor (participante O7 com 167 milissegundos) e a maior duração média de regressão (participante O9 com 373,29 milissegundos).

É interessante observar que as menores durações médias de regressão registradas (participantes O5 e O7 com 167 e 190,43 ms, respectivamente) são dos dois participantes que tiveram as menores médias de duração de fixação (208 e 206,2 milissegundos, respectivamente), mostrando que essas medidas estão relacionadas ao custo de processamento também na legendagem, apesar das circunstâncias já mencionadas em relação à especificidade da medida de regressão em textos dinâmicos. Ainda em relação a isso, o participante O9, que apresentou a maior média de duração de regressão dentre os ouvintes, também esteve entre os que apresentaram as maiores médias de duração de fixação (a 6ª maior com 279,3 milissegundos).

Em relação aos participantes surdos, o Gráfico 16 mostra suas médias de duração de regressão.

**Gráfico 16 – Média de duração de regressão por participante surdo**



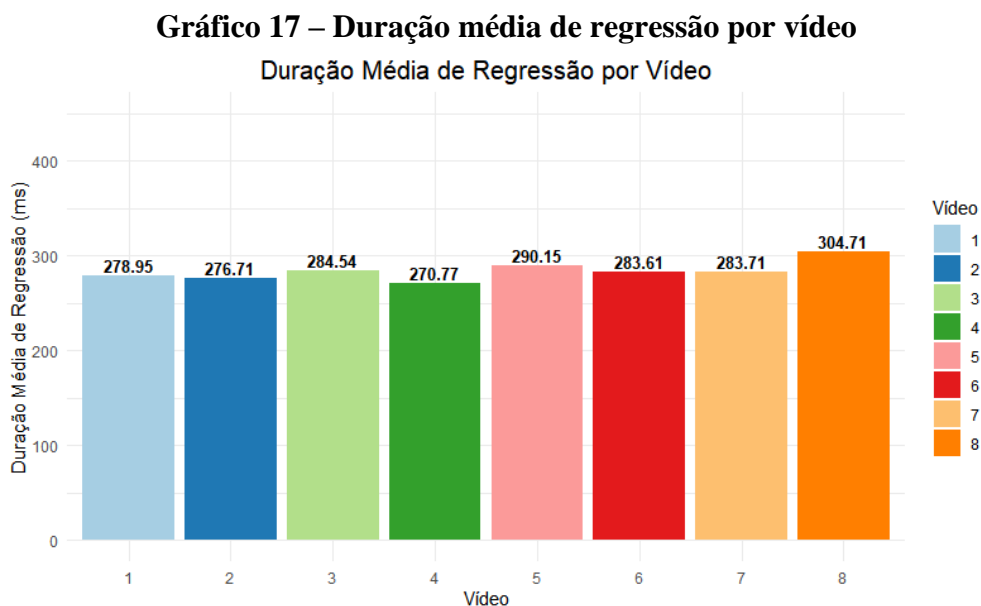
Fonte: Elaborado pelo autor.

Conforme já explicitado em relação à heterogeneidade esperada do grupo de surdos, a variação de médias de regressão é ainda maior, sendo a menor média registrada a de S18 (162,25 ms) e a maior de S6 (458,33 ms). Na menor duração média, é interessante notar que S18 também teve a menor duração média de fixação dentre os surdos (254 milissegundos). No que diz respeito à maior duração média de regressão apresentada, S6 não esteve entre os de maior duração média de fixação, visto que teve apenas a 11ª maior média na medida (303,7 ms). Contudo, o mesmo participante realizou apenas dois movimentos de regressão no experimento,



um número pequeno de ocorrências que não permite conclusões mais definitivas sobre processamento de S6 em relação a essa medida.

O Gráfico 17 traz as médias de fixação por vídeo utilizado como item experimental no estudo.

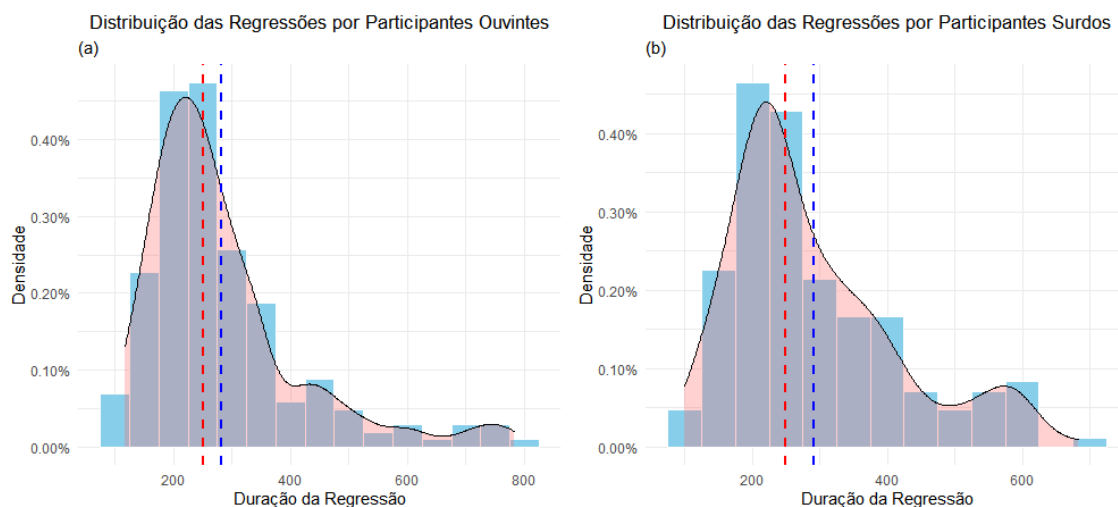


Fonte: Elaborado pelo autor.

A partir do Gráfico 17, é possível ver que, assim como na variável de duração média de fixação, há uma disparidade entre as médias de regressão entre os vídeos experimentais. Mais uma vez, inclusive, o Vídeo 8 (304,71 ms) teve um tempo médio maior do que os outros, apresentando uma diferença substancial para o 2º vídeo com o maior custo de processamento em termos de regressão, o Vídeo 5 (290,15 ms). Em seguida, em ordem decrescente, aparecem os Vídeos 3 (284,54 ms), 7 (283,71 ms), 6 (283,61 ms), 2 (278,95 ms) e 1 (278,95 ms). Em último lugar, com o menor tempo médio de regressão entre os vídeos e até um pouco distante dos outros, está o Vídeo 4 (270,77 ms). Essa diferença entre itens experimentais sobre a medida analisada foi o que fez, novamente, com que checássemos se se tratava de uma variável de confusão no estudo, o que será apresentado mais adiante por ocasião da descrição dos resultados da análise estatística.

No Gráfico 18 está a distribuição de ocorrências de regressão por ambos os grupos de participantes em relação à sua duração em milissegundos.

### Gráfico 18 – Distribuição das regressões de ouvintes e surdos

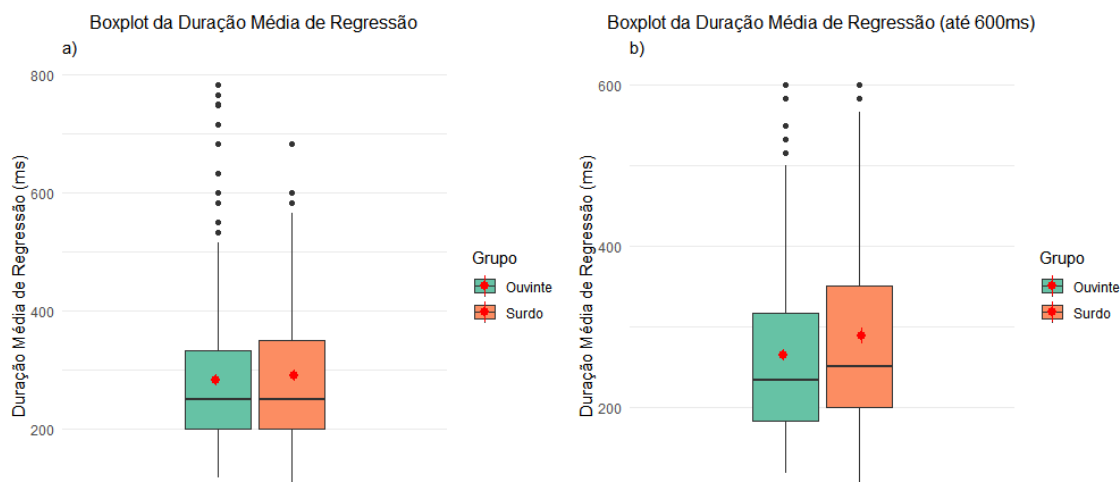


Fonte: Elaborado pelo autor.

Pelos histogramas acima é possível observar que o comportamento dos dois grupos é bem similar na distribuição de regressões em relação à sua duração. As linhas que representam a mediana (vermelha) e a média (azul) são bem próximas uma da outra em ambos os grupos, conforme a Tabela 17. Alguns poucos casos de duração maior do que 600 milissegundos, contudo, fazem com que a cauda à direita fique bastante alongada e, dessa forma, seja uma distribuição mais próxima da não normal.

Devido a essa distribuição, fizemos novamente uma análise por *boxplot* em busca de possíveis *outliers* para preparar os dados para a análise estatística dos dados, o que pode ser visto abaixo no Gráfico 19.

### Gráfico 19 – Boxplots de duração média das regressões antes e depois da filtragem



Fonte: Elaborado pelo autor.

A tarefa de excluir os dados foi pautada por diminuir as observações que estivessem fora do intervalo interquartil na visualização do *boxplot*. Ao mesmo tempo, considerou-se que o menor número possível de dados deveria ser descartado dada a pouca quantidade de observações. Sendo assim, chegamos ao valor de 600 milissegundos, o qual não eliminou todos os *outliers*, mas limpou consideravelmente o número destes. Após a filtragem, restaram 362 casos (ao invés de 371, como anteriormente), a partir dos dados de antes e depois dessa filtragem.

Com base nessa filtragem, é válido novamente apresentar os dados da estatística descritiva em relação aos tempos de duração média da regressão para fins de visualização dos resultados da ANOVA (vide Tabelas 21, 22 e 23).

**Tabela 21 – Duração média de regressão por Grupo (após filtragem)**

Grupo	Duração de Regressão (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>Ouvinte</b>	264,7	101,87	7,29	233
<b>Surdo</b>	288,25	122,2	9,45	250

Fonte: Elaborada pelo autor.

**Tabela 22 – Duração média de regressão por Número de Linhas (após filtragem)**

Número de Linhas	Duração de Regressão (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>1</b>	283,95	117,53	9,06	250
<b>2</b>	268,3	107,08	7,68	233

Fonte: Elaborada pelo autor.

**Tabela 23 – Duração média de regressão por Velocidade (após filtragem)**

Velocidade	Duração De Regressão (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>145</b>	270,53	105,36	7,32	233
<b>180</b>	282,30	120,69	9,69	250

Fonte: Elaborada pelo autor.

Com os dados apresentados acima, rodamos uma Análise de Variância (ANOVA) tendo como variáveis categóricas independentes Grupo, Número de Linhas, Velocidade e, por último, Vídeo enquanto possível variável de confusão. No que diz respeito a esta, o resultado advindo da ANOVA mostrou que a diferença entre vídeos não foi significativa ( $f(7, 351) =$

0,444,  $p = 0,87$ ), o que foi confirmado na comparação entre pares múltiplos através do Teste de *Tukey*. No que tange às outras variáveis, a categoria Grupo foi significativa ( $f(1, 351) = 3,963$ ,  $p = 0,0473$ ), enquanto Número de Linhas ( $f(1, 351) = 0,1829$ ,  $p = 0,1829$ ) e Velocidade ( $f(1, 351) = 0,967$ ,  $p = 0,3262$ ) não se mostraram significativas. Por fim, não foi detectada nenhuma interação no teste, mas a interação *Grupo:Velocidade* foi marginalmente significativa ( $f(1, 347) = 3,367$ ,  $p = 0,0674$ ). Ao examinarmos as comparações múltiplas entre pares na interação entre as duas variáveis, observamos que na interação *Surdo:180-Ouvinte:180* houve significância (0,0349), indicando que, nessa velocidade, os surdos têm uma média maior do que os ouvintes em 48,84 ms nas legendas de 180 ppm.

Concomitantemente à ANOVA, rodamos, primeiramente, um modelo linear múltiplo (sem efeitos mistos) tendo como efeito fixo as variáveis de Grupo, Número de Linhas, Velocidade e Vídeo. Nossa intenção, com esse primeiro modelo, era apenas de confirmar a não influência da variável de Vídeo no teste estatístico, assim como foi feito com a ANOVA acima. Dito isto, como a variável Vídeo não se mostrou significativa no resultado do modelo e nem no Teste de *Tukey* sobre as comparações entre pares múltiplos, percebemos que Vídeo, no caso da variável de regressão, não se constitui em uma variável de confusão. Por essa razão, não precisaria ter nenhum dos seus itens retirado.

Após essa operação, rodamos outro modelo, um linear múltiplo misto tendo como efeito fixo Grupo, Número de Linhas e Velocidade e, como efeito aleatório, intercepto de participante e *slope* aleatório de Grupo por participante, sendo que este foi o único modelo com efeitos mistos e todas as variáveis a convergir. Os resultados do modelo mostraram que nenhuma das variáveis foi significante: Grupo ( $\beta = 22,13$ ,  $p = 0,201$ ), Número de Linhas ( $\beta = -14,73$ ,  $p = 0,207$ ) e Velocidade ( $\beta = 11,70$ ,  $p = 0,315$ ).

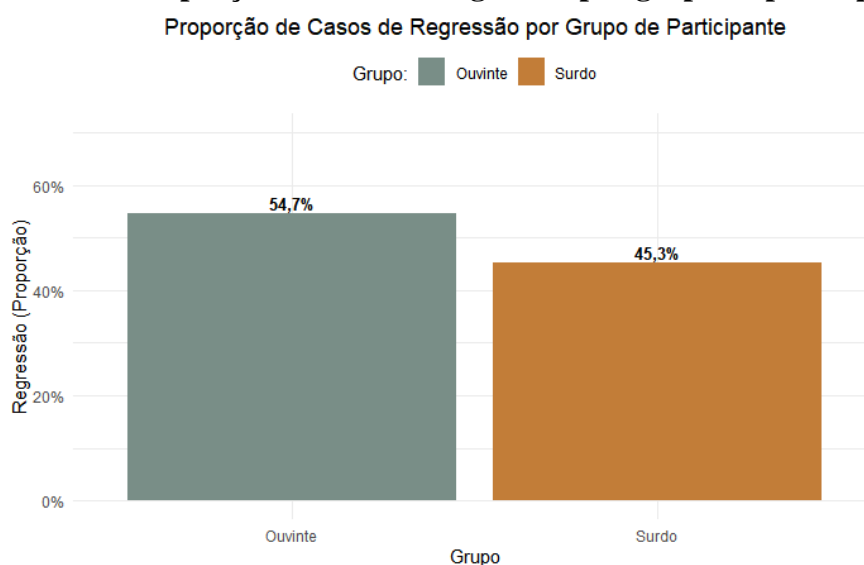
Além da significância apenas no nível de Grupo reportada pela ANOVA, não foi possível prever valores, dada a não significância dos efeitos fixos testados. Uma coisa a se considerar, contudo, é a baixa quantidade de dados disponíveis para esse momento da análise: 362 ocorrências, apenas. Assim, essas diferenças palpáveis entre os níveis das variáveis independentes podem não ter sido significativas no modelo dada a baixa quantidade de dados coletados dessa medida (duração média de regressão). Essa questão já havia sido apontada anteriormente e se trata, primariamente, de uma característica específica desse tipo de dado; ou seja, o fato de não ter sido objetivo do estudo “forçar” o participante a fazer regressões durante o processamento da legenda faz com que a quantidade de ocorrência desse fenômeno seja imprevisível.

Dito isto, passemos agora à segunda parte de descrição dos resultados sobre regressão, agora no que diz respeito à quantidade de movimentos regressivos realizados – além da probabilidade de sua ocorrência – de acordo com as variáveis independentes estudadas.

Primeiramente, examinemos a proporção de ocorrências de regressão a partir de nossas variáveis independentes: Grupo, Número de Linhas e Velocidade. Nesse momento, apenas as ocorrências de regressão importam (mais à frente, as legendas com ocorrências de regressão e não regressão também irão ser levadas em conta). Além disso, diferentemente da ANOVA e do Modelo de Regressão Linear, o Modelo de Regressão Logística, a partir do qual serão analisadas as ocorrências ou não de regressão observadas, não necessita que os dados trabalhados sejam paramétricos – justamente por investigar proporções distintas. Sendo assim, nenhuma observação das 371 ocorrências de regressão foi descartada.

No que diz respeito à variável por Grupo, os ouvintes realizaram 203 regressões (54,7% das ocorrências) em comparação aos surdos, que realizaram 168 regressões (45,3% das ocorrências), conforme o Gráfico 20.

**Gráfico 20 – Proporção de casos de regressão por grupo de participante**



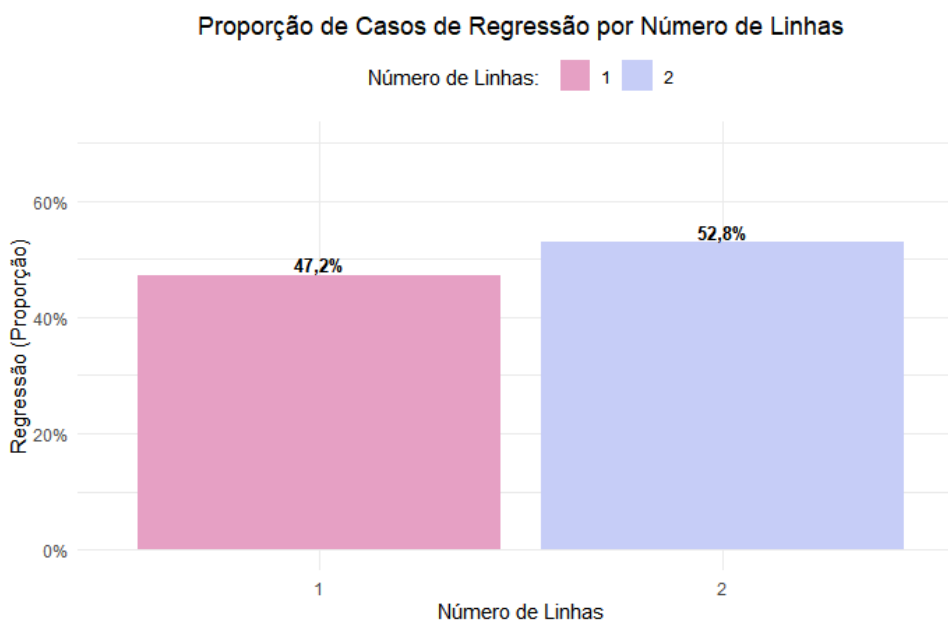
Fonte: Elaborado pelo autor.

À primeira vista, pode parecer simplesmente que os ouvintes tiveram mais dificuldade em relação a essa medida do que os surdos. Contudo, também há a necessidade de se comparar esses números com o total de fixações realizadas pelos dois grupos como um todo, visto que os ouvintes realizaram mais fixações do que os surdos no experimento. Nesse sentido, os ouvintes realizaram mais fixações (2.159) do que os surdos (1.909), em um universo de 4.068 fixações observadas, sendo que as regressões (203) corresponderam a 9,40% das fixações dos

ouvintes. Em relação aos surdos, as regressões (168) corresponderam a 8,80% de suas fixações. Ou seja, os ouvintes realizaram mais regressões tanto no universo de ocorrências de regressão do experimento quanto proporcionalmente ao número de fixações realizadas. A porcentagem de movimentos de regressão em ouvintes e surdos do experimento foi bem próxima, mas ainda abaixo do número indicado por Rayner, Barbara e Pollatsek (2013), que seria de 10 a 15% do total de fixações realizadas.

Em relação à variável Número de Linhas, as legendas de 1 linha tiveram 175 regressões (47,2% das ocorrências), enquanto as legendas de 2 linhas apresentaram 196 casos de regressão (52,8 % dos casos), conforme pode ser visualizado no Gráfico 21.

**Gráfico 21 – Proporção de casos de regressão por número de linhas**

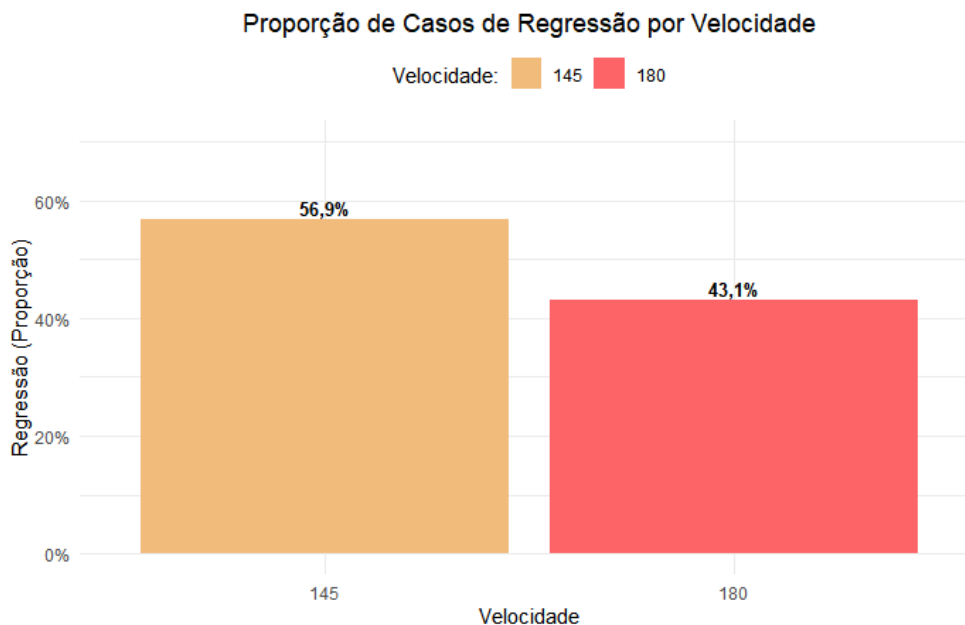


Fonte: Elaborado pelo autor.

Assim como apresentamos na proporção de casos de regressão por Grupo, investigamos a proporção de regressões em relação às fixações realizadas em legendas de 1 linha e 2 linhas. Nesse sentido, em um universo de 1.966 fixações realizadas em legendas de 1 linha, as 175 regressões corresponderam a 8,90% das fixações realizadas. Em relação às legendas de 2 linhas, as 196 ocorrências de regressão representaram 9,32% das 2.102 fixações realizadas nesse nível da variável, ratificando um número proporcional maior de regressões em legendas de 2 linhas, tanto levando em consideração o total de regressões do experimento, quanto proporcionalmente o número de fixações realizadas.

No que diz respeito à variável Velocidade, as legendas de 145 ppm tiveram 211 casos de regressão (56,9% dos casos), enquanto as legendas de 180 ppm tiveram 160 ocorrências (43,1% dos casos), conforme pode ser visualizado pelo Gráfico 22.

**Gráfico 22 – Proporção de casos de regressão por velocidade**



Fonte: Elaborado pelo autor.

Estabelecendo uma relação entre o número de fixações realizadas e o número de regressões, observamos que, dentre as 2.148 fixações em legendas de 145 ppm, os casos de regressão (211) representaram 9,82% das fixações, enquanto dentre as 1.920 fixações em legendas de 180 ppm, os casos de regressão (160) representaram 8,33% das ocorrências.

A Tabela 24 traz as frequências absolutas e relativas (porcentagem) das ocorrências de regressão em todas as condições experimentais do trabalho.

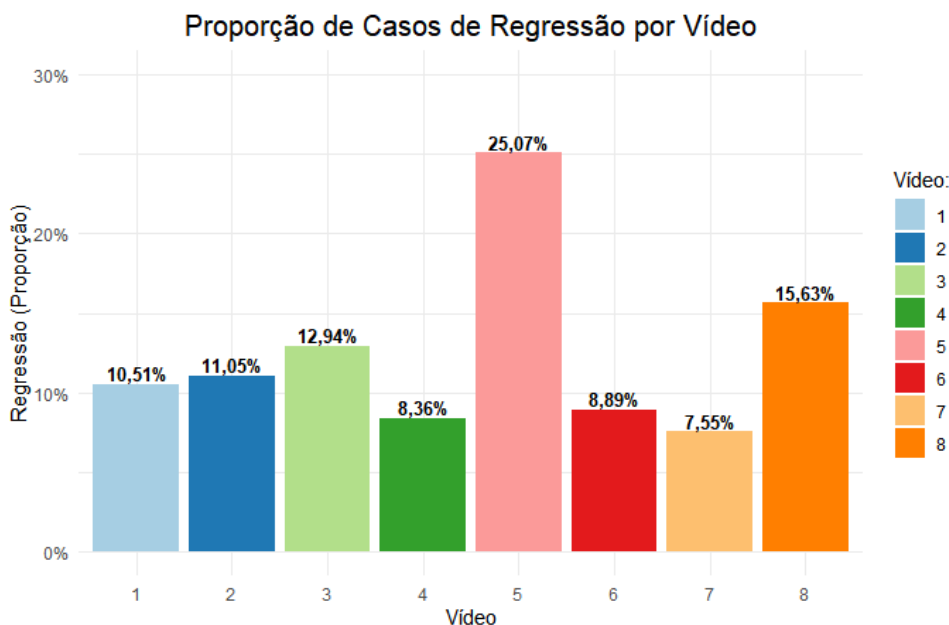
**Tabela 24 – Ocorrências de regressão por condição experimental**

Grupo	Linha	Velocidade	Frequência	
			Absoluta	Relativa
<b>Ouvinte</b>	1	145	50	13,48
	1	180	47	12,67
	2	145	64	17,25
	2	180	42	11,32
<b>Surdo</b>	1	145	48	12,94
	1	180	30	8,08
	2	145	49	13,21
	2	180	41	11,05

Fonte: Elaborada pelo autor.

Segundo as informações dispostas acima, os ouvintes realizaram mais regressões em legendas com velocidade de 145 ppm: 1L145 (50 casos ou 13,48% do total) e 2L145 (64 casos ou 17,25% do total). De maneira oposta, as legendas com velocidade de 180 ppm apresentaram menos casos de regressão nesse grupo de participantes. Enquanto 1L180 teve 47 casos de regressão (12,67% do total), 2L180 teve o menor número de regressões com 42 (11,32%). Em relação aos surdos, os resultados foram bem semelhantes. Os surdos também apresentaram o maior número de regressões nas condições experimentais de 1L145 (48 casos ou 12,94% do total) e 2L145 (49 casos ou 13,21%). Essa pequena diferença entre legendas de 1 linha e 2 linhas entre surdos sugere que o número de linhas nas condições experimentais de 145 ppm influenciou mais no comportamento de ouvintes do que de surdos. Em relação às condições experimentais de 180 ppm, contudo, o resultado nos surdos foi um pouco diferente dos ouvintes: 1L180 teve menos regressões (30 casos ou 8,08% do total) do que 2L180 (41 casos ou 11,05% do total).



**Gráfico 23 – Proporção de casos de regressão por vídeo**

Fonte: Elaborado pelo autor.

Como vemos no Gráfico 23, o vídeo com o maior número de ocorrências – bem à frente dos outros – foi o 5 com 93 ocorrências de regressão (25,07%), seguido do 8 com 38 ocorrências (15,6%), do 3 com 48 ocorrências (12,94%), do 2 com 41 ocorrências (11,05%) e do 1 com 39 ocorrências de regressão (10,51%). Abaixo destes, ainda em ordem decrescente, estiveram o Vídeo 6 com 33 ocorrências (8,89%), o 4 com 31 (8,36%) e, por último, o 7 com 28 ocorrências de regressão (7,55%). Sobre isso, os vídeos com maior ocorrência de regressão também foram aqueles com o maior número de palavras: Vídeo 5 com 23 palavras e Vídeo 8 com 22 palavras. Os dois também estiveram entre aqueles com o maior número de caracteres: enquanto o Vídeo 5 teve 102 caracteres e foi o 1º neste quesito, o vídeo 8 foi o 3º com 83 caracteres. Outro vídeo com um alto número de palavras e caracteres foi o 3, com 18 palavras (3º maior) e 84 caracteres (2º maior), que também ficou entre os vídeos com o maior número de casos de regressão no experimento.

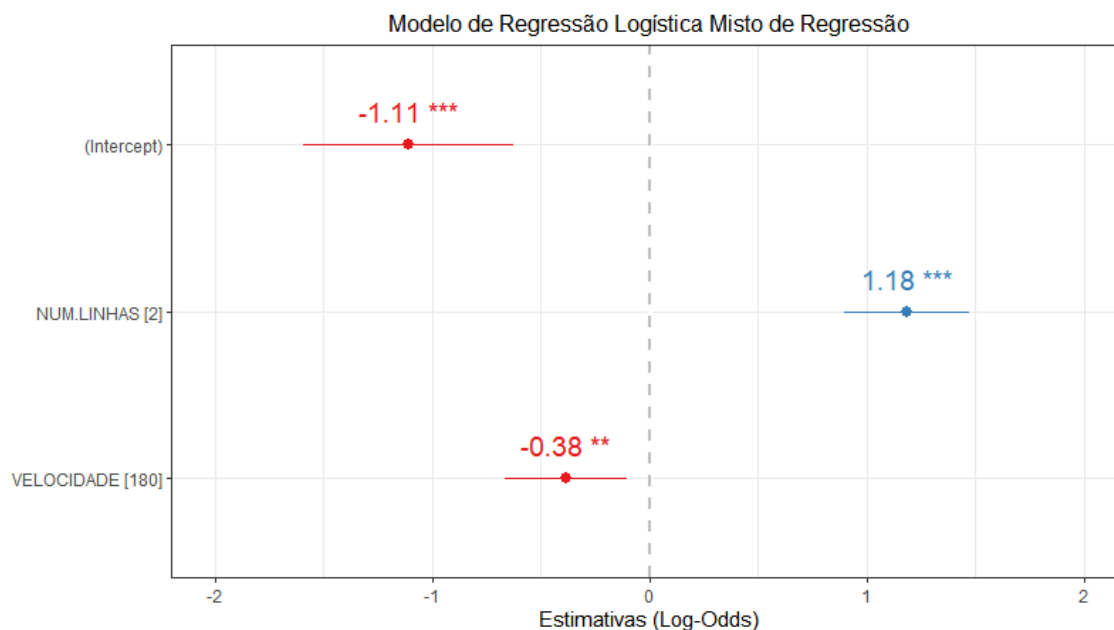
Essa questão nos faz refletir sobre a relação de ocorrências de regressão em um texto dinâmico como o vídeo legendado não só como uma possível dificuldade de processamento causada por questões técnicas e linguísticas, mas também como algo motivado pelo tempo de exposição da legenda em tela. Isso acontece porque os vídeos com a maior quantidade de caracteres também são aqueles que, proporcionalmente, passaram mais tempo em tela. Esse ponto merece ser discutido mais profundamente e será colocado mais adiante. Além disso, esse também é um motivo para que não desconsideremos nenhum dos dados

provenientes dessa medida. Como o tempo de exposição e o número de palavras das legendas não foram controlados, torna-se complicado considerar esses números díspares de dados de vídeos diferentes como sendo prova de qual vídeo se trata de uma variável de confusão nessa medida específica.

De modo a investigar essas proporções e sua relação com a ocorrência ou não de regressão durante o processamento de textos legendados, rodamos um modelo de regressão logística com Grupo, Número de Linhas e Velocidade como efeitos fixos e interceptos aleatórios de participante e vídeo. O contraste dos níveis dos efeitos fixos foi feito por meio de *dummy coding*<sup>77</sup>, tendo como nível de referência (em uma ordem alfabética e numérica) as variáveis de 1 linha, 145 ppm e Ouvinte (Grupo). Após retirarmos a variável Grupo por não ter sido significativa ( $\beta = -0,3958$ ,  $p = 0,1278$ ) para atingirmos um modelo mais ajustado, uma comparação por modelos aninhados mostrou que as variáveis de Número de Linhas e Grupo contribuem significativamente para o modelo ( $X^2 = 75,969$ ,  $p < 0,0001$ ). Esse modelo ajustado indicou que as legendas de 2 linhas (com 145 ppm) têm uma probabilidade maior de ter regressões ( $\beta = 1,1840$ ,  $p < 0,0001$ ), e as legendas de 180 ppm (com 1 linha), uma probabilidade menor de ter regressões ( $\beta = -0,3823$ ,  $p = 0,0075$ ), mostrando-se significativas. Além disso, nenhuma interação entre as variáveis foi identificada. No Gráfico 24, é possível visualizar os efeitos fixos supracitados.

---

<sup>77</sup> Esse tipo de contraste entre os níveis das variáveis é uma codificação padrão em modelos lineares (GODOY, 2019). Nesse tipo de contraste, um nível assume o valor igual a 0 (nível de referência e valor do intercepto) e outro, valor igual a 1. Para efeitos desse trabalho, todos os níveis de referência seguem uma ordem alfabética, sendo eles: Ouvinte (Grupo), 1 linha (Número de Linhas) e 145 ppm (Velocidade). Ainda segundo Godoy (2019, p. 78): “A tabela de coeficientes de um modelo linear com contrastes do tipo *dummy* mostra resultados referentes aos efeitos simples das variáveis envolvidas. Quando nosso experimento tem apenas uma variável de 2 níveis, isso não é um problema: afinal, o objetivo último da análise será contrastar A1 e A2”. Entretanto, segundo a autora, “isso pode ser um problema quando queremos reportar efeitos principais envolvendo mais de uma variável e suas interações” (GODOY, 2019, p. 78). Como não estipulamos hipóteses para a interação entre as duas variáveis independentes testadas no estudo (Número de Linhas e Velocidade), a utilização do *dummy coding* satisfaz, por ora, as necessidades deste trabalho. Contudo, vislumbramos a aplicação do *sum coding* – tipo de codificação que opõe os níveis das variáveis – em um momento futuro. Nesse tipo de contraste, ao invés de termos como intercepto o nível de referência em ordem alfabética e numérica das variáveis, temos a média geral das observações, e os coeficientes são os valores dos efeitos principais dessas variáveis testadas.

**Gráfico 24 – Modelo de Regressão Logística Misto de Regressão**

Fonte: Elaborada pelo autor.

Além dos efeitos fixos mencionados anteriormente, outro ponto a se notar no gráfico é o valor do intercepto, que, neste caso específico, interessa para o nosso trabalho. Como o tipo de dado que se trabalha com regressão logística é binário e a codificação dos dados relativos à regressão foi feita com “0 = não regressão” e “1 = regressão”, o que o coeficiente do intercepto está dizendo é que a probabilidade de se ter regressão é menor na legenda de 1 linha com 145 ppm ( $\beta = -1,1102$ ,  $p < 0,0001$ ). Para uma compreensão de como estão previstas as chances de se ter regressão ou não em todas as condições experimentais, a Tabela 25 traz os valores previstos pelo modelo.

**Tabela 25 – Valores previstos pelo modelo para ocorrência de regressão**

Linhas	Velocidade	Log-odds	Odds ratio	Probabilidade
1	145	-1,1102	0,3294	0,2478
1	180	-1,4925	0,2248	0,1835
2	145	0,0738	1,0765	0,5184
2	180	-0,3085	0,7345	0,4234

Fonte: Elaborada pelo autor.

Ao examinarmos a Tabela 25, vemos as previsões do modelo ajustado para as condições experimentais de 1 linha e 145 ppm, 1 linha e 180 ppm, 2 linhas e 145 ppm e 2 linhas

e 180 ppm em *log-odds*, *odds ratio*<sup>78</sup> e probabilidade. Visto que esta é um pouco mais intuitiva, vamos nos ater a ela. A partir da coluna de probabilidade, podemos ver que nas condições experimentais 1L145, 1L180 e 2L180, a probabilidade de ocorrer regressão é menor. Transformando em porcentagem, a probabilidade de 1L145 ter regressão é de 24,78%, de 1L180 é de 18,35% e de 2L180 é de 42,34%. A única condição experimental que favorece a ocorrência de regressão é a de 2L145, com uma probabilidade de 51,84%. Proporcionalmente, ao se dividir a probabilidade das legendas de 2 linhas por 2, percebemos que 2L180 ainda teve uma probabilidade maior de ocorrência do que 1L145 (21,17% vs. 18,35%), assim como 2L145 teve uma probabilidade ligeiramente maior de ocorrência do que 1L145 (25,92% vs. 24,78%).

Levando em consideração as hipóteses do trabalho e a partir da noção de que a regressão é um sinal de desconforto nos estudos sobre leitura, o fato de a legenda de 2 linhas ter favorecido a ocorrência de regressão não só de forma absoluta, como proporcionalmente, parece ir de encontro à nossa hipótese de que as legendas de 2 linhas seriam mais confortáveis do que as de 1 linha. Além disso, o efeito fixo de Velocidade também favoreceu a regressão, visto que a diferença entre as velocidades de 145 ppm e 180 ppm se mostrou significativa, contrariando a expectativa de que não haveria diferenças significativas no processamento de legendas nas duas velocidades pesquisadas.

Ao contrário de Número de Linhas e Velocidade, a variável independente Grupo não se mostrou significativa, indicando que os participantes surdos não realizam mais regressões do que os ouvintes de maneira significativa, apesar de serem leitores de L2. Inclusive, apesar de essa diferença não ter sido significativa, os ouvintes realizaram mais movimentos regressivos do que os surdos. Dessa forma, os resultados dessa medida (ocorrência de regressão) devem ser contextualizados com o que já havia sido ressaltado anteriormente: a ocorrência da regressão ou não no estudo de leitura de legendas pode ser tão associada a um custo de processamento? Dada a natureza dinâmica do texto legendado, a associação da

---

<sup>78</sup> “*Odds ratio*” ou “razão das chances”, em português, é um valor que pode ir do 0 ao infinito (GODOY, 2019). A razão das chances no nosso estudo funciona assim: digamos que em 100 legendas existam 45 ocorrências de regressão. Essas 45 ocorrências são entendidas nesse trabalho como “1”, visto que o valor de referência é 0 e representa a não ocorrência de um movimento regressivo em uma legenda. Dessa forma, podemos pensar em uma razão de 45:55, em que 45 representa o “sucesso” dessa ocorrência e 55 o seu “fracasso”. O resultado dessa razão é 0,81; ou seja, a cada legenda ocorre 0,81 regressão. Isso também pode ser lido ao contrário, como 55:45, que significa que a cada 1,22 legenda, ocorre 1 regressão. A chance é uma outra maneira de expressar **probabilidade**, que, por sua vez, é um número de 0 a 1. Utilizando este mesmo exemplo, a probabilidade de ocorrer uma regressão é de 0,45 ou, em **porcentagem**, 45%. O “*log-odds*”, por sua vez, é o logaritmo das chances. Como o *log-odds* é um valor que pode ser negativo ou positivo, ele é o valor padrão utilizado na regressão logística, visto que ele aponta se a direção da diferença entre o sucesso e o fracasso de uma ocorrência (regressão, neste caso) é negativa ou positiva. Por isso, o valor em *log-odds* é utilizado para representar graficamente esse tipo de modelo, conforme pode ser observado no Gráfico 24.

ocorrência ou não da regressão parece estar mais pautada no tempo maior disponível em tela (seja em legendas com velocidade menor ou com 2 linhas) do que em uma dificuldade do processamento em si. Isso será mais explorado durante a discussão dos nossos resultados mais adiante.

#### 4.1.4 Releitura

A variável resposta de releitura diz respeito às fixações após a finalização da leitura do bloco de legendas e pode estar relacionada a uma dificuldade de processamento. De maneira similar à regressão, os dados da releitura precisam ser qualificados, visto que o ritmo da leitura é definido pelo tempo de exposição da legenda, sendo este controlado pela própria legenda, e não pelo participante do experimento; sendo assim, os resultados também precisam ser considerados dentro desse contexto.

Essa medida foi analisada em duas perspectivas: i) duração média da releitura (em milissegundos); e ii) proporção de casos de releitura. Assim como foi feito com as regressões na subseção anterior, os dados serão apresentados de maneira separada em relação a essas duas perspectivas, visto que pressupõem visualização de dados e testes estatísticos diferentes.

A etapa de análise sobre a movimentação ocular dos participantes identificou 225 fixações (de até 800 milissegundos) que consistiam em releitura, um número menor ainda do que o de regressões. As fixações de releitura com mais de 800 milissegundos foram descartadas por, historicamente, os estudos de processamento com rastreamento ocular não considerarem fixações com tempo maior de duração. A estatística descritiva dividida por variável independente em relação à duração média de releitura pode ser vista nas Tabelas 26, 27 e 28, abaixo.

**Tabela 26 – Duração média de releitura por Grupo**

Grupo	Duração de releitura (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>Ouvinte</b>	298, 78	141, 26	12, 94	266
<b>Surdo</b>	320, 91	139, 45	13, 54	291, 5

Fonte: Elaborada pelo autor.

**Tabela 27 – Duração média de releitura por Número de Linhas**

Número de Linhas	Duração de Releitura (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>1</b>	316,50	133,86	14,35	283
<b>2</b>	304,60	144,87	12,33	266

Fonte: Elaborada pelo autor.

**Tabela 28 – Duração média de releitura por Velocidade**

Velocidade	Duração de Releitura (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>145</b>	308,47	140,56	10,81	267
<b>180</b>	311,41	141,69	18,93	267

Fonte: Elaborada pelo autor.

As estatísticas descritivas relativas à duração média de releitura dispostas nas tabelas acima estão de acordo como o que hipotetizamos sobre nossas variáveis: 1) a duração média de releitura em legendas de surdos (320,91 ms) foi maior do que de ouvintes (298,78 ms); 2) essa mesma duração de releitura foi maior em legendas de 1 linha (316,5 ms) do que em 2 linhas (304,6 ms); e, por último, 3) a duração média de releitura em legendas de 145 e 180 ppm foi bem próxima (308,47 ms e 311,41 ms, respectivamente). Esses dados, contudo, serão qualificados mais adiante por ocasião da descrição dos resultados da análise estatística dos dados.

Em relação ao cruzamento das condições experimentais (1L145, 1L180, 2L145 e 2L180), a Tabela 29, abaixo, traz a duração média de releitura por grupo de participantes surdos e ouvintes.

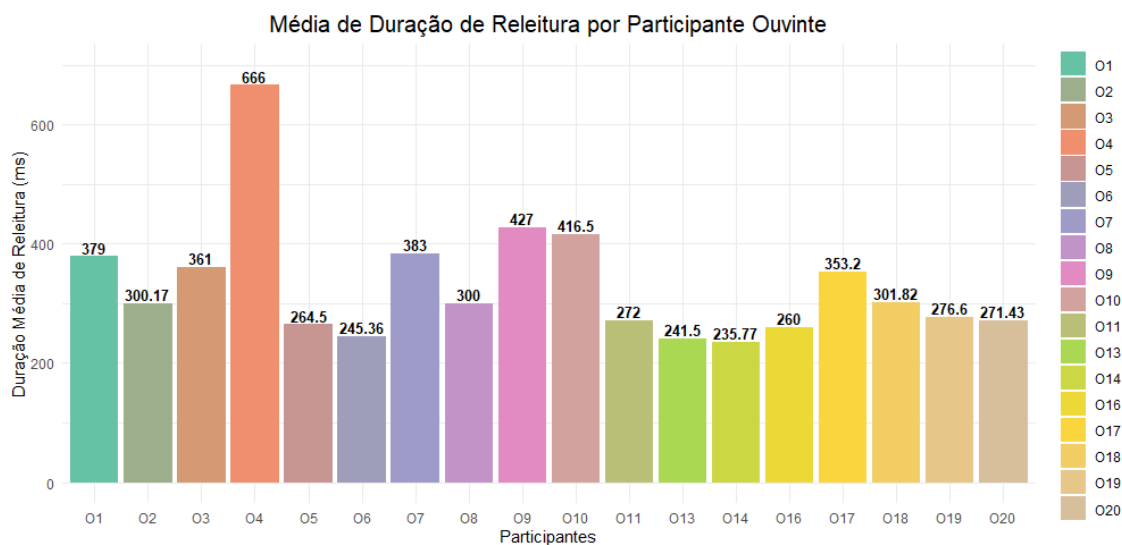
**Tabela 29 – Duração média de releitura por condição experimental**

Grupo	Condição Experimental	Duração de Releitura (ms)
<b>Surdo</b>	1L145	348,35
	1L180	266,41
	2L145	306,22
	2L180	366,50
<b>Ouvinte</b>	1L145	286,82
	1L180	343,38
	2L145	297,31
	2L180	291,09

Fonte: Elaborada pelo autor.

Com relação ao cruzamento das variáveis independentes em torno das condições experimentais utilizadas no estudo, vemos pela Tabela 29 que os participantes surdos tiveram um desempenho melhor em relação à duração média de releitura nas condições experimentais 1L180 (266,41 ms) e 2L145 (306,22 ms), sendo grande a diferença de 1L180 para as outras condições experimentais. O maior custo de processamento em relação a releitura foi em 2L180 (366,50 ms), seguido por 1L145(348,35 ms). Os participantes ouvintes, por sua vez, tiveram um melhor desempenho em velocidades distintas dos surdos: 1L145 (286,82 ms) e 2L180 (291,09 ms). As condições experimentais em que os ouvintes tiveram um maior custo de processamento foram em 1L180 (343,38 ms) e 2L145 (297,31 ms). A condição experimental de 1L180 em ouvintes, inclusive, afastou-se bastante das outras condições, as quais apresentaram uma pequena diferença. Nos surdos, entretanto, a diferença entre todas as condições experimentais foi bem maior, apontando uma certa dispersão dos dados no grupo em questão.

Sobre o desempenho dos participantes ouvintes em relação a essa medida, o Gráfico 25 traz essa informação.

**Gráfico 25 – Média de duração de releitura por participante ouvinte**

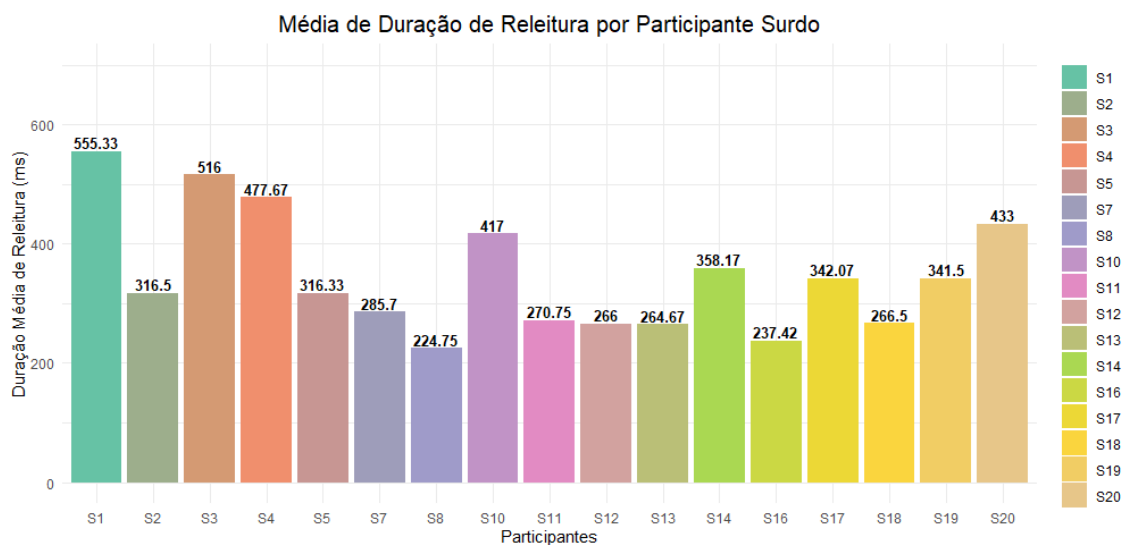
Fonte: Elaborado pelo autor.

Como é possível observar no Gráfico 25, há uma heterogeneidade acentuada em relação às médias de duração de releitura entre os ouvintes. Nesse sentido, O4 se destaca em relação a sua alta média (666 ms), bem acima da segunda maior média apresentada por O9 (427 ms). Ademais, os dois participantes em questão foram, respectivamente, o 5º e 6º maiores tempos de duração média de fixação no experimento (respectivamente 280,2 ms e 279,3 ms), conforme apresentado anteriormente. Essa média de O6, contudo, precisa ser contextualizada: o participante em questão só fez essa fixação de releitura no experimento, enquanto O9, para fins de comparação, que também apresentou uma média alta, mas acentuadamente inferior, realizou 8 fixações de releitura.

Já dentre os participantes ouvintes que tiveram as menores médias de duração de releitura, podemos destacar O14 (235,77 ms) e O13 (241,5 ms) com os menores custos de processamento em releitura. Nesse sentido, esses mesmos participantes também apresentaram um bom desempenho em processamento considerando a duração média de fixação. Enquanto O14 teve o 3º menor tempo médio de fixação (231,3 ms) entre ouvintes, O13 teve o 6º menor tempo médio de fixação no grupo de participantes (256,6 ms). Além disso, O14 teve o segundo maior número de fixações de releitura com 13, o que sugere que, junto a um número baixo de tempo médio de releitura e de fixação como um todo de O14, a noção de releitura está associada mais a um tempo maior de exposição da legenda em tela do que um desconforto *per se*.

Os participantes surdos, por sua vez, tiveram seus tempos médios de releitura distribuídos da seguinte maneira.



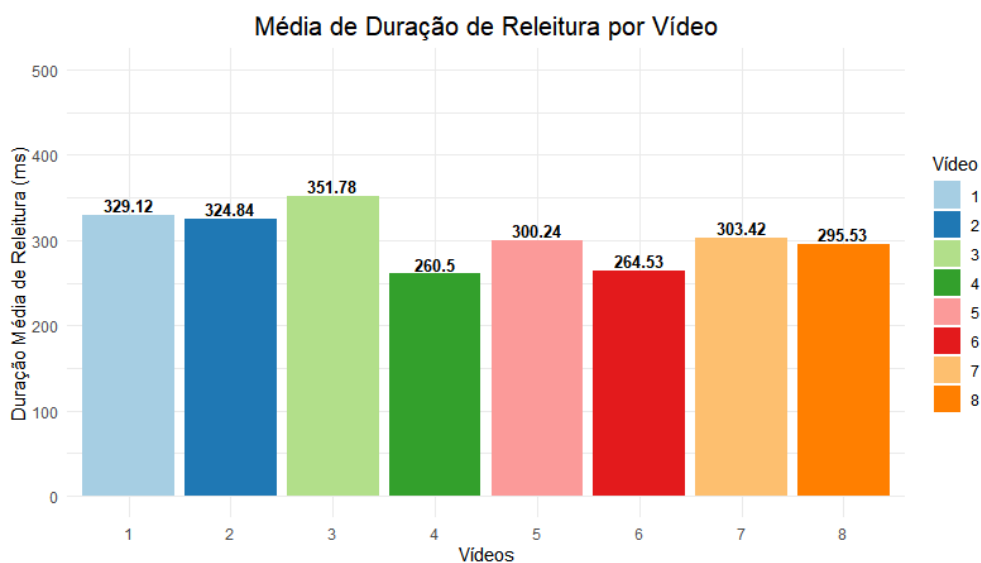
**Gráfico 26 – Média de duração de releitura por participante surdo**

Fonte: Elaborado pelo autor.

Heterogêneo como o grupo de ouvintes em relação a essa medida, os surdos que tiveram o maior tempo médio de fixação foram S1 (555,33 ms) e S3 (516 ms), médias altas quando comparadas com as dos outros participantes. Ademais, S1 não figurou entre os maiores tempos médios de fixação no grupo de surdos, sendo o 9º menor tempo médio com 303 ms. Diferentemente, S3 esteve entre os maiores tempos médios de fixação, sendo o 2º maior com 356,2 ms. Contudo, os dois participantes tiveram poucas fixações de releitura; S1 teve três e S3, apenas uma.

Em relação aos menores tempos de releitura, podemos destacar os participantes S8 (224,75 ms) e S16 (237,42 ms) como os menores tempos médios, respectivamente. Esses mesmos participantes apresentaram um custo baixo de processamento no que diz respeito às fixações como um todo, sendo S8 o 4º menor tempo médio de fixação (270,7 ms) e S16 o 3º (265,4 ms). S16, inclusive, apresentou o 3º maior número de ocorrências de releitura entre os surdos (12), diferentemente de S8, que apresentou apenas 4.

No Gráfico 27, podemos ver as médias de duração de releitura por vídeo (item) utilizado no experimento.

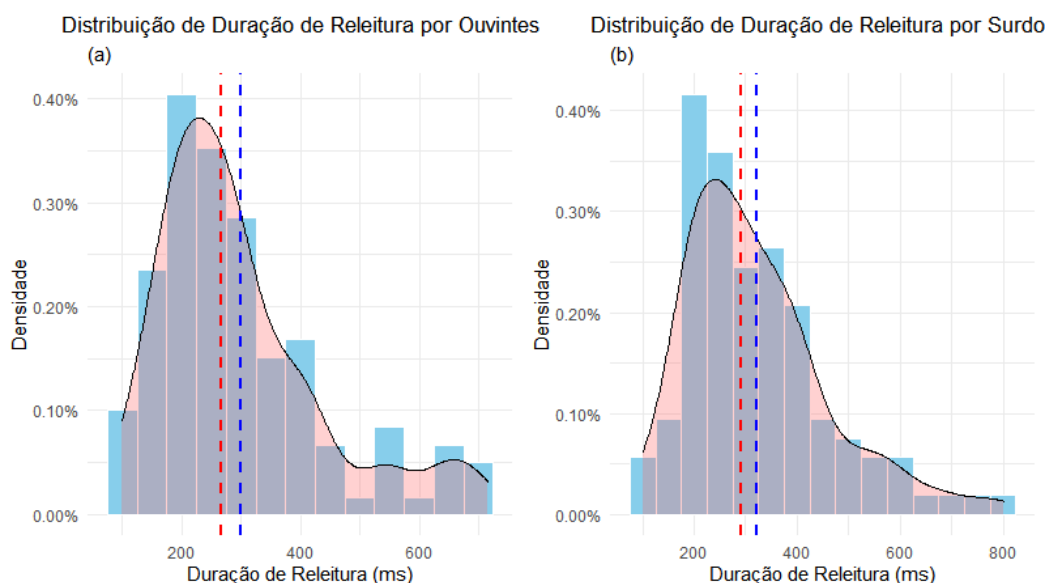
**Gráfico 27 – Média de duração de releitura por vídeo**

Fonte: Elaborado pelo autor.

Segundo os dados acima, de maneira similar aos dados encontrados na média de duração média de fixação, houve uma disparidade dos tempos médios entre os vídeos utilizados no experimento. Nesse sentido, o Vídeo 3 foi o que teve o maior tempo médio de releitura (351,7 ms), seguido dos vídeos 2 (324,84 ms), 1 (329,12 ms), 7 (303,42 ms) e 5 (300,24 ms). Os únicos abaixo dos 300 milissegundos foram os vídeos 6 (264,53 ms) e 4 (260,5 ms). Essa disparidade, inclusive, fez com que fosse necessário investigar Vídeo enquanto possível variável perturbadora, o que será apresentado na análise estatística dessa medida mais adiante.

No Gráfico 28, há a distribuição das 225 ocorrências de releitura e suas durações nos dois grupos de participantes.

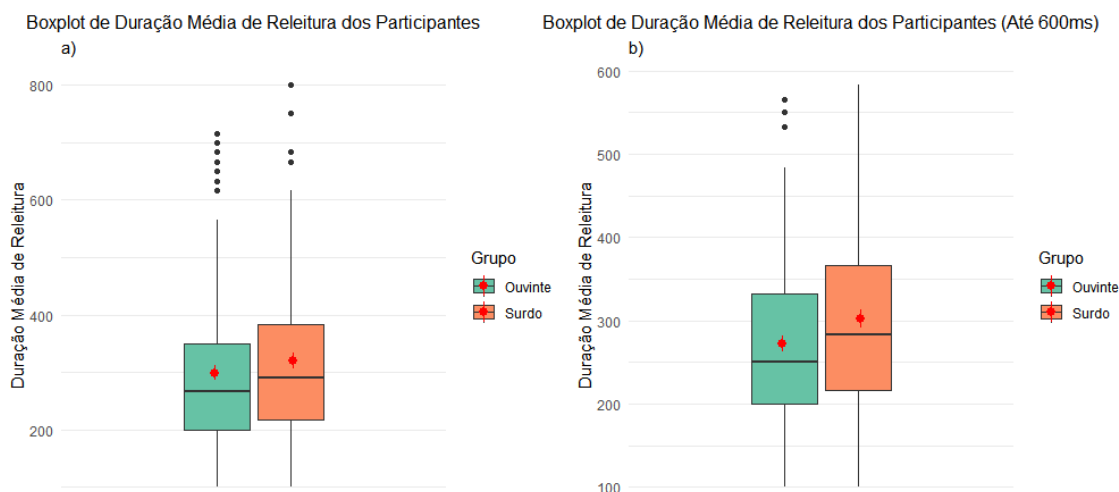
### Gráfico 28 – Distribuição de releitura em ouvintes e surdos



Fonte: Elaborado pelo autor.

Ao observarmos os histogramas em questão, percebemos que o maior número de fixações de releitura se situou na casa de 200-250 milissegundos nos dois grupos. Contudo, os valores altos à direita dos histogramas de ouvintes e surdos fazem com que a média (representada pela linha azul) esteja à direita da mediana (representada pela linha vermelha) de maneira bem semelhante em ambos os grupos (apesar de o grupo de surdos conter valores ainda mais altos, na casa dos 800 milissegundos), o que torna a distribuição dos dados mais próxima da não normal, como tem sido observado com dados de movimentação ocular neste trabalho no que tange às variáveis numéricas contínuas trabalhadas.

Sendo assim, de maneira análoga às outras medidas numéricas contínuas que temos analisado aqui, observamos a distribuição de casos de releitura em ambos os grupos através de *boxplot* para definir um limite para *outliers* e, assim, eliminar alguns dados que possam desviar a análise estatística. O Gráfico 29, abaixo, apresenta *boxplots* antes e depois da filtragem estabelecida.

**Gráfico 29 – Boxplot de duração média de releitura antes e após filtragem**

Fonte: Elaborado pelo autor.

No Gráfico 29, percebemos que nos *boxplots* à esquerda, os quais tinham ocorrências de releitura de até 800 milissegundos, há vários *outliers*. Os *boxplots* dos dois grupos pendem para a direita; contudo, após o limite estabelecido de 600 milissegundos, são encontrados poucos *outliers*, sendo todos no grupo de ouvintes. Assim, decidimos utilizar esse limite de tempo para a análise estatística dos dados, o que resultou em 212 ocorrências de releitura. Não eliminamos todos os *outliers* pois, dado o número reduzido de dados dessa variável, limitaria ainda mais a análise estatística.

Após essa filtragem, novas médias para as variáveis independentes de Grupo, Número de Linhas e Velocidade foram estabelecidas, as quais podem ser vistas nas Tabelas 30, 31 e 32 abaixo.

**Tabela 30 – Duração média de releitura por Grupo após filtragem**

Grupo	Duração de releitura (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>Ouvinte</b>	272,44	104,49	9,91	250
<b>Surdo</b>	302	112,02	11,14	283

Fonte: Elaborada pelo autor.

**Tabela 31 – Duração média de releitura por Número de Linha após filtragem**

Número de Linhas	Duração de Releitura (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>1</b>	294,78	103,17	11,39	283
<b>2</b>	281,31	112,44	9,86	266

Fonte: Elaborada pelo autor.

**Tabela 32 – Duração média de releitura por Número de Linha após filtragem**

Velocidade	Duração de Releitura (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>145</b>	287,83	112,84	8,92	267
<b>180</b>	282,50	96,66	13,40	266

Fonte: Elaborada pelo autor.

A partir dos dados das Tabelas acima, rodamos uma Análise de Variância (ANOVA) com as observações de releitura de até 600 milissegundos e tendo como variáveis independentes Grupo, Número de Linhas, Velocidade e Vídeo como provável variável de confusão. Esta análise mostrou que Vídeo não foi significativo ( $f(7, 201) = 0,916$ ,  $p = 0,4949$ ), permitindo que pudéssemos seguir sem nos preocuparmos em retirar nenhuma observação de nenhum dos vídeos. Um Teste de *Tukey post-hoc* confirmou que nenhuma das comparações entre pares foi significativa. Nesta mesma análise, a variável Grupo se mostrou significativa ( $f(1, 201) = 3,911$ ,  $p = 0,0493$ ), tendo o grupo de surdos uma média significativamente maior de duração de releitura do que os ouvintes. As diferenças na ordem de Número de Linhas ( $f(1, 201) = 0,479$ ,  $p = 0,4895$ ) e Velocidade ( $f(1, 201) = 0,029$ ,  $p = 0,8655$ ) não se mostraram significativas.

Concomitantemente à ANOVA, rodamos um modelo linear múltiplo (sem efeitos mistos) tendo como efeito fixo as variáveis de Grupo, Número de Linhas, Velocidade e Vídeo, assim como fizemos com as outras variáveis. A análise desse modelo linear mostrou que, novamente, nenhum dos 7 vídeos foi significativo em relação ao Vídeo 1, utilizado como nível de referência na análise do modelo. Outro Teste de *Tukey* confirmou que nenhuma das comparações entre pares múltiplos de Vídeo foi significativa.

Após isso, rodamos outro modelo, dessa vez um modelo linear múltiplo misto tendo como efeito fixo as variáveis de Grupo, Número de Linhas e Velocidade e como efeito aleatório o intercepto de participante.

**Tabela 33 – Modelo de Regressão Linear Misto de Duração Média de Releitura**

<b>Coefficiente</b>	<i>Estimativas</i>	<i>Conf.int (95%)</i>	<i>P-value</i>
<b>Intercepto</b>	281,97	247,81 – 316,13	<0,001
<b>Grupo (Surdo)</b>	30,43	-6,80 – 67,66	0,109
<b>Número de linhas (2)</b>	-9,81	-40,08 – 20,46	0,525
<b>Velocidade (180)</b>	1,85	-33,15– 36,85	0,918
<b>R<sup>2</sup> Marginal/R<sup>2</sup> Condicional</b> 0,022 / 0,096			

Fonte: Elaborada pelo autor.

Assim como na variável resposta de duração média de regressão, o modelo acima não mostrou nenhum dos efeitos fixos do experimento como significativo (Tabela 33). Apesar de os coeficientes de duração média de releitura em 2 linhas na variável de Número de Linhas ( $\beta = -9,81$ ,  $p = 0,525$ ) e por participantes surdos na variável Grupo ( $\beta = 30,43$ ,  $p = 0,109$ ) refletirem uma diferença considerável entre níveis nas previsões do modelo, é provável que a pouca quantidade de observações sobre essa medida (apenas 212 ocorrências utilizadas na modelagem) tenha feito com que não fosse alcançado um nível de significância como demarcado pelo valor-p das variáveis em questão. A pequena diferença na variável de Velocidade no nível de 180 ppm ( $\beta = 1,85$ ,  $p = 0,918$ ), entretanto, está de acordo com a hipótese prévia sobre a não significância entre níveis da variável Velocidade no processamento de surdos e ouvintes no experimento.

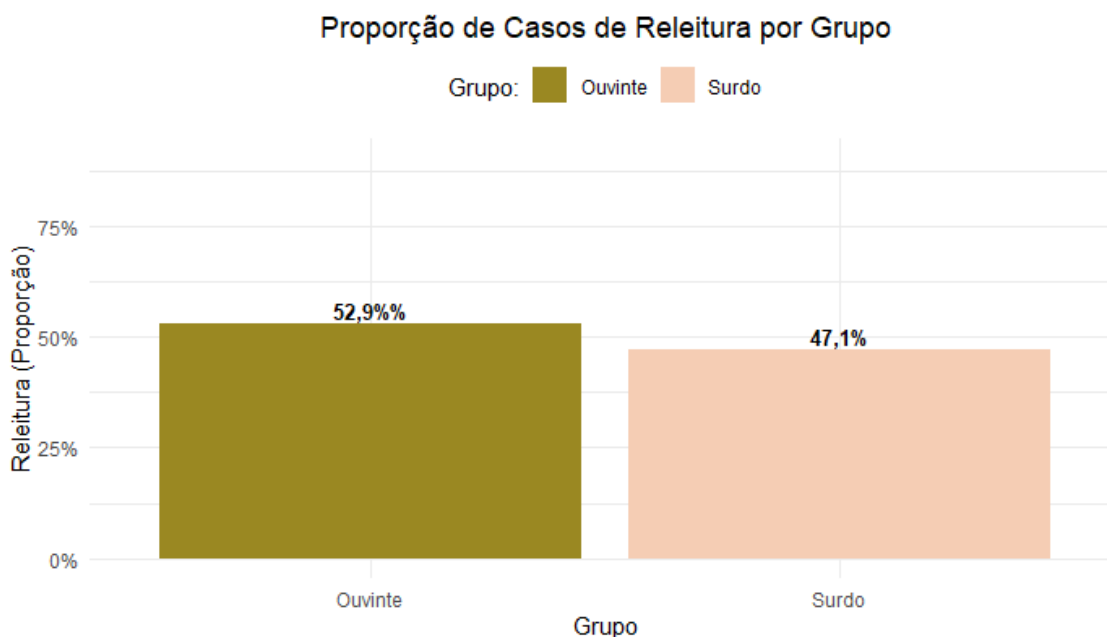
Agora, passemos à descrição dos resultados sobre a ocorrência de releitura nas condições experimentais do trabalho. Aqui, assim como fizemos com a medida de regressão na subseção anterior, discorreremos sobre as ocorrências da medida (releitura) no experimento e sobre a probabilidade de essa ocorrência ser extrapolada para a população (regressão logística); isto é, como o modelo de regressão consegue prever estimativas para o fenômeno da releitura para além da amostra do estudo, considerando o contexto mais amplo da população, no nosso caso, o processamento de legendas de 1 e 2 linhas (lentas e rápidas) por surdos e ouvintes brasileiros.

Antes de investigar estatisticamente a ocorrência de releitura, é necessário que examinemos em termos de proporção como as 225 ocorrências de releitura foram distribuídas nas variáveis independentes: Grupo, Número de Linhas e Velocidade.

Primeiramente, no que tange à distribuição de releitura por Grupo, a inspeção manual de movimentação ocular realizada identificou que 119 das 225 ocorrências foram em fixações realizadas por ouvintes (52,9% do total), enquanto 106 ocorrências de releitura foram realizadas por surdos (47,1% do total). Inicialmente, essa distribuição contraria a expectativa

do trabalho em relação a uma maior incidência de releitura por parte dos surdos, dada a sua relação com português como L2. Por essa razão, teriam maior dificuldade de leitura atrelada à ocorrência dessa medida especificamente, dada a noção de releitura nos estudos sobre processamento. A questão da releitura enquanto fator demonstrativo de custo de processamento de vídeos legendados será colocada durante a discussão dos dados. Essa proporção pode ser vista no Gráfico 30.

**Gráfico 30 – Proporção de casos de releitura por Grupo**



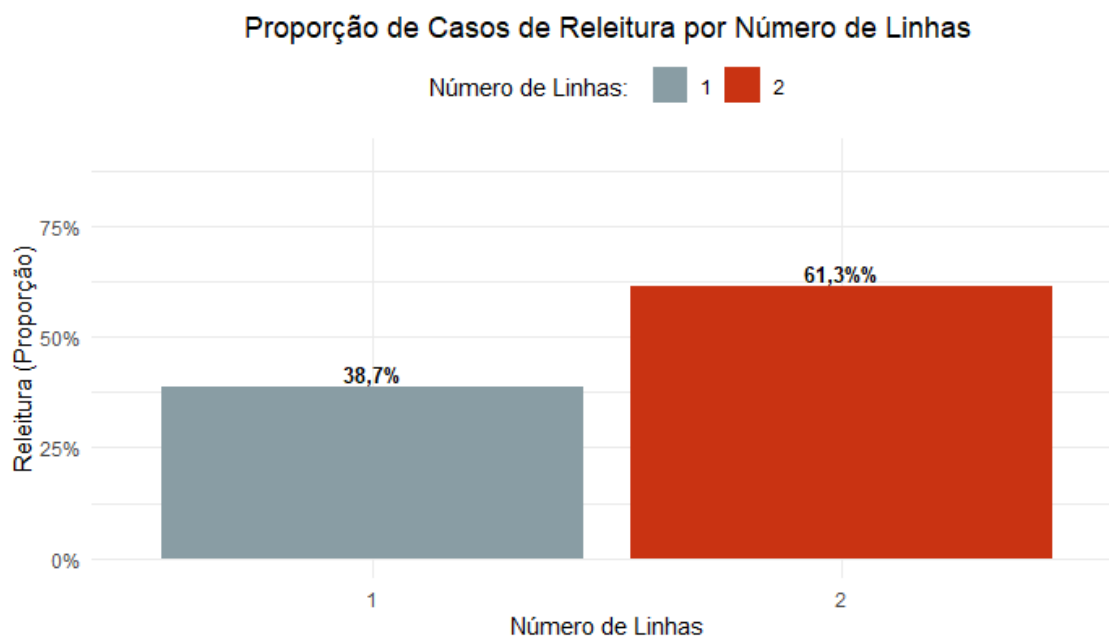
Fonte: Elaborado pelo autor.

Em uma relação com o total de fixações realizadas<sup>79</sup> (4.072), as 119 ocorrências de releitura por ouvintes representaram 5,51% das suas 2.159 fixações realizadas; já 106 ocorrências de releitura por surdos, os quais fizeram 1.913 fixações no total, representaram 5,54% desse total de releitura, uma proporção bem parecida. Sendo assim, mesmo realizando mais fixações do que surdos, a porcentagem de releitura por ouvintes foi bem semelhante.

Em relação ao Número de Linhas, 87 ocorrências de releitura foram em legendas de 1 linha (38,7%) e 138 ocorrências, nas de 2 linhas (61,3%), uma proporção de releitura nas legendas de 2 linhas menor do que o dobro da proporção em legendas de 1 linha<sup>80</sup>. Abaixo, no Gráfico 31, a proporção de ocorrências de releitura por Número de Linhas.

<sup>79</sup> Fixações regulares de até 600 ms e fixações de releitura de até 800 ms.

<sup>80</sup> Como as legendas de 2 linhas ficaram pelo menos duas vezes mais tempo em tela do que as legendas de 1 linha, seria normal caso elas tivessem o dobro do número de ocorrências de releitura do que as legendas de 1 linha.

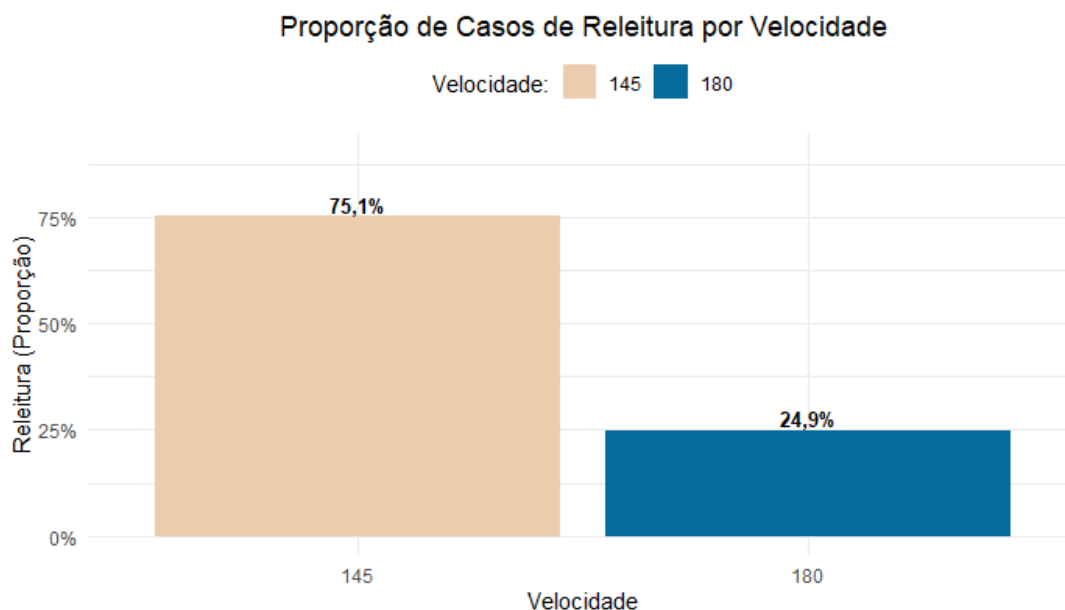
**Gráfico 31 – Proporção de casos de releitura por Número de Linhas**

Fonte: Elaborado pelo autor.

Fazendo uma correlação com o total de fixações realizadas, os casos de releitura (87) representaram 4,42% do âmbito de 1.964 fixações realizadas em legendas de 1 linha. Em legendas de 2 linhas (138), por sua vez, foram 6,54% no total de 2.108 fixações. Dessa forma, tanto de forma absoluta quanto proporcionalmente em relação ao total de fixações, as ocorrências de releitura foram maiores em legendas de 2 linhas.

No que diz respeito à variável de Velocidade, as legendas de 145 ppm apresentam um número bem maior de releituras do que as de 180 ppm: 169 (75,1% dos casos) contra 56 ocorrências (24,9% dos casos), inicialmente contrariando nossas expectativas em relação a uma diferença não tão substancial assim. Essa proporção pode ser vista no Gráfico 32.



**Gráfico 32 – Proporção de casos de releitura por Velocidade**

Fonte: Elaborado pelo autor.

Fazendo uma correlação com o total de fixações realizadas em ambas as velocidades, dentre as 2.153 fixações realizadas em legendas de 145 ppm por ambos os grupos de participantes, as 169 ocorrências de releitura representaram 7,84% dessas fixações, enquanto que as 56 ocorrências em legendas de 180 ppm representaram 2,91% das 1.919 fixações nesta velocidade, uma diferença aparentemente bem significativa.

Na Tabela 34 podemos ver a distribuição dos 225 casos de releitura a partir do cruzamento das variáveis do estudo.

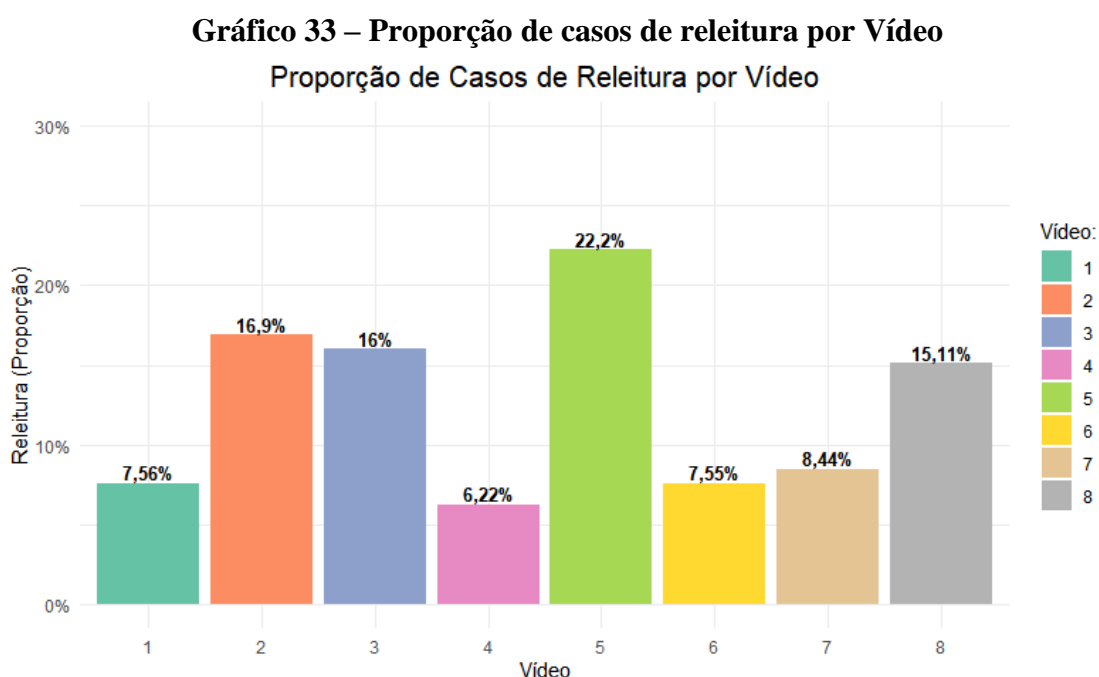
**Tabela 34 – Ocorrências de releitura por condição experimental**

Grupo	Linha	Velocidade	Frequência	Frequência
			Absoluta	Relativa
<b>Ouvinte</b>	1	145	28	12,45%
	1	180	13	5,78%
	2	145	<b>57</b>	25,33%
	2	180	21	9,33%
<b>Surdo</b>	1	145	34	15,11%
	1	180	12	5,33%
	2	145	<b>50</b>	22,22%
	2	180	10	4,45%

Fonte: Elaborada pelo autor.

Com os dados apresentados na Tabela 34, podemos observar que, dentre os ouvintes, as legendas de 145 ppm foram as que apresentaram um maior número de casos de releitura: 57, ou 25,33% do total (2L145) e 28, ou 12,45% do total (1L145). As condições experimentais de 180 ppm, por sua vez, foram as que apresentaram o menor número de casos de releitura entre ouvintes: 28, ou 12,45% do total (1L180) e 57, ou 25,33% do total (2L180). Esse padrão se repetiu entre os surdos em ambas as velocidades. Nas condições experimentais de 145 ppm, que apresentaram mais ocorrências de releitura, os surdos fizeram 50 (22,22%) e 34 (15,11%) releituras nas condições 2L145 e 1L145, respectivamente. Em relação às condições de 180 ppm, por sua vez, os surdos realizaram 12 (5,33%) e 10 (4,45%) releituras, respectivamente, nas condições experimentais de 1L180 e 2L180. Esses resultados sugeriram uma influência maior do tempo de exposição das legendas em tela para a ocorrência de releitura, sobre o que iremos nos debruçar de maneira mais detalhada à frente, na análise estatística dessa medida.

Abaixo, no Gráfico 33, apresentamos a proporção de casos de releitura por vídeo utilizado no experimento.



Fonte: Elaborado pelo autor.

De acordo com a análise descritiva dos casos de releitura, o item experimental com mais casos foi o Vídeo 5, com 50 casos (22,2% do total), seguido dos Vídeos 2, com 38 casos, (16,9%) e 3, com 36 casos (16%). Logo em seguida, aparece o Vídeo 8 com 34 casos de releitura

(15,11% dos casos), seguido dos Vídeos 7, com 19 casos (8,44%), 6, com 17 casos (7,55%) e 1, com 17 casos também (7,56%) e, por fim, do Vídeo 4, com 14 casos de releitura (6,22%). O fato de o Vídeo 5, que foi o vídeo com o maior número de palavras (23) e caracteres (102) no estudo, ser o item experimental com o maior número de ocorrências de releitura sugere novamente uma relação do tempo de exposição com a releitura da legenda, visto que legendas maiores ficam mais tempo expostas em tela e, dessa forma, teriam maior probabilidade de ser relidas.

Apesar de uma diferença notável na ocorrência de releitura por item experimental, decidimos não investigar os efeitos de vídeo como possível variável de confusão: 1) devido ao número reduzido de dados (225); e 2) pelo fato de a relação entre tempo de exposição e item experimental (vídeo) não ter sido controlada no experimento, algo que, a partir da análise descritiva, parece ser a questão envolvida no fenômeno de releitura.

Como um todo, os 20 participantes estiveram expostos a 480 inserções de legenda, sendo 240 legendas de 145 ppm e outras 240 legendas de 180 ppm. Dessas 480 inserções de legenda, 320 foram de legenda de 1 linha e 160 legendas de 2 linhas. Dentre estas, 160 foram de 1L145ppm, 160 de 1L180ppm, 80 de 2L145 e 80 de 2L180. Em relação à identificação de casos de releitura para essa medida, as que tinham releitura eram marcadas com “1” e as que não se constituíam em casos de releitura eram identificadas com “0”, organizando assim os dados para a análise binária pressuposta pela regressão logística. É a partir dessa lógica que o teste estatístico a seguir foi rodado em relação à probabilidade de ocorrência da medida de releitura nas condições experimentais do estudo.

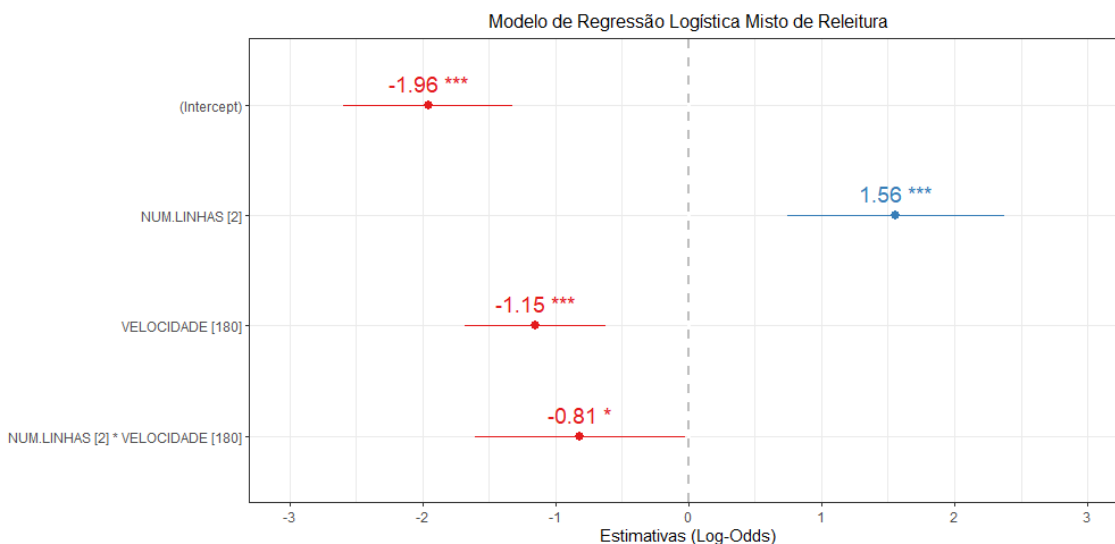
De modo a investigar essas proporções e sua relação com a ocorrência de releitura durante o processamento de textos legendados, rodamos um primeiro modelo de regressão logística misto com Grupo, Número de Linhas e Velocidade como efeitos fixos e interceptos aleatórios de participante e vídeo (item) e *slopes* aleatórios de Número de Linhas e Grupo por participante e Número de Linhas por vídeo. O contraste dos níveis dos efeitos fixos foi feito por meio de *dummy coding*, tendo como nível de referência (em uma ordem alfabética e numérica) as variáveis de 1 linha, 145 ppm e Ouvinte (Grupo), conforme já fora mencionado anteriormente. Nesse modelo, as variáveis independentes de Número de Linhas no nível de 2 linhas ( $\beta = 1,4029$ ,  $p = 0,0004$ ) e de Velocidade no nível de 180 ppm ( $\beta = -1,5406$ ,  $p < 0,0001$ ) foram significativas, ao contrário da variável Grupo, a qual não obteve significância ( $\beta = -0,3507$ ,  $p = 0,4443$ ).

Após retirar a variável Grupo em busca de um modelo preditivo mais ajustado, rodamos novamente o modelo com os interceptos e *slopes* supracitados, juntamente às variáveis

independentes de Número de Linhas e Velocidade. As variáveis independentes, novamente, foram significativas e um modelo contendo interação entre *Velocidade:Número de Linhas* foi observado como potencialmente o mais ajustado. Uma comparação entre modelos aninhados mostrou que o modelo contendo a interação explica melhor a proporção de ocorrências de regressão no experimento do que o modelo sem interação e apenas com os efeitos fixos e aleatórios ( $X^2 = 3,9402$ ,  $p = 0,0471$ ). Além disso, outra comparação entre modelos aninhados mostrou que o Número de Linhas e a Velocidade contribuem significativamente para o modelo em relação a um modelo com apenas intercepto e efeitos aleatórios ( $X^2 = 77,254$ ,  $p < 0,0001$ ). Isso mostra que as variáveis de Número de Linhas e Velocidade influenciam a ocorrência de releituras por ouvintes e surdos.

O modelo mais ajustado encontrado para prever a presença de fixação de releitura em processamento de legendas teve o Número de Linhas ( $\beta = 1,5594$ ,  $p = 0,0001$ ) e Velocidade ( $\beta = -1,1493$ ,  $p < 0,0001$ ) como significativos, além de uma interação entre legendas de 2 linhas e Velocidade de 180 ppm ( $\beta = -0,8142$ ,  $p = 0,0433$ ). No Gráfico 34, é possível visualizar os efeitos simples e a interação entre as variáveis desse modelo.

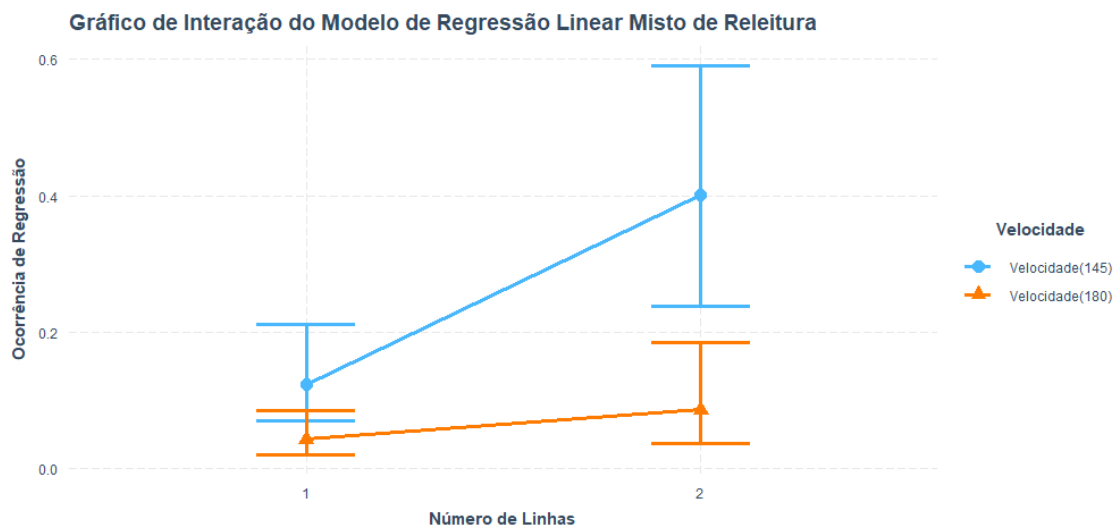
**Gráfico 34 – Modelo de Regressão Logística Misto de Releitura**



Fonte: Elaborado pelo autor.

No Gráfico 35, é possível visualizar como acontece a interação entre Número de Linhas e Velocidade.

### Gráfico 35 – Interação do Modelo de Regressão Linear Misto de Releitura



Fonte: Elaborado pelo autor.

O que o Gráfico 35 basicamente representa é que a mudança de 1 linha para 2 linhas significa um aumento dentro do nível de significância apenas quando na velocidade 145 ppm. Na Tabela 35, abaixo, é possível observar as previsões a partir do modelo nas condições experimentais de 1L145, 1L180, 2L145 e 2L180 em *log-odds*, *odds ratio* e probabilidade.

**Tabela 35 – Valores previstos pelo modelo para ocorrência de releitura**

Linhas	Velocidade	<i>Log-odds</i>	<i>Odds ratio</i>	Probabilidade
1	145	-1,9568	0,1408	0,1234
1	180	-3,1061	0,0447	0,0428
2	145	-0,3974	0,6720	0,4019
2	180	-2,3609	0,0943	0,0862

Fonte: Elaborada pelo autor.

Como colocado anteriormente com a regressão logística de regressão, vamos utilizar a probabilidade na Tabela 35 para falar sobre as previsões do modelo. Em relação às previsões de legendas com 180 ppm, segundo o modelo há uma probabilidade de 0,0862 (8,62%) de haver releitura em legendas com 2 linhas e 180 ppm, enquanto a probabilidade de ocorrência de releitura em legendas com 1 linha e 180 ppm é de 0,0428 (4,28%). Em relação a essas duas condições experimentais, é possível observar que a velocidade não parece exercer uma diferença na probabilidade de ocorrência de releitura, visto que é esperado que as legendas de 2 linhas tenham mais releituras do que de 1 linha. Isso ocorre, conforme já explicitado anteriormente, porque em nosso experimento legendas de 2 linhas significam o dobro de texto

e o dobro de tempo de exposição em comparação com legendas de 1 linha. Sendo assim, legendas de 1 linha e 2 linhas na velocidade de 180 ppm têm probabilidade de ocorrência de releitura bem proporcional.

Quando observamos os valores previstos para as legendas de 145 ppm, contudo, vemos que há uma diferença bem grande entre legendas com 1 linha e 2 linhas. Na condição experimental de 1 linha e 145 ppm, há uma probabilidade de 0,1234 (12,34 %) de ocorrência de regressão, mais do que na condição de 1 linha e 180 ppm, reforçando o que já foi discutido sobre a influência do tempo de exposição em tela nas medidas de regressão e também de releitura. A legenda de 2 linhas e 145 ppm, contudo, destaca-se em relação às outras condições experimentais quanto à probabilidade alta de ocorrência de releitura: 0,4019 (40,19%). Essa diferença se mostra grande quando comparamos com 2L180, a outra condição experimental de 2 linhas, que apresentou uma probabilidade de 0,0862 (8,62%) de ocorrência de releitura.

Proporcionalmente, ao dividir as probabilidades de ocorrência de releituras nas legendas de 2 linhas por 2, vemos que 2L145 tem uma probabilidade bem maior de ocorrência de releitura do que 1L145 (20,09% vs. 12,34%, respectivamente), enquanto 2L180 é bem próximo de 1L180 (4,31% vs. 4,28%, respectivamente), o que explica o efeito de interação no nosso modelo. Essas questões serão discutidas à luz da teoria de processamento de legendas mais à frente, durante a discussão dos dados.

#### 4.1.5 Encapsulamento

Outra variável interessante para observamos a influência do uso de 1 linha ou 2 linhas e nas velocidades de 145 ppm e 180 ppm é o encapsulamento. Como já dito anteriormente, há uma tendência a se gastar mais tempo na última fixação da sentença (o que vamos aqui correlacionar com o final da legenda, que pode ser ou não o final de um período). O que esse tempo maior de fixação significa? Segundo Just e Carpenter (1980), o final do processamento desse período acontece durante a última fixação, porque o leitor atrela referências que ainda não haviam sido associadas aos seus respectivos referentes durante a leitura. Como é o momento em que possíveis inconsistências e digressões deságuam, interessamos saber se as variáveis independentes do trabalho podem estar atreladas à última fixação da inserção de legenda.

A primeira parte de análise, inspeção sobre a movimentação ocular dos participantes, identificou 816 fixações finais nas inserções de legenda do experimento (fixações de encapsulamento) de até 800 milissegundos. Em relação às variáveis independentes do

estudo, a duração média de encapsulamento em milissegundos e outras estatísticas descritivas (desvio padrão, erro padrão e mediana) estão presentes nas Tabelas 36, 37 e 38.

**Tabela 36 – Duração média de encapsulamento por Grupo**

Grupo	Duração De Encapsulamento (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>Ouvinte</b>	303,49	139,79	6,67	266
<b>Surdo</b>	378,50	163,17	8,39	350

Fonte: Elaborada pelo autor.

**Tabela 37 – Duração média de encapsulamento por Número de Linha**

Número de Linhas	Duração de Encapsulamento (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>1</b>	343,92	157,50	6,73	300
<b>2</b>	326,68	151,13	9,21	267

Fonte: Elaborada pelo autor.

**Tabela 38 – Duração média de encapsulamento por Velocidade**

Velocidade	Duração de Encapsulamento (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>145</b>	340,13	153,04	7,55	300
<b>180</b>	336,33	158,20	7,85	300

Fonte: Elaborada pelo autor.

As estatísticas descritivas das tabelas acima estão em consonância com o que esperávamos a partir de nossas hipóteses: 1) a duração média de encapsulamento foi maior em 1 linha (343,92 ms) do que em 2 linhas (326,68 ms); 2) apesar de diferentes, a duração média de encapsulamento em legendas de 145 ppm (340,13 ms) e 180 ppm (336,33 ms) foi bem próxima; 3) a duração média de encapsulamento foi bem maior no grupo de surdos (378,50 ms) do que no de ouvintes (303,49 ms).

Em relação ao cruzamento das condições experimentais (1L145, 1L180, 2L145 e 2L180), a Tabela 39 abaixo traz a duração média de encapsulamento pelo grupo de participantes surdos e ouvintes.

**Tabela 39 – Duração média de encapsulamento por condição experimental**

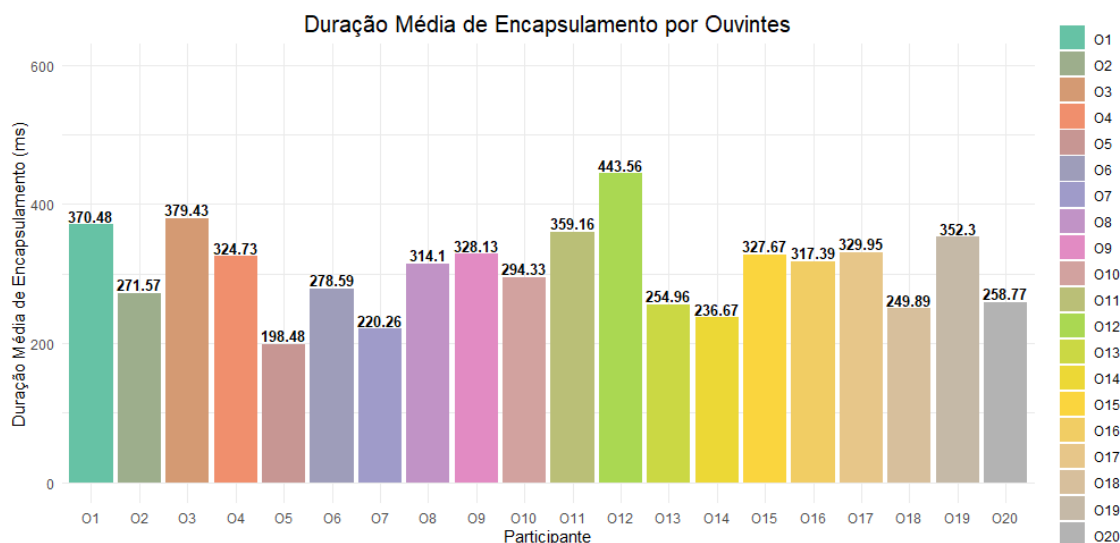
Grupo	Condição Experimental	Duração de Encapsulamento (ms)
<b>Surdo</b>	1L145	375,86
	1L180	399,74
	2L145	360,85
	2L180	352,16
<b>Ouvinte</b>	1L145	313,79
	1L180	292,49
	2L145	313,38
	2L180	294,43

Fonte: Elaborada pelo autor.

Ao observamos a Tabela 39, percebemos que, entre surdos, as condições experimentais com 2 linhas (2L145 e 2L180) proporcionaram um desempenho melhor, ou seja, um tempo menor de duração média de encapsulamento, do que as condições experimentais com 1 linha (1L145 e 1L180). Podemos fazer essa afirmação baseados no fato de que a menor média foi de 352,16 ms (2L180), média não muito distante da de 2L145 (360,85 ms). No que diz respeito ao grupo de ouvintes, a variável independente de Número de Linhas não pareceu exercer papel preponderante na duração média de encapsulamento, sendo as médias das condições 1L145 (313,79 ms) e 2L145 (313,38 ms), assim como 1L180 (292,49 ms) e 2L180 (294,43 ms), as quais compartilhavam a mesma velocidade, bem parecidas. Inclusive, o menor tempo de encapsulamento entre os participantes ouvintes foi em uma condição experimental de 1 linha, a 1L180 (292,49 ms).

O Gráfico 36 traz duração média de encapsulamento por participantes do grupo de ouvintes.

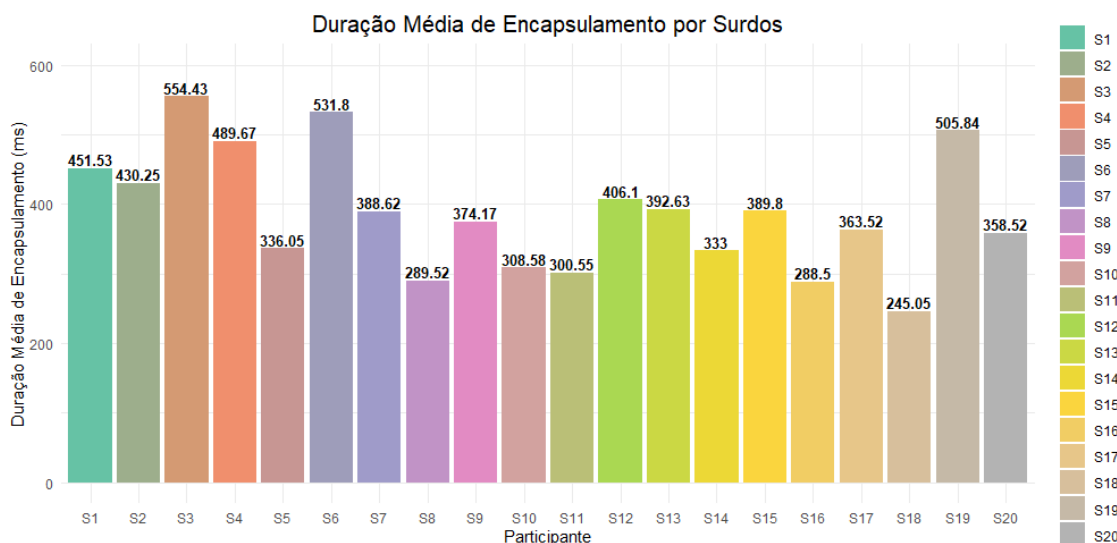


**Gráfico 36 – Duração média de encapsulamento por ouvintes**

Fonte: Elaborado pelo autor.

Como é possível perceber pelo Gráfico 36, há uma variabilidade grande entre as durações médias de encapsulamento entre os participantes ouvintes do estudo, indicando uma heterogeneidade significativa dentro do grupo. Correlacionando o desempenho dos participantes do grupo de ouvintes nessa medida/resposta com o seu desempenho na duração média de fixação (apresentada anteriormente), é possível observar uma certa linearidade entre os dois indicadores de custo no processamento de legendas. O12, por exemplo, que teve o pior desempenho na duração média de encapsulamento (443,56 ms), também apresentou o maior tempo de duração média de fixação (305,9ms). O5, por sua vez, que apresentou a menor duração média de encapsulamento (198,48 ms), obteve o 2º menor tempo de duração média de fixação (208 ms).

A duração média de encapsulamento entre os participantes do grupo de surdos, por sua vez, pode ser observada no Gráfico 37.

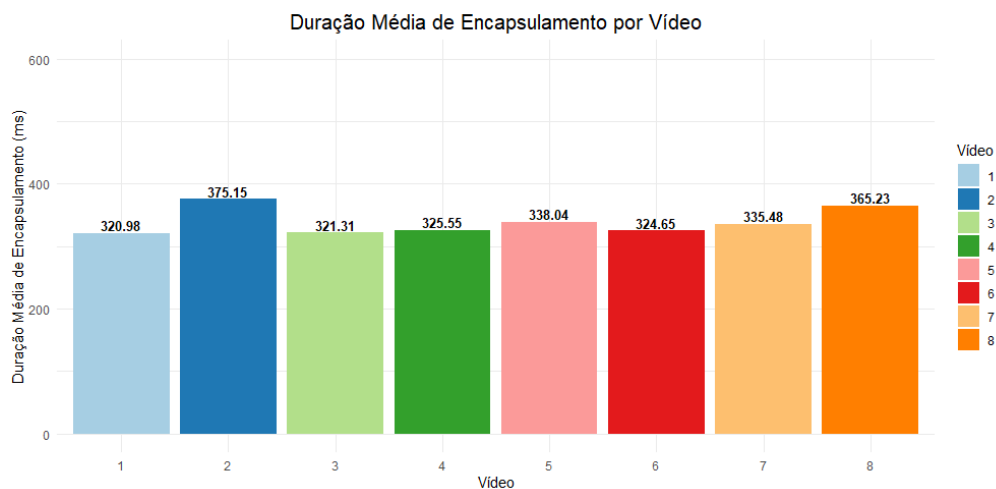
**Gráfico 37 – Duração média de encapsulamento por surdos**

Fonte: Elaborado pelo autor.

Assim como os ouvintes, a duração média de encapsulamento por participantes surdos apresentou uma variância grande entre os participantes. Continuando o tipo de correlação feita na apresentação dos dados dos ouvintes, observamos que o S3, que teve o maior tempo de duração nessa variável resposta (554,43 ms), também teve o 2º maior tempo de duração média de fixação (356,9 ms). Já S18, que teve o menor tempo de duração média de encapsulamento (245,05 ms), também teve o menor tempo de duração média de fixação entre os surdos (254 ms).

De forma análoga ao desempenho entre os participantes ouvintes, é possível perceber que a correlação entre a duração média de fixação e a duração média de encapsulamento, duas medidas utilizadas para investigar o custo de processamento durante a leitura, também pode ser observada nos participantes surdos utilizados como exemplo acima.

Em relação ao tempo por vídeo utilizado como estímulo, no Gráfico 38 estão as médias de duração de encapsulamento dos participantes do experimento.

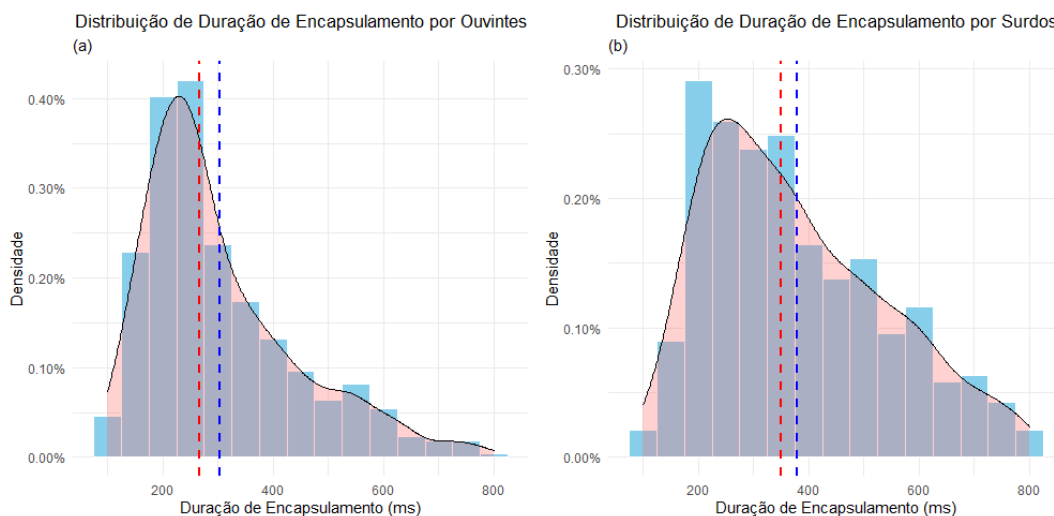
**Gráfico 38 – Duração média de encapsulamento por vídeo**

Fonte: Elaborado pelo autor.

No Gráfico 38 é possível observar que os Vídeos 1, 3, 4, 5, 6 e 7 tiveram médias de certa forma aproximadas de encapsulamento, uma distância de cerca de 17,06 milissegundos entre os Vídeos 1 (320,98 ms) e 5 (338,04 ms). Os vídeos 2 (375,15 ms) e 8 (365,23 ms), entretanto, afastam-se de maneira destacada em relação às médias de encapsulamento. O Vídeo 8, aliás, figurou novamente de maneira negativa em uma medida investigada nesse trabalho, apontando realmente que se trata de um vídeo com um processamento mais difícil do que os outros. Assim como nas outras variáveis analisadas, essa diferença de média de alguns vídeos nos fez checar se havia alguma distinção estatisticamente significativa entre estímulos, o que será apresentado mais à frente nesta seção.

Em relação à distribuição das 816 fixações de encapsulamento e suas durações, podemos observar o Gráfico 39, que mostra essa distribuição entre os dois grupos de participantes.

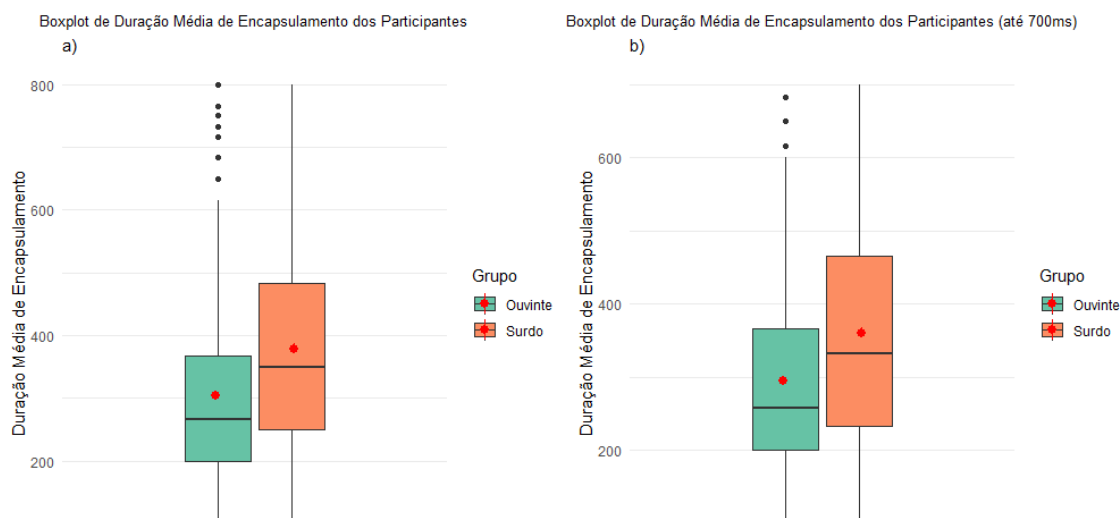
### Gráfico 39 – Distribuição de duração de encapsulamento por grupo



Fonte: Elaborado pelo autor.

No Gráfico 39 é possível perceber que as fixações de encapsulamento entre 250 e 300 milissegundos foram as que mais ocorrem entre os ouvintes, enquanto nos surdos foi entre 200 e 250 milissegundos. Contudo, no grupo de surdos há uma densidade maior de ocorrência proporcional a partir da casa dos 300 milissegundos até 800, medida esperada, visto que a duração média de encapsulamento foi maior entre os surdos do que entre os ouvintes. Além disso, em ambos os grupos, há valores altos à direita dos histogramas que puxam o valor da média (representada pela linha azul) para a direita da mediana (representada pela linha vermelha), sinalizando uma distribuição mais próxima da não normal, como é comum com dados oriundos de rastreamento ocular e como já foi visto nas outras variáveis dependentes numéricas contínuas analisadas até então.

Assim como nas outras variáveis numéricas contínuas, observamos a distribuição das ocorrências de encapsulamento entre ouvintes e surdos através de *boxplot* para identificação de uma linha de corte para *outliers* em função da posterior análise estatística dos dados. O Gráfico 40 representa essa questão.

**Gráfico 40 – Boxplot de duração média de encapsulamento antes e depois da filtragem**

Fonte: Elaborado pelo autor.

A partir do Gráfico 40, é possível perceber que, inicialmente, com fixações de encapsulamento de até 800 milissegundos, há um número considerável de *outliers* no *boxplot* de ouvintes, à esquerda do gráfico. Ao definirmos uma filtragem de fixações de encapsulamento de até 700 milissegundos, visível à direita no gráfico, o número de *outliers* no *boxplot* de ouvintes é bastante reduzido. Dessa forma, decidimos utilizar esse limite para tentar aproximar os dados a uma distribuição mais normal e, além disso, não perder observações de uma medida já com dados reduzidos dentro do nosso experimento. Assim, após a filtragem, restaram 791 observações sobre encapsulamento.

A seguir, discorreremos sobre os resultados dos testes estatísticos da ANOVA, do Teste T e do Modelo de Regressão Linear Misto a partir dos números dessas 791 observações após filtragem. Apresentamos nas Tabelas 40, 41 e 42 as estatísticas descritivas de duração média de encapsulamento por variável.

**Tabela 40 – Duração média de encapsulamento por Grupo após filtragem**

Grupo	Duração de Encapsulamento (ms)	Desvio Padrão	Erro Padrão	Mediana
Ouvinte	295,27	127,21	6,13	258
Surdo	360,89	144,66	7,61	333

Fonte: Elaborada pelo autor.

**Tabela 41 – Duração média de encapsulamento por Número de Linha após filtragem**

Número de Linhas	Duração de Encapsulamento (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>1</b>	328,33	138,06	6,01	300
<b>2</b>	318,99	141,68	8,71	267

Fonte: Elaborada pelo autor.

**Tabela 42 – Duração média de encapsulamento por Velocidade após filtragem**

Velocidade	Duração de Encapsulamento (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>145</b>	328,21	138,73	6,95	283
<b>180</b>	322,19	139,91	7,05	283

Fonte: Elaborada pelo autor.

Rodamos, primeiramente, uma Análise de Variância (ANOVA), tendo Vídeo enquanto preditor ao lado de Número de Linhas, Grupo e Velocidade, de modo a detectar se os Vídeos 2 e 8, especialmente, impactavam os resultados e, dessa forma, deveriam ser retirados da análise. Essa ANOVA mostrou que Vídeo foi significativo ( $f(7, 780) = 2,327$ ,  $p = 0,0236$ ). Ao rodarmos uma Teste de *Tukey* de comparação de pares múltipla, pudemos observar que, em uma comparação entre os Vídeos 2 e 6, o valor-p ajustado se mostrou significativo ( $p = 0,0495$ ) e em uma comparação entre os Vídeos 2 e 3 o valor-p ajustado foi marginalmente significativo ( $p = 0,0672$ ), sugerindo a variação grande entre o Vídeo 2 e os outros estímulos utilizados no experimento que já havia sido possível visualizar na análise descritiva anteriormente. Dessa forma, decidimos rodar as ANOVAs com e sem o Vídeo 2 em questão de modo a verificar se a sua inclusão influenciaria os resultados estatísticos.

Por um lado, a ANOVA contendo todos os vídeos mostrou que a diferença entre surdos e ouvintes foi significativa ( $f(1, 787) = 45,979$ ,  $p < 0,0001$ ), enquanto a diferença tanto entre 1 e 2 linhas ( $f(1, 787) = 0,323$ ,  $p = 0,570$ ) quanto entre as velocidades de 145 e 180 ppm ( $f(1, 787) = 0,430$ ,  $p = 0,512$ ) não se mostrou significativa. A ANOVA sem o Vídeo 2, por outro lado, com 696 observações após a retirada dos dados relativos ao vídeo, também mostrou diferença significativa entre surdos e ouvintes ( $f(1, 692) = 37,433$ ,  $p < 0,0001$ ), mas não entre 1 e 2 linhas ( $f(1, 692) = 0,087$ ,  $p = 0,768$ ) e também não entre as velocidades de 145 e 180 ppm ( $f(1, 692) = 2,206$ ,  $p = 0,138$ ).

Além da ANOVA descrita acima, rodamos dois modelos de regressão linear misto: com e sem o Vídeo 2. Em primeiro lugar, utilizamos um modelo com todos os vídeos e os preditores de Grupo, Número de Linhas e Velocidade (efeitos fixos), tendo como intercepto aleatório participante e vídeo (item) e *slopes* aleatórios de Grupo, Número de Linhas e Velocidade por participante. Esse modelo, assim como a ANOVA, retornou significância apenas para a variável de Grupo, tendo Ouvinte como referência ( $\beta = -68,518$ ,  $p = 0,0011$ ), enquanto Número de Linhas ( $\beta = -6,619$ ,  $p = 0,5606$ ) e Velocidade ( $\beta = -3,353$ ,  $p = 0,7434$ ) não se mostraram significativas. Em segundo lugar, rodamos um outro modelo sem o Vídeo 2. Contudo, os resultados não divergiram do modelo com todos os vídeos de maneira significativa. Esse modelo teve como preditores Grupo, Número de Linhas e Velocidade (efeito fixos) e interceptos aleatórios de participante e vídeo (item); essa foi a versão que convergiu (sem *slopes* aleatórios, diferentemente do modelo anterior). Esse modelo, semelhantemente ao modelo com todos os vídeos, retornou a variável de Grupo como significativa, tendo Ouvinte como nível de referência ( $\beta = 68,542$ ,  $p = 0,0014$ ), enquanto Número de Linhas ( $\beta = 1,820$ ,  $p = 0,8532$ ) e Velocidade ( $\beta = -9,751$ ,  $p = 0,2937$ ) não foram significativas.

Como Número de Linhas e Velocidade não se mostraram significativas na análise múltipla, procuramos um modelo mais ajustado tendo apenas Grupo enquanto preditor (efeito fixo). Além de Grupo como preditor, utilizamos participante e vídeo (item) como efeitos aleatórios e Número de Linhas e Velocidade por participante como *slopes* aleatórios. No primeiro modelo, com todos os vídeos, Grupo foi significante ( $\beta = 67,77$ ,  $p = 0,0013$ ), prevendo aumento de 67,77 milissegundos de duração média de encapsulamento quando passamos de Ouvintes para Surdos. Esse modelo apresentou um  $R^2$  condicional de 0,240, o que quer dizer que a duração média de encapsulamento pode ser explicada em 24% a partir da variável independente de Grupo. Em uma comparação entre modelos aninhados, esse modelo se mostrou significativo ( $X^2 = 9,7848$ ,  $p = 0,017$ ), indicando que a variável Grupo tem relevância na explicação da duração média de encapsulamento.

Sem o Vídeo 2, tendo o mesmo efeito fixo de Grupo e os mesmos efeitos aleatórios descritos no parágrafo acima, Grupo também foi significante, mas o coeficiente de duração média de encapsulamento mudou um pouco ( $\beta = 73,30$ ,  $p = 0,0006$ ). Esse modelo também apresentou um  $R^2$  condicional semelhante, de 0,277, ou que o modelo consegue explicar 27,7% do fenômeno analisado, uma porcentagem maior do que o modelo com todos os vídeos. Em uma comparação entre modelos aninhados, esse modelo também se mostrou significativo ( $X^2 = 10,802$ ,  $p = 0,0010$ ).

Na Tabela 43, podemos observar a previsão de duração média de encapsulamento de surdos e ouvintes em legendas de 1 e 2 linhas a partir dos modelos apresentados.

**Tabela 43 – Valores previstos pelo modelo para duração média de encapsulamento**

Grupo	Modelo	Duração de Encapsulamento (ms)
<b>Ouvinte</b>	Modelo com todos os vídeos	298,61
<b>Surdo</b>	Modelo com todos os vídeos	366,38
<b>Ouvinte</b>	Modelo sem o Vídeo 2	288,92
<b>Surdo</b>	Modelo sem o Vídeo 2	361,5

Fonte: Elaborada pelo autor.

Apesar de estarmos mostrando os modelos com e sem um vídeo possivelmente perturbador, é interessante ressaltar que o  $R^2$  condicional dos dois modelos é bem parecido, com leve superioridade do modelo sem o vídeo, que explica melhor o fenômeno do encapsulamento por surdos e ouvintes. Essa medida será discutida à luz da teoria do processamento de legendas mais à frente, durante a discussão dos dados.

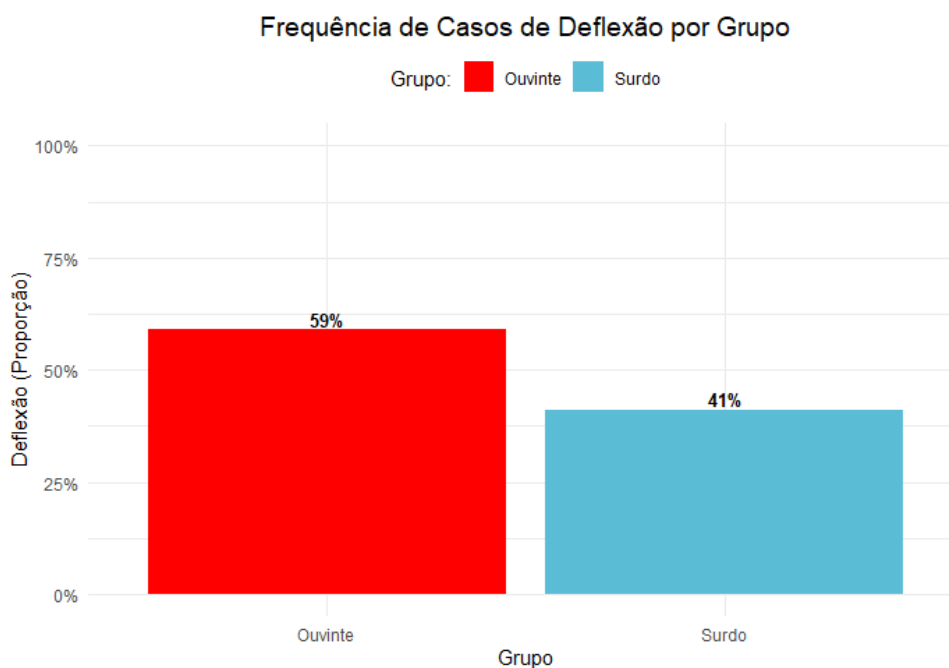
#### 4.1.6 Deflexão

Outra medida que, historicamente, tem sido utilizada para se observar o processamento de legendas é o da deflexão. Essa medida diz respeito à interrupção do processo de leitura da legenda para visita à imagem e há, na literatura da área, registro de duas formas de como ela pode ocorrer: i) durante a leitura da legenda; e ii) após a finalização da leitura da legenda, há um retorno da imagem para a legenda e, conseqüentemente, há um processo de releitura. No contexto desse trabalho, analisamos apenas o primeiro exemplo de deflexão. Apesar de parecer um movimento natural olhar a imagem em um produto audiovisual, esse movimento antes de completar o processamento da legenda (seja uma sentença, um sintagma etc.) se trata de uma digressão e, dessa forma, um número alto de deflexões pode comprometer esse processamento (KRUGER; SZARKOWSKA; KREJTZ, 2015).



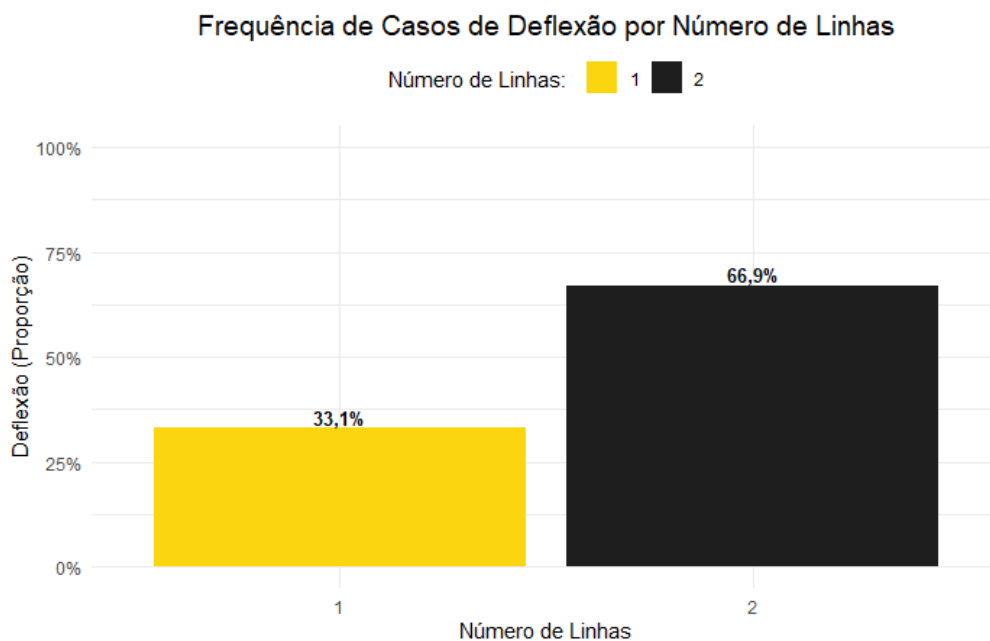
A inspeção da movimentação ocular dos participantes surdos e ouvintes indicou 139 movimentos deflexivos realizados, distribuídos entre as condições experimentais do estudo. Abaixo, no Gráfico 41, podemos ver a frequência relativa das ocorrências de deflexão na variável independente de Grupo.

**Gráfico 41 – Frequência de casos de deflexão por Grupo**



Fonte: Elaborado pelo autor.

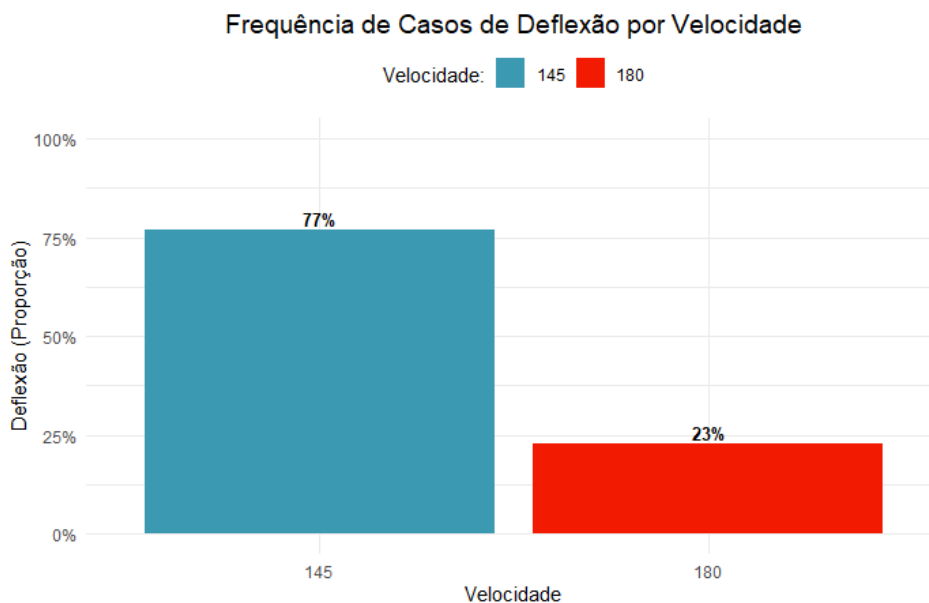
De acordo com os dados encontrados, há uma diferença nos níveis entre os participantes ouvintes e surdos no âmbito da variável de Grupo. Enquanto os ouvintes realizaram um número maior de movimentos deflexivos (82 ocorrências ou 59% do total), os surdos realizaram 57 deflexões (41% do total).

**Gráfico 42 – Frequência de casos de deflexão por Número de Linhas**

Fonte: Elaborado pelo autor.

O Gráfico 42 acima mostra que houve um número maior de movimentos deflexivos em legendas de 2 linhas (93 ocorrências ou 66,9% do total) em relação ao número em legendas de 1 linha (46 ocorrências ou 33,1% dos casos). Contudo, levando em conta a proporção por número de linhas, as legendas de 1 e 2 linhas tiveram um número bem parecido (33,1% e 33,45%, respectivamente).

Abaixo, no Gráfico 43, temos a frequência relativa de casos de deflexão por Velocidade.

**Gráfico 43 – Frequência de casos de deflexão por Velocidade**

Fonte: Elaborado pelo autor.

Em relação à variável Velocidade, podemos observar que houve uma discrepância aparentemente significativa entre os seus níveis. Enquanto as legendas com 180 ppm apresentaram 32 ocorrências de deflexão (23%), as de 145 ppm tiveram 107 movimentos deflexivos (77%), uma diferença considerável.

Abaixo, na Tabela 44, podemos ver a distribuição dos 139 casos de deflexão entre as variáveis e condições experimentais do estudo.

**Tabela 44 – Ocorrências de deflexão por condição experimental**

Grupo	Linha	Velocidade	Frequência	
			Absoluta	Relativa
<b>Ouvinte</b>	1	145	20	14,39
	1	180	4	2,88
	2	145	41	29,50
	2	180	17	12,23
<b>Surdo</b>	1	145	15	10,79
	1	180	7	5,03
	2	145	31	22,30
	2	180	4	2,88

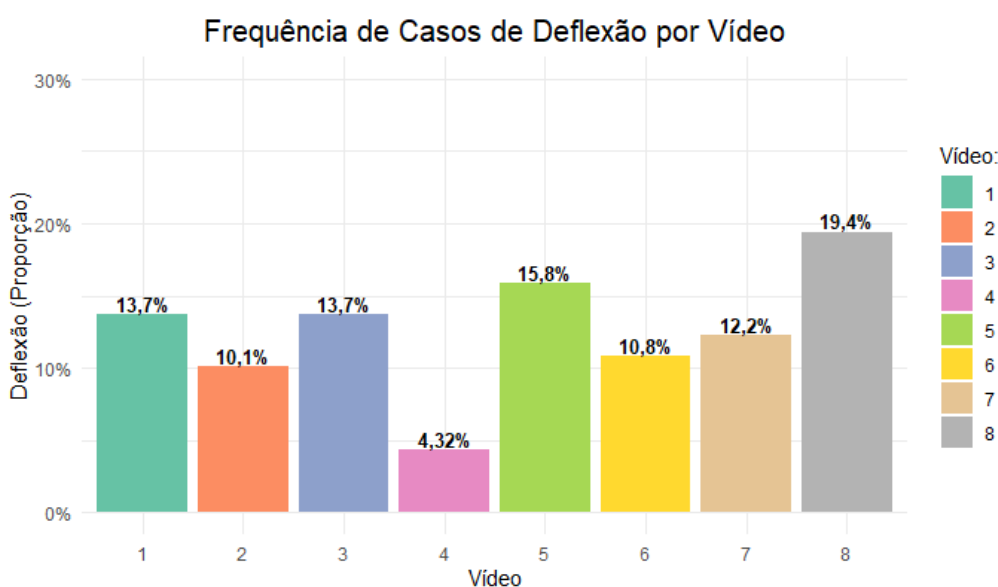
Fonte: Elaborada pelo autor.

A partir dos dados da Tabela 44, podemos observar que, dentre os ouvintes, as duas condições experimentais com 145 ppm se destacaram com o maior número de ocorrências de

deflexão. Enquanto a condição experimental 1L145 apresentou 20 casos de deflexão (14,39% do total), a de 2L145 apresentou 41 movimentos deflexivos (29,5% do total), sendo a categoria com maior número de casos entre todas. Os surdos, da mesma forma, também apresentaram um maior número de deflexões nas condições experimentais com 145 ppm: 15 ocorrências na condição 1L145 (10,79% do total) e 31 ocorrências na condição 2L145 (22,3% do total). Em ambos os grupos, foi na condição experimental 2L145 que ocorreu o maior número de movimentos deflexivos.

Em relação ao número de deflexões por vídeo utilizado no experimento, o Gráfico 44 traz as proporções dessa ocorrência.

**Gráfico 44 – Frequência relativa de casos de deflexão por vídeo**



Fonte: Elaborado pelo autor.

Como é possível observar no Gráfico 44, o Vídeo 8 apresentou o maior número de ocorrências (27 ou 19,4% do total), novamente indicando uma certa dificuldade dos participantes no processamento da deflexão. Logo após, em ordem decrescente, o Vídeo 5 apresentou 22 ocorrências de deflexão (15,8% do total), seguido dos Vídeos 1 e 3 (cada um com 19 ocorrências ou 13,7% do total), do 7, com 17 ocorrências (12,2% do total), do 6, com 15 ocorrências (10,8% do total), do 2, com 14 ocorrências (10,1%) e do 4, com apenas 6 ocorrências (4,32% do total).

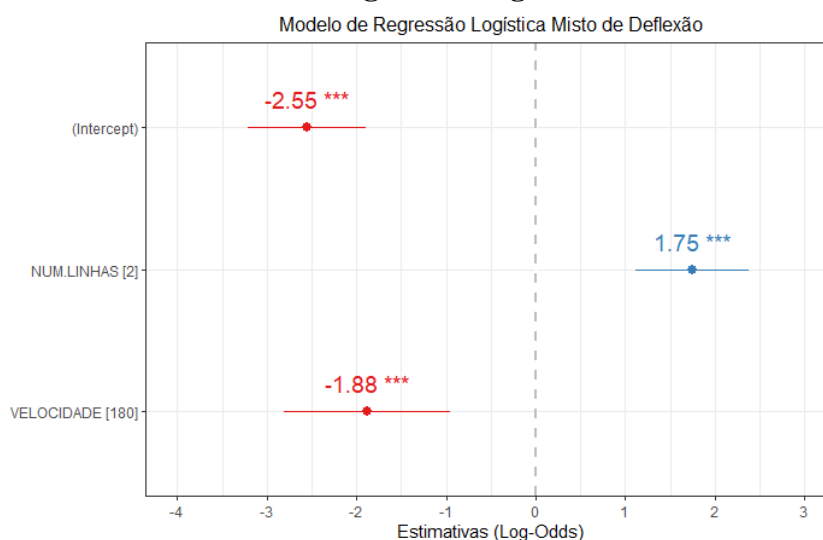
Diferentemente do que fizemos em outras variáveis ao investigar os efeitos de vídeo enquanto possível variável de confusão, decidimos não fazer isso com a variável de deflexão por dois motivos: 1) o número já reduzido de dados (apenas 139 movimentos deflexivos); e 2)

o fato de haver uma relação esperada entre o tipo de imagem em tela e essa movimentação de ida e volta entre áreas de interesse (área da imagem e da legenda).

A partir desses dados, rodamos um modelo de regressão logística misto tendo as variáveis de Grupo, Número de Linhas e Velocidade como preditores (efeitos fixos) e os interceptos aleatórios de participante e vídeo (item), assim como os *slopes* aleatórios de Número de Linhas e Velocidade por Participante. O contraste dos níveis dos efeitos fixos foi feito por meio de *dummy coding*, tendo como nível de referência (em uma ordem alfabética e numérica) as variáveis de 1 linha (Número de Linhas), 145 ppm (Velocidade) e Ouvinte (Grupo), conforme já fora mencionado anteriormente. Esse primeiro modelo que convergiu apresentou significância no intercepto ( $\beta = -2,3907$ ,  $p < 0,0001$ ), Número de Linhas ( $\beta = 1,8171$ ,  $p < 0,0001$ ) e Velocidade ( $\beta = -1,8378$ ,  $p < 0,0001$ ), mas não em Grupo ( $\beta = -0,4300$ ,  $p = 0,314$ ).

Dessa forma, rodamos outro modelo sem a variável Grupo, com os mesmos interceptos e *slopes* aleatórios supracitados, juntamente às variáveis independentes de Número de Linhas e Velocidade. Esse modelo retornou significância para o intercepto ( $\beta = -2,5536$ ,  $p < 0,0001$ ), Número de Linhas ( $\beta = 1,7483$ ) e Velocidade ( $\beta = -1,8791$ ,  $p < 0,0001$ ). Uma comparação entre modelos aninhados mostrou que as variáveis de Número de Linhas e Velocidade explicam melhor a ocorrência de deflexão no processamento de legendas do que um modelo sem preditores ( $X^2 = 30,646$ ,  $p < 0,0001$ ). No Gráfico 45, abaixo, podemos ver uma representação gráfica dos efeitos das variáveis em questão.

**Gráfico 45 – Modelo de Regressão Logística Misto de Deflexão**



Fonte: Elaborado pelo autor.

Na Tabela 45, abaixo, podemos ver os valores previstos pelo modelo para surdos e ouvintes.

**Tabela 45 – Valores previstos para a ocorrência de deflexão**

Linhas	Velocidade	<i>Log-odds</i>	<i>Odds ratio</i>	Probabilidade
1	145	-2,5536	0,0778	0,0721
1	180	-4,4333	0,0118	0,0117
2	145	-0,8053	0,4469	0,3088
2	180	-2,6844	0,0682	0,0639

Fonte: Elaborada pelo autor.

Como havíamos discutido anteriormente, a expectativa era de que as legendas de 2 linhas realmente tivessem mais ocorrências de deflexão e, dessa forma, fosse significativo em relação às de 1 linha, dado que foram expostas pelo dobro do tempo. De acordo com a manipulação das legendas, 1 inserção de legenda de 2 linhas corresponde a 2 inserções de 1 linha. Dessa forma, seria normal se as de 2 linhas (2L145 e 2L180) tivessem o dobro de probabilidade de ocorrências de deflexão.

Nesse sentido, a condição experimental com a menor probabilidade de ocorrência de deflexão foi a de 1 linha e 180 ppm, com probabilidade de 0,0117 (1,17 %), menor também proporcionalmente do que a condição correlata a esta em termos de velocidade, a de 2 linhas e 180 ppm, que apresentou uma probabilidade de 0,0639 (6,39%) de ocorrência de deflexão, um pouco maior do que seria esperado em termos de proporção (3,19%) quando comparado com 1L180. As condições experimentais de 145 ppm, por sua vez, apresentaram probabilidade maior de deflexão em comparação com as de 180. Em relação a isto, as legendas de 1 linha e 145 ppm apresentaram uma probabilidade de ocorrência de 0,0721 (7,21%), enquanto as de 2 linhas e 145 ppm apresentaram uma probabilidade de 0,3088 (30,88%); 2L145, inclusive, teve a maior probabilidade de ocorrência de deflexão entre as condições experimentais. Além disso, 2L145 também teve, proporcionalmente, uma probabilidade maior de ocorrência de deflexão do que 1L145, considerando a metade da probabilidade de 2L145 (15,44%). Esses resultados serão analisados à luz da teoria sobre o processamento de legendas mais à frente, durante a discussão dos dados.

#### 4.1.7 Latência

A latência diz respeito ao tempo de intervalo entre a mudança de atenção (através da movimentação ocular) e as áreas de interesse. No caso da legendagem, diz respeito ao tempo que o espectador leva para sair da imagem e ir para a legenda a partir do momento em que esta aparece. Diferentemente das outras medidas analisadas até aqui, essa variável resposta diz respeito à alocação de atenção, e não ao processamento leitor *per se*. Contudo, pensamos que é uma variável interessante na análise do Número de Linhas, principalmente, visto que alguns dos principais estudos que se dedicaram à investigação desse parâmetro (D'YDEWALLE *et al.*, 1987; D'YDEWALLE; DE BRUYCKER, 2007). Segundo esses autores, quanto mais tempo os espectadores demoram para fixar na área da legenda, pior, porque o menor tempo para ler a legenda em tela pode significar a perda de conteúdo linguístico.

Para que pudéssemos identificar esse tempo de latência, fizemos a subtração entre o momento em que a legenda aparecia em tela durante o experimento e o tempo que o participante demorava até sair da imagem para ir à área da legenda. Sendo assim, essa primeira parte de inspeção da movimentação ocular dos participantes identificou 736 observações em relação a essa variável. Diferentemente do que ocorre nos estudos de movimentação ocular em relação às fixações, os quais normalmente consideram como válidas fixações entre 80 e 800 milissegundos, não são claros na literatura os tempos mínimo e máximo a serem considerados para uma latência válida entre imagem e legenda. Dessa forma, inicialmente não retiramos nenhuma observação de latência do conjunto de dados, fazendo isso apenas posteriormente para a análise estatística.

Nas Tabelas 46, 47 e 48, podemos ver a relação entre as variáveis independentes do estudo, a duração média de latência em milissegundos e outras estatísticas descritivas (desvio padrão, erro padrão e mediana).

**Tabela 46 – Duração média de latência por Grupo**

Grupo	Duração de Latência (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>Ouvinte</b>	369,49	161,21	8,44	339,5
<b>Surdo</b>	395,75	194,66	10,09	353,5

Fonte: Elaborada pelo autor.

**Tabela 47 – Duração média de latência por Número de Linhas**

Número de Linhas	Duração de Latência (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>1</b>	389,39	166,43	7,72	362
<b>2</b>	371,46	199,06	12,06	330,5

Fonte: Elaborada pelo autor.

**Tabela 48 – Duração média de latência por Velocidade**

Velocidade	Duração de Latência (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>145</b>	394,61	203,15	10,95	361,5
<b>180</b>	372,37	154,82	7,81	340,5

Fonte: Elaborada pelo autor.

As estatísticas descritivas das tabelas acima sugerem que nem todas as nossas hipóteses foram confirmadas: 1) a duração média de latência realmente foi maior em 1 linha (389,39 ms) do que em 2 linhas (371,46 ms), não apenas proporcionalmente, mas em números absolutos; 2) a duração média de latência em legendas de 145 ppm (394,61 ms) foi bem maior do que nas legendas de 180 ppm (372,37 ms); e 3) a duração média de encapsulamento foi bem maior no grupo de surdos (395,75 ms) do que no de ouvintes (369,49 ms).

Em relação ao cruzamento das condições experimentais (1L145, 1L180, 2L145 e 2L180), a Tabela 49 abaixo traz a duração média de latência pelo grupo de participantes surdos e ouvintes.

**Tabela 49 – Duração média de latência por condição experimental em surdos e ouvintes**

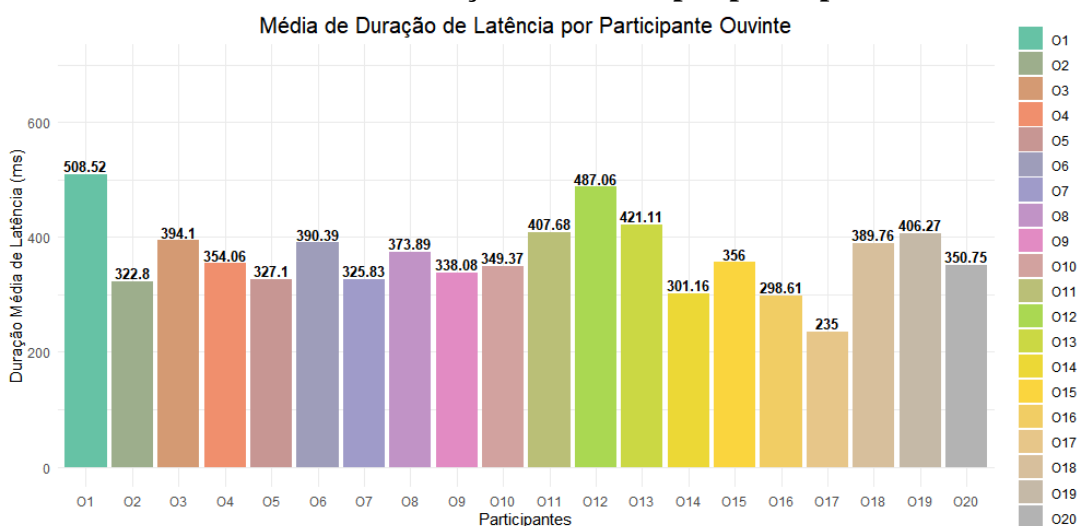
Grupo	Condição Experimental	Duração de Latência (ms)
<b>Surdo</b>	1L145	433,63
	1L180	375,43
	2L145	396,24
	2L180	377,29
<b>Ouvinte</b>	1L145	389,65
	1L180	365,93
	2L145	338,06
	2L180	372,65

Fonte: Elaborada pelo autor.



Ao observarmos a Tabela acima, podemos observar que, no grupo de surdos, as condições experimentais de 180 ppm (1L180 e 2L180) apresentaram tempos médios de latência menores, inclusive sendo bem próximos um do outro, do que as condições experimentais de 145 ppm (1L145 e 2L145). Dentre estas, a legenda na condição experimental de 1L145 foi a que passou mais tempo para ser fixada (433,63 ms). Por outro lado, entre os participantes ouvintes, a correlação que podemos observar é que a condição 1L145 também foi a que apresentou o maior tempo médio de latência (389,65 ms). Contudo, a condição experimental com a mesma velocidade com 2 linhas (2L145) foi a que apresentou o menor tempo médio de latência nesse grupo de participantes (338,06 ms), sugerindo que a velocidade interfere na recepção dos surdos em relação à latência, mas apenas na variável independente de 1 linha. O Gráfico 46 traz duração média de latência por ouvintes.

**Gráfico 46 – Média de duração de latência por participante ouvinte**



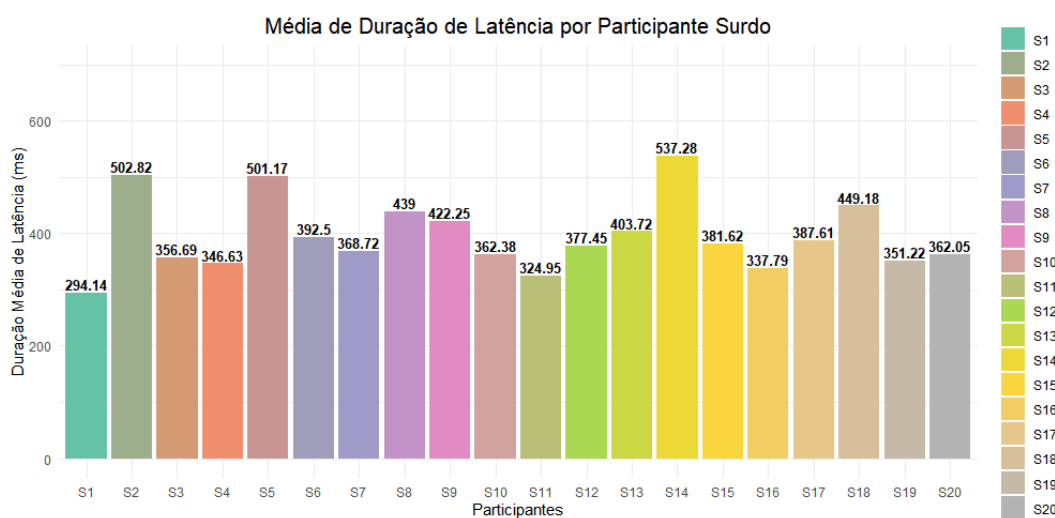
Fonte: Elaborado pelo autor.

Por um lado, de acordo com o Gráfico 46, podemos observar que os participantes O1 e O12 foram os que demoraram mais tempo para acessar a legenda após o início da sua exibição (508,52 ms e 487,06 ms, respectivamente). Apesar de não haver uma correspondência na literatura entre o tempo médio de fixação na legenda e o tempo médio de latência, é interessante observar que o participante O12 foi o que apresentou o maior tempo médio de fixação entre os ouvintes, conforme visto anteriormente (305,9 ms). O participante O1, contudo, não apresentou tendência parecida, visto que foi apenas o 13º participante com maior tempo médio de fixação (260,9 ms). Por outro lado, dentre os participantes que tiveram o menor tempo médio de latência, podemos destacar os dois participantes com tempos menores, O17

(235 ms) e O16 (298,61 ms). Ainda na correlação iniciada acima no parágrafo, O16 foi o que apresentou o 5º menor tempo médio de fixação entre os ouvintes (249,3 ms); O17, por sua vez, foi apenas o 9º menor tempo médio de fixação (267,2 ms).

Os surdos, por sua vez, tiveram suas médias de duração de latência distribuídas da seguinte forma:

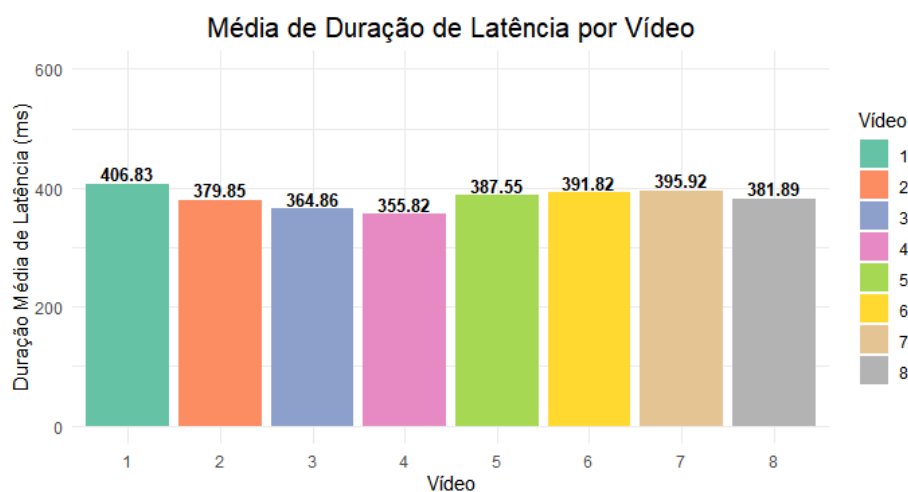
**Gráfico 47 – Média de duração de latência por participante surdo**



Fonte: Elaborado pelo autor.

No Gráfico 47, podemos observar S14 e S2 foram os que tiveram os maiores tempos médios de latência entre os surdos (537,28 ms e 502,82 ms, respectivamente). S2, inclusive, foi o que teve o maior tempo médio de fixação entre os surdos (359,6 ms), enquanto o S14 teve o 8º maior tempo médio de fixação (310,4 ms). No que diz respeito aos menores tempos médios de latência, podemos destacar os participantes S1 (294,14 ms), com um tempo bem abaixo dos outros, e o participante S11 (324,95 ms). Em uma correlação com os tempos médios de fixação dos mesmos participantes, S11 apresentou o 2º menor tempo médio de fixação (261,8 ms) entre os surdos, enquanto o S1 teve o 9º menor tempo médio (303 ms).

No Gráfico 48 podemos observar qual foi o tempo médio de latência para cada vídeo no experimento.

**Gráfico 48 – Média de duração de latência por vídeo**

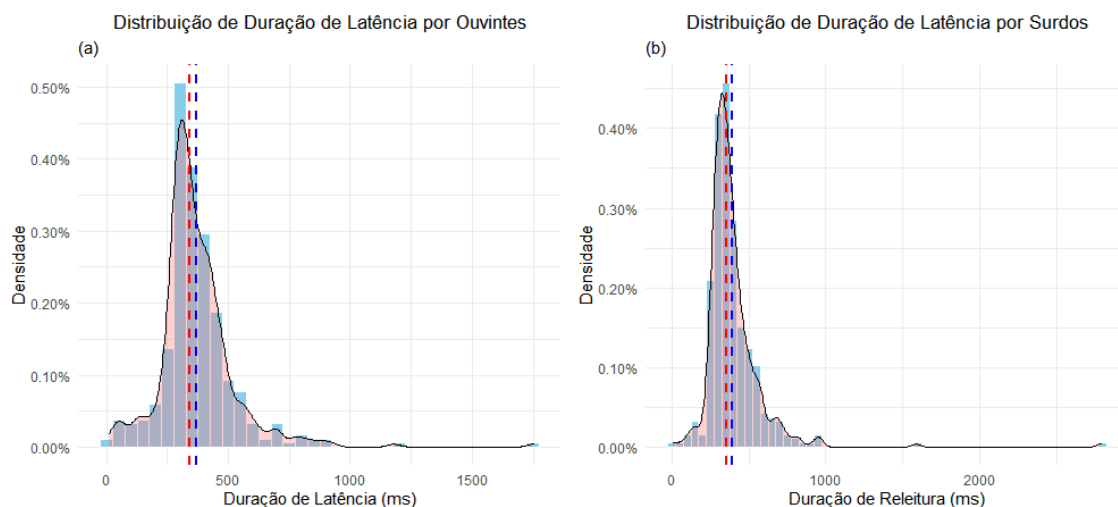
Fonte: Elaborado pelo autor.

Em relação ao que podemos ver no gráfico acima, houve uma certa disparidade entre as médias apresentadas pelos vídeos, principalmente o tempo médio de latência baixo apresentado pelo Vídeo 4 (355,82 ms); quando comparado com o Vídeo 1, que teve o tempo médio mais alto (406,83 ms). Há uma diferença de cerca de 51,01 milissegundos entre os dois. Em uma ordem decrescente, após o Vídeo 1, vem o 7 (395,92 ms), o 6 (391,82 ms), seguido do Vídeo 5 (387,55 ms), do 2 (379,85 ms), do 3 (364,86 ms) e, por último, do 4, já mencionado.

A performance dos participantes no Vídeo 4, especificamente, leva-nos a crer que, com o fato de ser aquele em que a fala é mais espaçada, é provável que essa segmentação retórica influencie o tempo de latência entre imagem e legenda. Dessa maneira, como 1) essa manipulação iria requerer um processo de ajuste de velocidade da fala e legenda bem específico e não foi algo planejado para esta pesquisa e 2) não é algo apontado na literatura em legendagem como algo que possivelmente influencie o tempo de latência, decidimos não retirar nenhum dos vídeos em questão e deixar os efeitos mistos da modelagem de regressão linear serem o único responsável por balancear os valores distintos dos vídeos (itens).

No Gráfico 49 podemos observar a distribuição de ocorrências de latência em dois histogramas relativos a ouvintes e surdos.

### Gráfico 49 – Distribuição de duração de latência entre os participantes

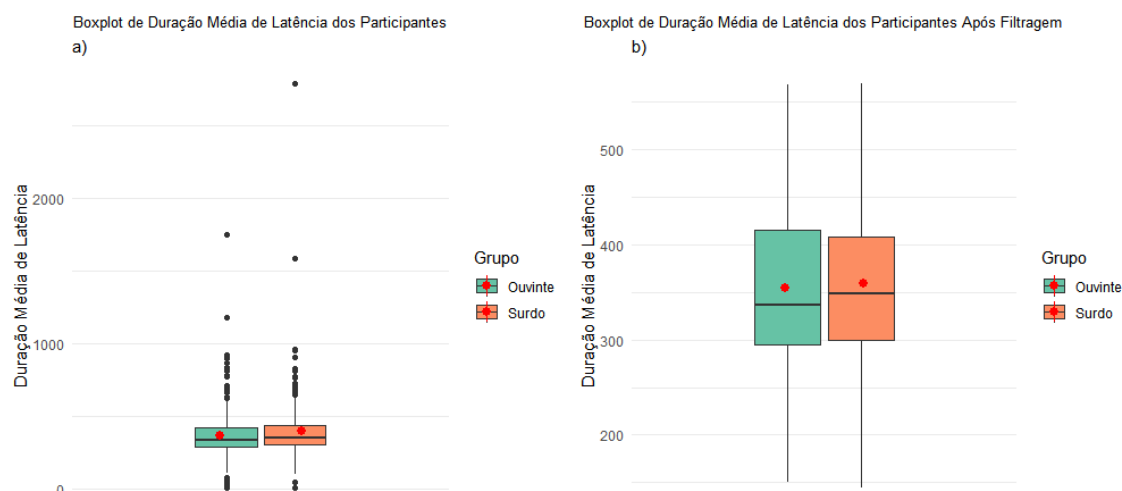


Fonte: Elaborado pelo autor

No Gráfico 49, podemos perceber que a maior parte das ocorrências de latência entre imagem e legenda por parte dos ouvintes ficou entre 300 e 350 milissegundos, enquanto entre os participantes surdos ficou entre 350 e 400 milissegundos. De maneira similar aos outros dados de tipo numérico e contínuo que já discutimos ao longo desta seção, valores altos dos histogramas de ambos os grupos de participantes à direita fazem com que a média (representada pela linha azul) esteja localizada à direita da mediana (representada pela linha vermelha), sinalizando mais uma vez uma distribuição não normal dos dados. Contudo, uma diferença entre os dados numéricos de latência é que não estabelecemos um limite para a validade dos dados – diferentemente do limite de 600-800 milissegundos que utilizamos para as outras medidas de fixação – visto que, como enfatizado anteriormente, não há um registro para isso na literatura em Legendagem e Rastreamento Ocular. Isso fez com que valores muito altos, como de 1750 milissegundos (ouvintes) e 2876 milissegundos (surdos), deixassem as caudas dos histogramas bem alongada para a direita.

Sendo assim, para decidir o que retirar ou não dos dados para a análise estatística, recorreremos novamente à distribuição das ocorrências de latência entre ouvintes e surdos através de *boxplot* para o estabelecimento de *outliers*. O Gráfico 50 tem, à esquerda, um *boxplot* sem a filtragem e, à direita, um *boxplot* com dados acima de 140 milissegundos até 570 milissegundos, recorte este no qual eliminamos os valores fora do limite superior e inferior do *boxplot*.

**Gráfico 50 – Boxplot de duração média de latência dos participantes antes e após a filtragem**



Fonte: Elaborado pelo autor.

Após essa filtragem, restaram 652 observações relativas ao tempo de latência. Dessa forma, outros valores foram utilizados para a análise estatística da influência das variáveis independentes do estudo no fenômeno da duração média de latência. Nesse sentido, as Tabelas 50, 51 e 52 trazem os valores de duração média de latência e outras estatísticas descritivas (desvio padrão, erro padrão e mediana).

**Tabela 50 – Duração média de latência por Número de Linha após filtragem**

Número de Linhas	Duração de Latência (ms)	Desvio Padrão	Erro Padrão	Mediana
1	366,69	86,53	4,32	355
2	341,95	78,11	4,93	326

Fonte: Elaborada pelo autor.

**Tabela 51 – Duração média de latência por Velocidade após filtragem**

Velocidade	Duração de Latência (ms)	Desvio Padrão	Erro Padrão	Mediana
145	367,48	86,68	5,03	353,5
180	348,58	81,21	4,30	332,5

Fonte: Elaborada pelo autor.

**Tabela 52 – Duração média de latência por Grupo após filtragem**

Grupo	Duração de Latência (ms)	Desvio Padrão	Erro Padrão	Mediana
<b>Ouvinte</b>	354,52	84,57	4,72	337
<b>Surdo</b>	359,73	83,88	4,61	348

Fonte: Elaborada pelo autor.

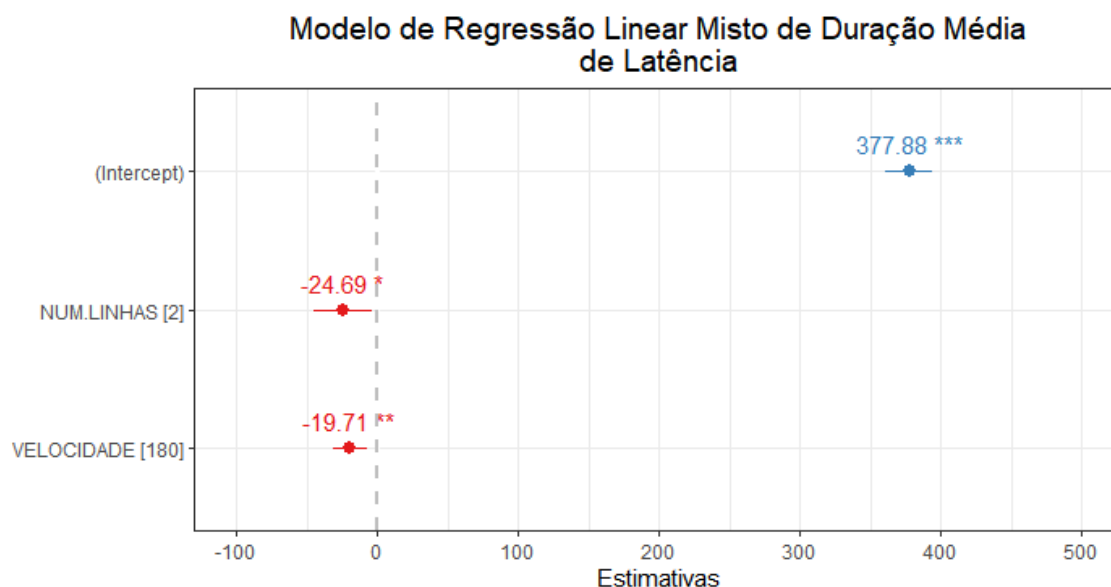
Utilizando os dados acima, rodamos uma Análise de Variância (ANOVA) tendo Grupo, Número de Linhas e Velocidade como preditores. Os resultados mostraram que Grupo não foi significativo ( $f(1,648) = 0,771$ ,  $p = 0,3801$ ), enquanto Número de Linhas ( $f(1,648) = 13,743$ ,  $p = 0,0002$ ) e Velocidade ( $f(1,648) = 9,108$ ,  $p = 0,0026$ ) foram significativos. Conforme os resultados, Número de Linhas esteve em consonância com as hipóteses do trabalho de acordo com a ANOVA, já Grupo e Velocidade contrariaram nossas expectativas.

Além da ANOVA, rodamos um modelo de regressão linear misto de modo a ajustar os valores encontrados no experimento para uma população maior de participantes. Assim, ajustamos esse modelo tendo Grupo, Número de Linhas e Velocidade como efeitos fixos, interceptos aleatórios de participante e vídeo (item) e *slope* aleatório de Número de Linhas por vídeo. Assim como nos modelos anteriores, o contraste dos níveis relativos aos efeitos fixos foi feito através de *dummy coding*, tendo como nível de referência (em uma ordem alfabética e numérica) as variáveis de 1 linha, 145 ppm e Ouvinte (Grupo). Como Grupo não foi significativo nesse modelo ( $\beta = 7,241$ ,  $p = 0,5187$ ), ajustamos ainda mais o modelo ao retirarmos essa variável. Esse modelo mais ajustado indicou que Número de Linhas ( $\beta = -24,709$ ,  $p = 0,0533$ ) foi marginalmente significativo e Velocidade ( $\beta = -19,752$ ,  $p = 0,0011$ ) foi significativo. Esse mesmo modelo teve como  $R^2$  marginal 0,032 e como  $R^2$  condicional 0,196; ou seja, com a adição de efeitos mistos, as variáveis de Número de Linhas e Velocidade conseguem explicar 19,6% do fenômeno da latência, um efeito não tão forte.

Dada a proximidade da significância de Número de Linhas com o 95% de intervalo de confiança que estamos adotando ao longo do trabalho, decidimos utilizar esse coeficiente para a previsão de valores do modelo. Além disso, uma comparação entre modelos aninhados mostrou que o modelo em questão apresenta uma diferença significativa quando comparado com um modelo sem variável preditora, apenas com intercepto ( $X^2 = 14,778$ ,  $p = 0,0006$ ).

No Gráfico 51, é possível visualizar os valores dos efeitos fixos supracitados e o valor do intercepto (valor quando os níveis das variáveis estão, em ordem alfabética ou numérica, em 1 linha e na velocidade de 145 ppm).

Gráfico 51 – Modelo de Regressão Linear Misto de Duração Média de Latência

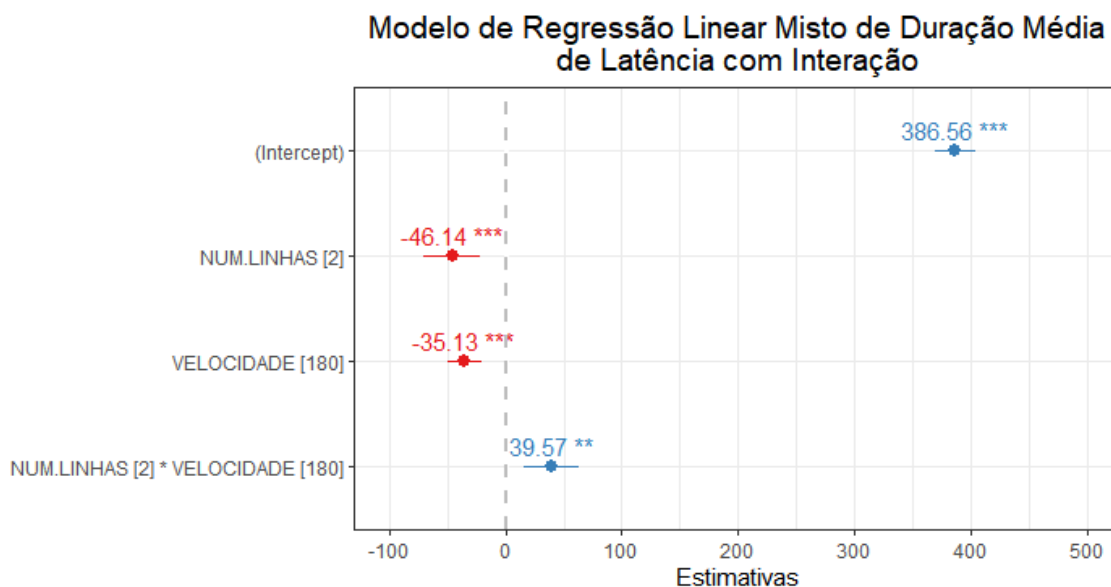


Fonte: Elaborado pelo autor.

Contudo, um outro modelo que checou a interação entre as variáveis mostrou uma interação significativa *Número de Linhas:Velocidade*. Esse novo modelo, tendo os mesmos efeitos fixos e mistos do modelo anterior com a adição da interação entre Número de Linhas e Velocidade, também apresentou significância em Número de Linhas ( $\beta = -46,141$ ,  $p = 0,0023$ ), Velocidade ( $\beta = -35.131$ ,  $p < 0,0001$ ) e na interação *Número de Linhas:Velocidade* ( $\beta = 39,573$ ,  $p = 0,0013$ ), este último coeficiente sendo específico para Número de Linhas no nível de 2 linhas e Velocidade no nível de 180 ppm. Além disso, a escolha por esse modelo ao invés do anterior se justificou por ele ter se mostrado significativo em uma comparação com um modelo apenas com intercepto ( $X^2 = 25,051$ ,  $p < 0,0001$ ) e também quando comparado com o modelo anterior sem interação ( $X^2 = 10,273$ ,  $p = 0,0013$ ), tendo ainda apresentado  $R^2$  marginal e  $R^2$  condicional melhores (0,045 e 0,209); ou seja, apesar de ainda se tratar de efeito pequeno sobre o fenômeno, o modelo com interação entre as variáveis de Número de Linhas e Velocidade consegue explicar 20,9% do fenômeno da latência, um pouco maior do que o efeito anterior.

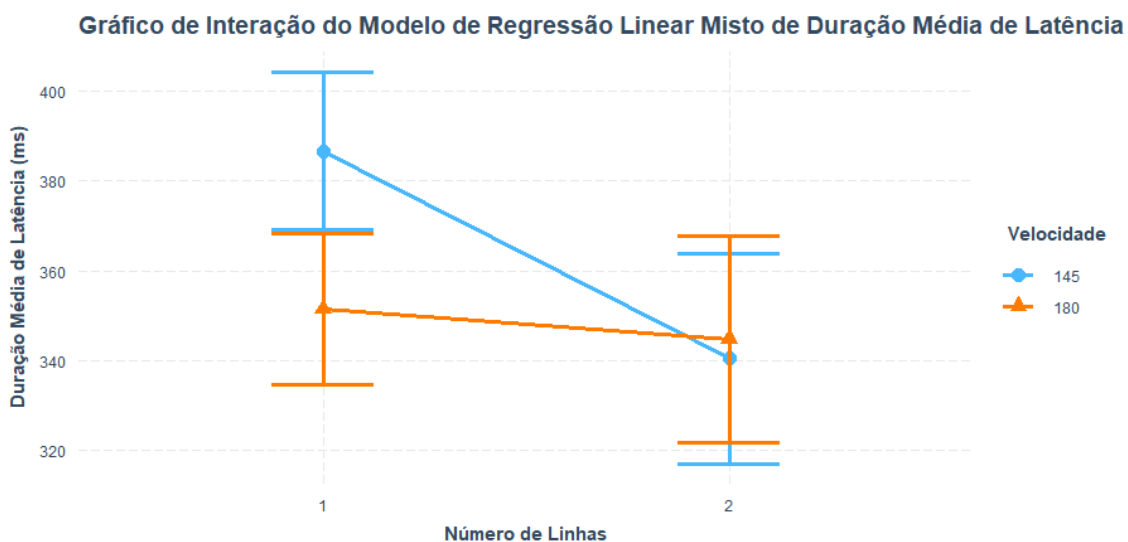
Abaixo, nos Gráficos 52 e 53, podemos ver os efeitos do modelo em questão e a interação entre as variáveis independentes, respectivamente.

**Gráfico 52 – Modelo de Regressão Linear Misto de Duração Média de Latência com Interação**



Fonte: Elaborado pelo autor.

**Gráfico 53 – Interação do Modelo de Regressão Linear Misto de Duração Média de Latência**



Fonte: Elaborado pelo autor.

A partir do Gráfico 53, podemos entender melhor como funciona essa interação entre as variáveis. No gráfico é possível perceber que, entre 1 e 2 linhas, não há uma diferença significativa no tempo médio de latência na velocidade de 180 ppm. Entretanto, essa diferença



entre níveis em Número de Linhas é significativa na velocidade de 145 ppm, tendo uma redução de tempo médio de latência de 1 para 2 linhas.

Dessa maneira, tendo em vista que não houve diferença significativa entre ouvintes e surdos nessa variável resposta, podemos prever, levando em consideração os efeitos principais e a interação do modelo, os valores de latência média nas condições experimentais de 1L145, 1L180, 2L145 e 2L180 para surdos e ouvintes, conforme podem ser vistos na Tabela 53.

**Tabela 53 – Valores previstos para duração média de latência a partir do modelo**

Linha	Velocidade	Duração de Latência (ms)
1	145	386,56
1	180	351,43
2	145	340,20
2	180	344,86

Fonte: Elaborado pelo autor.

De acordo com os valores apresentados na Tabela 53, as duas condições experimentais com 2 linhas (2L145 e 2L180) apresentaram os menores valores do modelo (340,2 ms e 344,86 ms, respectivamente). As condições experimentais de 1 linha (1L145 e 1L180), por sua vez, apresentaram os maiores valores de tempo médio de latência (386,56 ms e 351,43 ms), sendo, inclusive, uma diferença significativa, como é possível observar no Gráfico 53 de interação, em que os intervalos de confiança entre 1L145 e 1L180 não se tocam.

O que esse tempo bem maior na condição experimental 1L145 significa para o processamento de legendas? De maneira prática, significa que o espectador demora muito mais a acessar o texto da legenda e, por isso, pode acabar perdendo informação relevante da legenda. Duas legendas seguidas de 1 linha e 145 ppm, por exemplo, levariam 773,12 milissegundos para serem acessadas, ao passo que uma legenda de 2 linhas e 145 ppm (contendo a mesma quantidade de texto que as duas legendas de 1 linha na mesma velocidade) levaria 340,2 milissegundos para ser acessada. Dessa forma, a legenda de 2 linhas seria mais confortável (em ambas as velocidades, visto que a legenda de 1 linha e 180 ppm também tem um tempo médio de latência maior do que a de 2 linhas e 180 ppm) de ser processada, pois haveria tempo suficiente para a sua leitura e, após finalizada, o leitor poderia ir para a imagem e, quando aparecesse outra legenda de 2 linhas, voltar à legenda, ou resolver alguma ambiguidade lexical através de regressão e releitura. A discussão teórica desses resultados será pauta da próxima seção.

## 4.2 Discussão dos resultados

Nesta subseção, iremos retomar os resultados das medidas de rastreamento ocular ao longo desta seção – duração média de fixação, número de fixações por legenda, duração média e ocorrência de regressão, duração média e ocorrência de releitura, deflexão, duração média de encapsulamento e duração média de latência – relacionando-os com as duas hipóteses do trabalho: a) Legendas de 1 linha induzem um processamento menos linear e, dessa forma, mais custoso do que legendas de 2 linhas; b) Legendas rápidas (180 palavras por minuto) e lentas (145 palavras por minuto) não são significativamente diferentes no processamento de surdos e ouvintes. Como várias medidas foram analisadas e serão dispostas ao longo desta seção, ao final fazemos um balanço de como o resultado de cada uma dessas medidas contribui para que as hipóteses do trabalho sejam confirmadas ou refutadas.

Os resultados de cada medida apresentada anteriormente são resumidos a seguir e, em cada medida, discutimos como cada variável independente manipulada – Grupo, Número de Linhas e Velocidade – está relacionada às expectativas do trabalho e à literatura sobre Leitura, Rastreamento Ocular e Tradução Audiovisual-Legendagem que embasa este estudo.

Começamos então com a medida de duração média de fixação, a medida mais relevante para tarefas de processamento de informação de acordo com o postulado de que nossa atenção está focada no processamento foveal. Sobre isso, apesar de ser possível separar atenção e olhar com tarefas de discriminação (POSNER, 1980 *apud* RAYNER, 1998) em um processo complexo como a leitura, já que a conexão entre atenção e mente é bem próxima (RAYNER, 1998). Logo, um tempo maior de duração média de fixação está relacionado a um custo maior de processamento, o que na leitura de textos impressos pode estar relacionado a fatores como vocabulário, sintaxe, dentre outros.

Dada a natureza complexa e dinâmica do produto audiovisual legendado, além de suas diferentes manifestações a partir de características visuais do vídeo legendado e aspectos técnicos desse tipo de tradução, não há uma duração média prevista na literatura em rastreamento ocular para esse tipo de texto levando em consideração a inspeção da movimentação ocular que fizemos aqui. D'Ydewalle *et al.* (1985) estimou um tempo médio de 124 ms por palavra por parte de adultos expostos a vídeos legendados. Contudo, i) essa média é bem baixa e bem abaixo do tempo médio de fixação até de textos impressos, que é de 200 a 250 ms (RAYNER, 1998); e ii) esse tempo estimado de duração de fixação não foi medido diretamente, mas inferido a partir do tempo total gasto na área da legenda dividido pelo número de palavras. Assim, os resultados expostos anteriormente serão aqui contextualizados com

dados sobre a leitura de textos impressos (JUST; CARPENTER, 1980; RAYNER, 1998; REICHLER; RAYNER; POLLATSEK, 2013), os achados sobre o processamento de legendas e movimentação ocular de surdos brasileiros (MONTEIRO, 2016; SOUZA, 2020; VIEIRA, 2016), sobre o processamento de legendas de 1 e 2 linhas por ouvintes (D'YDEWALLE *et al.*, 1987; D'YDEWALLE *et al.*, 1991; D'YDEWALLE; DE BRUYCKER, 2007; PRAET *et al.*, 1990) e sobre as velocidades de legenda de 145 e 180 ppm por surdos e ouvintes brasileiros (MONTEIRO, 2016; SOUZA, 2020; VIEIRA, 2016).

Assim como era esperado, os surdos apresentaram uma média maior de fixação juntando todas as condições experimentais (299,07 ms) do que os ouvintes (263,17 ms), contando com as 4.038 observações sobre fixação de até 600 milissegundos do experimento. Essa diferença foi confirmada estatisticamente pela ANOVA e pelo Modelo de Regressão Linear Misto. Apesar de todos os surdos que participaram do estudo cursarem o nível superior ou já serem graduados, o que os coloca em contato com a língua portuguesa em níveis acadêmicos, são usuários da língua portuguesa como L2 e, dessa forma, apresentam uma duração média maior de fixação (RAYNER; LIVERSEDGE, 2004; RAYNER; BARBARA; POLLATSEK, 2013).

Essa tendência confirmou dados encontrados por Vieira (2016) e Monteiro (2016) com documentários e propagandas políticas, respectivamente, expostos a surdos e ouvintes brasileiros. Apesar de ambas as pesquisas terem utilizado condições experimentais relativas à segmentação – boa e má<sup>81</sup> – e velocidade (145 e 180 ppm), encontramos médias bem parecidas. Por um lado, Vieira (2016), ao analisar legendas de 2 linhas bem segmentadas de um documentário, observou uma média de 258,44 ms para ouvintes e 284,24 ms para surdos, números aproximados com o que identificamos na presente pesquisa. Por outro lado, Monteiro (2016), ao analisar legendas de 1 linha – característica do gênero – de legendas bem segmentadas de propagandas políticas, observou médias ligeiramente maiores do que Vieira (2016) em ambos os grupos de participantes: 261,24 ms para ouvintes e 292,5 ms para surdos. Essa diferença pode se justificar pelos gêneros distintos – documentário e propaganda política, mas também há que se considerar que as legendas de Monteiro (2016) foram todas de 1 linha, o que, de acordo com a nossa hipótese alternativa e os resultados de nossa pesquisa, pode proporcionar um maior custo de processamento para os espectadores. Além disso, pode se justificar pelo fato de Monteiro (2016) ter utilizado uma metodologia de análise das fixações

---

<sup>81</sup> Para efeitos de comparação, entretanto, só utilizamos as médias de duração de fixação relativas às condições com boa segmentação, visto que também pautamos as legendas manipuladas no experimento com esse parâmetro linguístico.

diferente de Vieira (2016). Enquanto Vieira (2016) fez um acompanhamento do movimento do olhar dos participantes de modo a verificar quais fixações realmente constituíam processos de leitura, Monteiro (2016) utilizou a análise automática<sup>82</sup> do *Tobii Studio*.

Ao considerar os dados desta pesquisa e de Monteiro (2016) e Vieira (2016) com surdos e ouvintes brasileiros expostos a vídeos legendados de diferentes gêneros, podemos estabelecer uma margem de 250 a 270 milissegundos para ouvintes e de 280 a 300 milissegundos para o processamento desse tipo de texto dinâmico por surdos do país. Esses números são maiores do que os sugeridos por Rayner (1998) para a leitura de textos impressos, cerca de 200 a 250 milissegundos<sup>83</sup>. Isso, contudo, não causa surpresa quando analisamos a natureza multimodal e dinâmica do texto legendado, cuja relação com o canal sonoro e imagético (apenas o imagético para surdos e no contexto geral desta pesquisa) é complexa e ainda deve ser objeto de escrutínio, mas estão bem acima de médias observadas por d'Ydewalle *et al.* (1985), por exemplo, de 124 ms.

No que tange à variável de Número de Linhas, nossa hipótese era de que surdos e ouvintes teriam uma média maior de fixação em legendas de 1 linha do que 2 linhas, o que foi confirmado. Enquanto as legendas de 1 apresentaram média de 283,24 ms, as legendas de 2 linhas tiveram em média fixações de 277,07 ms considerando ambos os grupos de participantes. Separando por Grupo, os ouvintes tiveram médias de fixação de 267,08 ms em legendas de 1 linha e 259,65 ms em legendas de 2 linhas. Os participantes surdos, por sua vez, também tiveram um maior custo de processamento em legendas de 1 linha (300,79 ms) do que 2 linhas (297,41 ms), mas uma diferença menor. O resultado sobre a duração das fixações nas legendas de 1 linha foi próximo do encontrado por Monteiro (261,24 ms para ouvintes e 292,5 ms para surdos – apesar de um pouco maiores), que testou com vídeos com tempos de 1 minuto.

Estatisticamente, ao perceber que os vídeos (estímulos experimentais da pesquisa) consistiam uma possível variável de confusão, realizamos testes estatísticos com e sem os vídeos em questão. Enquanto os testes de Análise de Variância (ANOVA) flutuaram entre marginalmente significantes e não significantes em relação a um maior tempo de duração média de fixação em legendas de 1 linha, o Teste T Unicaudal – especificando a hipótese alternativa

---

<sup>82</sup> Souza (2020) analisou os dados de Monteiro (2016) de maneira semelhante a Vieira (2016). Além disso, a autora ajustou um modelo de regressão linear misto com os dados em questão, assim como fizemos aqui. Apesar de as estimativas para duração média de fixação em milissegundos em Souza (2020) não estarem claras, e por isso não serão reportadas aqui, a autora aponta que a duração da fixação cai quando a segmentação está boa em ouvintes e surdos. Além disso, a autora observou que surdos e ouvintes realizaram fixação mais curtas na condição rápida (180 ppm) e bem segmentada.

<sup>83</sup> Conforme já apontado anteriormente, a leitura silenciosa, estimada em uma média de 225 milissegundos, seria o tipo de atividade de leitura de textos impressos apontado por Rayner (1998) mais semelhante à tarefa de ler legendas.

de que legendas de 1 linha teriam uma média maior e não simplesmente que as médias dos dois níveis da variável seriam diferentes, foi significativo tanto com todos os vídeos quanto sem os vídeos perturbadores. Além disso, conseguimos convergir um modelo ajustado de Regressão Linear Misto sem os vídeos perturbadores, a partir do qual estimamos um efeito de 7,51 milissegundos de redução da duração média de fixação quando passamos do nível de 1 linha para 2 linhas de legenda nos dois grupos. Esse modelo, cujo  $R^2$  condicional estimou como explicativo de 9,4% do fenômeno de duração média de fixação, percentual baixo de explicação, estimou uma duração média de fixação de 260,1 ms (ouvintes) e 294,05 ms (surdos) em legendas de 1 linha e 252,59 ms (ouvintes) e 286,54 ms (surdos) em legendas de 2 linhas.

Esses números diferem das médias de fixação encontradas em d'Ydewalle e De Bruycker (2007) em adultos ( $n = 12$ ) e crianças ( $n = 8$ ) holandesas processando um filme em desenho animado sueco em duas condições: i) com legendas em sueco e áudio em holandês (legendagem padrão); e ii) legendas em holandês e áudio em sueco (legendagem reversa). Os dados do trabalho em questão mostraram que na condição reversa (legenda na L1), houve significância na diferença de média de fixação entre legendas de 2 linhas (212 ms) e legendas de 1 linha (242 ms), mas não foi significativa na legendagem padrão (210 ms e 215 ms, respectivamente). Além disso, entre adultos não houve significância (182 ms e 190 ms), mas o resultado foi diferente em crianças com as legendas de 2 linhas sendo mais custosas do que de 1 linha (241 ms e 268 ms).

Outra medida utilizada por d'Ydewalle e De Bruycker (2007), assim como por d'Ydewalle *et al.* (1987) e d'Ydewalle *et al.* (1991), como sinônimo de custo de processamento foi o tempo proporcional gasto na área da legenda nas condições de 1 e 2 linhas. Nesse sentido, para d'Ydewalle *et al.* (1987) a diferença entre 2 linhas e 1 linha em relação ao tempo gasto na área da legenda era semelhante na regra de 6 segundos de exposição na tela, mas quando em 2 e 4 segundos, tempo de exposição menor e, conseqüentemente, velocidades de legenda mais rápidas, os participantes se mantiveram mais tempo na área da legenda. D'Ydewalle *et al.* (1991), por sua vez, observaram que participantes ouvintes americanos lendo legendas intralinguísticas passaram mais tempo lendo legendas de 2 linhas (21,87%) do que legendas de 1 linha (16,37%); d'Ydewalle e De Bruycker (2007), da mesma forma, constataram essa diferença na mesma direção, com participantes passando 37% do tempo lendo legenda de 2 linhas (37%) do que 1 linha (31%).

Não utilizamos essa medida de tempo gasto na área da legenda como indicador de conforto na recepção de legendas em nosso estudo por ser uma medida mais de alocação de atenção do que processamento propriamente dito. Além disso, nem toda fixação realizada na

área da legenda significa realmente um processo de leitura, o que o cálculo de tempo gasto na área da legenda não leva em consideração, dificultando a contextualização dos dados dessas pesquisas anteriores com a medida de duração média de fixação aqui apresentada.

Em relação à variável de Velocidade e seus dois níveis (145 e 180 ppm) apresentados no estudo, não houve diferença significativa, convergindo para a hipótese do trabalho de que não haveria diferença significativa entre legendas nessas duas velocidades. Enquanto as legendas de 145 ppm tiveram como média de fixação 280,17 ms, as legendas de 180 ppm tiveram 279,9 ms, isso considerando ambos os grupos. Em relação aos ouvintes, estes apresentaram uma média de 265,50 ms em 145 ppm, tendo um desempenho ligeiramente melhor no nível de 180 ppm, que foi de 260,61 ms. Os surdos, por sua vez, tiveram média de duração de fixação ligeiramente melhor em 145 ppm (296,36 ms) do que em 180 ppm (302,18 ms) em legendas de 1 e 2 linhas. Os resultados foram bem parecidos com os de Monteiro (2016) em propagandas políticas de 1 linha bem segmentadas em 145 ppm (266,25 ms para ouvintes e 292,5 ms para surdos). Já em 180 ppm, os resultados diferiram um pouco mais. Enquanto os ouvintes tiveram uma média de 256,25 ms, melhor do que em 145 ppm, os surdos tiveram uma média de 292,5 ms, a mesma nos dois níveis de Velocidade.

Vieira (2016) teve uma média próxima do número encontrado na pesquisa na condição lenta (145 ppm) bem segmentada (278 ms), mas na rápida bem segmentada houve uma diferença grande para a nossa média (263,5 ms). Observando a questão por grupo, os surdos do estudo de Vieira apresentaram uma média próxima à do nosso trabalho na condição de 145 ppm (300 ms), mas a condição de 180 ppm bem mais baixa (268 ms), direção oposta dos nossos dados. As médias dos ouvintes, por sua vez, foram praticamente iguais em 145 ppm (257 ms) e 180 ppm (259 ms).

A não significância na variável Velocidade e a falta de interação entre *Grupo:Velocidade* mostram que os dois grupos processam as duas velocidades (145 ppm e 180 ppm) da mesma maneira, assim como previa nossa hipótese alternativa. Sendo assim, nossos achados reforçam os resultados de Araújo e Nascimento (2011) sobre a boa recepção de surdos a vídeos legendados com até 180 ppm. Além disso, apesar de serem similares aos resultados de Monteiro (2016) sobre o processamento de legendas de 1 linha em 145 e 180 ppm por ouvintes e surdos, diferem dos resultados de Souza (2020), que apontaram um tempo menor de fixação para surdos e ouvintes nas legendas rápidas (180 ppm) e bem segmentadas. Os resultados de Souza (2020), inclusive, são similares aos de Vieira (2016), que apontaram uma diferença significativa nos níveis de 145 e 180 ppm por participantes surdos, visto que as legendas de 180 ppm bem segmentadas tiveram um tempo menor de duração média nesse grupo

específico de participantes. O fato de os resultados sobre a variável Velocidade neste trabalho não terem convergido com os achados de Vieira (2016) e Souza (2020) pode ter sido motivado pelo *design* experimental distinto das pesquisas em questão, as quais utilizaram vídeos maiores – de até 1 minuto – e um número menor de participantes surdos (N = 8) e ouvintes (N = 8).

Em relação aos dados encontrados por De Linde e Kay (1999) sobre a questão da Velocidade, nossos resultados vão de encontro ao dos autores, que postularam que, quanto maior a velocidade da legenda, mais rápidas as fixações, ou seja, mais curtas. Apesar de não termos encontrado significância para a variável de Velocidade em nosso estudo, é importante contextualizar o trabalho de De Linde e Kay (1999), que manipulou e testou legendas de velocidades distintas às do nosso trabalho (74 e 139 ppm em vez de 145 e 180 ppm).

Outra medida levada em consideração no nosso estudo é o número de fixações por legenda. Assim como fizemos acima com a duração média de fixação, vamos apresentar os resultados por variável independente – Grupo, Número de Linhas e Velocidade – e correlacioná-los com a teoria.

Nesse sentido, no que tange à variável Grupo, os ouvintes realizaram mais fixações por legenda do que os surdos e houve significância estatística nesta diferença na Análise de Variância (ANOVA). Sem considerar a filtragem dos dados nessa variável, os ouvintes tiveram 5,85 fixações, enquanto os surdos tiveram 5,41 por legenda. Sabemos que o número de fixações encontrado é particular ao contexto da nossa pesquisa, porque há uma relação direta entre o número de fixações e o número de palavras. Portanto, colocamos o número aqui mais para ilustrar os apontamentos e desdobramentos a partir das nossas hipóteses e da teoria.

Vieira (2016) e Monteiro<sup>84</sup> (2016) também observaram o número de fixações na área da legenda por surdos e ouvintes. Todavia, essa medida foi trabalhada pelas autoras de maneira um pouco distinta, visto que também observaram e compararam esse dado com o número de fixação na imagem, algo que não realizamos aqui. Além disso, as autoras utilizaram

---

<sup>84</sup> Utilizando os dados de Monteiro (2016) para análise estatística, Souza (2020) não explicita se surdos ou ouvintes tiveram um maior número de fixações e se foi uma diferença significativa; a autora, contudo, aponta que os surdos fizeram mais fixações nas legendas lentas (145 ppm) e bem segmentadas. Em relação a isso, a interpretação de Souza (2020) para a medida de número de fixações é distinta da do presente trabalho: para ela, quanto menos fixações, melhor. Dessa forma, como a condição rápida (180 ppm) bem segmentada foi a que apresentou o menor número de fixações para ambos os grupos, essa foi a que a autora apontou como a mais confortável. No contexto da legendagem, acreditamos que um menor número de fixações pode ser algo ruim, dada a característica dinâmica desse tipo de texto. Assim, um menor número de fixações pode significar que a legenda não foi lida por completo e conteúdo linguístico foi perdido. Da mesma forma, um número alto de fixações pode significar também que a legenda teve ocorrências de releitura e de regressão, o que seria algo ruim tomando por base os estudos sobre processamento de textos impressos, mesmo que a ocorrência de releitura e regressão em legendagem possa assumir uma concepção distinta (KRUGER; STEYN, 2014), conforme já foi discutido anteriormente. De qualquer forma, é uma medida que deve ser correlacionada com outras para que se mensure a sua influência sobre o processamento de legendas a partir das variáveis do estudo.

o número global de fixações realizadas na área da legenda, enquanto estamos utilizando a média de fixações por legenda. Mesmo assim, os resultados do nosso trabalho reforçaram os dados de Vieira e Monteiro nas condições experimentais com legendas bem segmentadas do seu trabalho, pois as autoras também identificaram um número maior de fixações por ouvintes do que por surdos, estando em consonância com o que falam Rayner e Liversedge (2004) e Rayner, Barbara e Pollatsek (2013) sobre o número maior de fixações e com duração menor por leitores mais proficientes; ou seja, surdos, por serem aprendizes de português como L2, tendem a realizar menos fixações e com duração mais longa.

Já em relação ao Número de Linhas, a ANOVA também apontou diferença estatística entre os níveis da variável (1 e 2 linhas), mas algo que já era naturalmente esperado, dado o tamanho maior de caracteres das legendas de 2 linhas. Em relação às médias dessa variável, as legendas de 1 linha apresentaram 3,69 fixações por legenda, enquanto as de 2 linhas apresentaram 7,46 fixações (ou 3,77 por linha de legenda), sendo ainda ligeiramente maior proporcionalmente do que as fixações em legendas de 1 linha. Essa diferença, entretanto, foi pequena e pode ter sido motivada pelo maior número de ocorrências de regressão e releitura nas legendas de 2 linhas. Dessa forma, um maior número de fixações em legendas de 2 linhas não nos auxilia a confirmar a hipótese de que as legendas de 2 linhas são mais confortáveis.

Como nenhum dos estudos apresentados que tiveram como tema a influência de legendas de 1 e 2 linhas no processamento de surdos e ouvintes utilizou essa medida, não há como fazer uma comparação direta com seus resultados; de qualquer forma, mais adiante, quando apresentarmos os valores previstos pelo Modelo de Regressão Linear Misto para essa variável, discutiremos com mais detalhes a influência do Número de Linhas nessa variável resposta.

Sobre a variável de Velocidade, a ANOVA apontou diferença significativa entre legendas de 145 e 180 ppm, indicando que um tempo maior em tela influencia essa variável dependente. As legendas de 145 ppm propiciariam mais fixações por legenda no nosso estudo com 6,02 fixações, enquanto as legendas de 180 ppm tiveram 5,23 fixações por legenda.

Os dados sobre variável Velocidade e o número de fixações por legenda podem ser relacionados aos dados de De Linde e Kay (1999), Vieira (2016) e Monteiro (2016). De Linde e Kay (1999) indicaram que um estímulo com legendas que eles consideraram de maior velocidade (139 ppm) propiciou um maior número de palavras por fixação do que um estímulo com legendas consideradas de menor velocidade (74 ppm); apesar de ser uma medida diferente do número de fixações por legenda aqui utilizado, são medidas de certa forma similares, visto que o número maior de palavras por fixação significa menos fixações. Destarte, a velocidade



menor também permitiu um número maior de fixações, assim como no nosso estudo. Em relação a Vieira, a autora observou, nas condições com boa segmentação do seu estudo, que os participantes ouvintes e surdos faziam mais fixações na área das legendas nas condições experimentais de 180 ppm do que de 145 ppm, diferentemente do que encontramos aqui. Com metodologia semelhante, mas com uma análise automática da movimentação ocular – e não um acompanhamento do olhar como fizemos aqui e como Vieira (2016) também fez –, Monteiro (2016) observou que os participantes ouvintes fizeram mais fixações na condição de 145 ppm do que na de 180 ppm; os surdos, entretanto, foram na direção oposta, fazendo mais fixações na legenda rápida (180 ppm) do que na lenta (145 ppm).

Essa diferença dos nossos resultados em relação a Monteiro (2016) e Vieira (2016), principalmente, pode estar relacionada a um gênero distinto aqui utilizado (filme em vez de propaganda política e documentário, respectivamente) e de como o imagético esteve presente nos três estudos. É válido ressaltar que a escolha das cenas do experimento foi pautada pela pouca movimentação em cena, assim como pela baixa saliência de aspectos visuais. Isso porque o intuito principal do trabalho estava relacionado ao processamento das legendas em interação com aspectos visuais que fossem de certa forma uniformes e não tirassem a alocação de atenção das legendas. Além disso, essa diferença dos nossos resultados com os de Vieira (2016) e Monteiro (2016) pode ter sido motivada pelo tamanho dos estímulos utilizados. Como os estímulos do presente estudo – de 8 a 15 segundos – foram mais curtos do que os utilizados pelas autoras – cerca de 1 minuto –, acreditamos que os participantes foram inclinados a ter que ler toda a legenda para a compreensão dos excertos, utilizando menos a imagem. Com o tempo maior de vídeo dos trabalhos de Vieira (2016) e Monteiro (2016), é possível que a imagem tenha sido mais utilizada como apoio de maneira mais acentuada. As propagandas políticas utilizadas por Monteiro (2016), inclusive, apresentavam bastante saliência no aspecto visual, dada a grande quantidade de informações imagéticas desse gênero, algo que deve ser considerado ao contrastarmos os nossos dados com os da autora. Como Souza (2020) também utilizou os mesmos dados, essa consideração também vale para os seus resultados.

Em relação ao Modelo de Regressão Linear Misto que convergimos para estimar valores do número de fixações por legenda, este indicou que Grupo e Velocidade não foram significativos para o número de fixações por legenda, diferentemente do que havia sido apontado pela ANOVA. Contudo, além de apontar novamente a significância de Número de Linhas, o modelo apontou uma interação interessante: enquanto a variação entre as condições experimentais de 1 linha (1L145 e 1L180) pouco muda entre surdos e ouvintes, as condições de 2 linhas (2L145 e 2L180) acentuam a diferença, principalmente na velocidade de 145 ppm,

em que ouvintes fazem ainda mais fixações. Novamente, como os ouvintes também fizeram mais regressões e releituras que os surdos, esse número maior de fixações parece motivado pelo maior tempo de exposição em tela, como no caso das legendas de 2 linhas e 145 ppm.

De maneira proporcional, o modelo estimou que a condição experimental com mais fixações para os surdos proporcionalmente é a de 1 linha e 145 ppm (3,65 fixações por linha de legenda). O número de fixações por linha de legenda na condição 2L145, entretanto, não é substancial (3,44 por linha de legenda). Para o grupo de ouvintes, por sua vez, a legenda de 2 linhas e 145 ppm é a que oferece mais fixações em cada linha de legenda (3,96 fixações por linha de legenda). A condição de 1L145 para ouvintes apresentou menos fixações proporcionalmente (3,65 fixações por legenda). O fato de surdos e ouvintes terem o mesmo número de fixações na condição 1L145 aponta que o tempo maior de exposição proporcionado pela condição experimental 2L145, por exemplo, faz com que o grupo mais proficiente – ouvintes – termine a leitura da legenda e realize mais fixações do que o necessário, como a maior ocorrência de regressão e releitura por ouvintes aponta, conforme será explicitado mais adiante. Esse modelo apresentou um  $R^2$  condicional de 0,715; ou seja, as variáveis de Grupo, Número de Linhas e Velocidade explicam 71,5% do número de fixações por legenda, um percentual de certa forma alto de explicação para o fenômeno de número de fixações no estudo.

Outra variável resposta analisada no trabalho foi a de regressão, mas em duas perspectivas: i) duração média de regressão; e ii) ocorrência de regressão.

Em relação à Variável Grupo, os participantes surdos realmente tiveram um tempo médio de regressão maior do que os ouvintes (290,6 ms vs. 282,82 ms), estando inicialmente em consonância com a nossa expectativa. Contudo, essa diferença não se provou significativa em nossos testes estatísticos; acreditamos que isso ocorreu, assim como nas outras variáveis independentes que serão apresentadas a seguir, muito por causa da baixa quantidade de regressões no experimento (371). Mesmo assim, os nossos resultados estão alinhados com o que foi observado por Vieira (2016), por exemplo, que também identificou uma média maior de duração de regressão em surdos do que em ouvintes (249,5 ms vs. 219 ms).

Sobre a Variável de Número de Linhas e a duração média de regressão, as legendas de 1 linha tiveram uma média maior do que as de 2 linhas (300,77 ms vs. 273,47 ms), sem, entretanto, ter havido uma diferença significativa. Apontamos, novamente, que é possível que a falta de significância dessa variável independente seja motivada pelo baixo número de regressões em nossos dados. Como os outros estudos que se debruçaram sobre a influência do número de linhas no processamento de legendas não utilizou essa medida – duração média de regressão –, não há como fazer essa comparação direta, mas sabemos que a regressão é um sinal

de desconforto na leitura de textos estáticos (RAYNER; BARBARA; POLLATSEK, 2013; RAYNER; LIVERSEDGE, 2004). Em relação a isso, todavia, há que se considerar a natureza do texto dinâmico (legenda) e o fato de o telespectador não ter controle sobre a apresentação do texto (KRUGER; STEYN, 2014).

No que tange à variável de Velocidade, as regressões realizadas em legendas de 145 ppm tiveram um tempo médio menor do que as de 180 ppm (270,53 ms vs. 282,3 ms), mas essa diferença, assim como nas outras variáveis, não foi significativa. Houve, entretanto, uma interação marginalmente significativa entre *Grupo:Velocidade*, especificamente entre surdos e ouvintes nas legendas de 180 ppm, sinalizando que a diferença de 48,84 ms a mais para os surdos nesse nível foi significativa e que os surdos tiveram um custo de processamento maior em regressões em legendas mais rápidas.

Outra parte da análise de regressões foi a ocorrência de regressões a partir das variáveis independentes do estudo. As 371 ocorrências de regressão representaram 9,11% das 4.068<sup>85</sup> fixações válidas, sendo um número próximo ao dos 10-15% apontados por Rayner (1998) como porcentagem de movimentos regressivos durante a leitura de textos estáticos. Em relação à variável Grupo, os ouvintes realizaram mais regressões do que os surdos (54,7% vs. 45,3%), indicando uma direção diferente da nossa hipótese alternativa; esse número foi maior por parte dos ouvintes também quando comparamos a porcentagem de regressões em relação ao número total de fixações (9,40% dos ouvintes vs. 8,80% dos surdos). Contudo, essa variável não foi significativa no nosso Modelo de Regressão Logística Misto de Regressão, indicando que ouvintes e surdos, apesar de terem o português como L1 e L2, respectivamente, não diferiram nesse indicativo de custo maior de processamento. Esses resultados divergem dos encontrados por Vieira (2016), cujo total de regressões por Grupo (surdos vs. ouvintes) apresentou uma significância para o número maior de regressões feitas pelos participantes ouvintes. Entretanto, há uma possibilidade de essa divergência estar relacionada à diferença do tamanho de dados sobre as fixações regressivas. Da mesma forma, Souza (2020) observou que os ouvintes fizeram mais regressões do que os participantes surdos, reforçando a ideia de que talvez a ocorrência de regressão não seja uma medida tão forte para indicar o custo de processamento de legendas. Isso porque os ouvintes, enquanto usuários nativos de português como L1, apresentaram mais ocorrências de regressão do que os surdos, que têm uma proficiência mais baixa em português, neste trabalho e nos outros estudos realizados com rastreamento ocular e LSE no Brasil (MONTEIRO, 2016; SOUZA, 2020; VIEIRA, 2016).

---

<sup>85</sup> Fixações válidas de até 600 milissegundos mais fixações de regressão de até 800 milissegundos.

Em relação ao Número de Linhas, as legendas de 2 linhas tiveram mais regressões do que as legendas de 1 linha (52,8% vs. 47,2%). Também representaram um número maior do total de fixações realizadas em ambos os níveis (9,32% vs. 8,90%). Essa variável foi significativa para o nosso Modelo, que terá seus valores estimados apontados mais adiante. Esses resultados divergem dos encontrados por d'Ydewalle e De Bruycker (2007), que observaram uma maior proporção de movimentos regressivos dentre as fixações em legendas de 1 linha do que de 2 linhas (46% vs. 34%). Os autores, contudo, utilizaram dois tipos de legenda (padrão e reversa) e trabalharam com adultos e crianças, o que faz com que a comparação com seus dados tenha que ser contextualizada. Além disso, o parâmetro de velocidade não foi trabalhado pelos autores, algo que notoriamente influencia a ocorrência de regressão, conforme veremos a seguir.

Em relação à medida de Velocidade, as legendas mais lentas, de 145 ppm, apresentaram um maior número de regressões do que as legendas de 180 ppm (56,9% vs. 43,1%) e representaram 9,82% (145 ppm) e 8,33% (180 ppm) no âmbito das fixações válidas para análise. Dada a significância dessa variável no nosso modelo, a nossa expectativa de que a variável de Velocidade não teria efeito sobre a ocorrência de regressões não foi confirmada, indicando que um maior tempo de exposição em tela, como é o caso de 145 ppm, proporciona um número maior de regressões. Esse resultado converge com o modelo ajustado por Souza (2020), que também observou que tanto os participantes surdos quanto ouvintes fizeram mais regressões nas legendas lentas. Os surdos, inclusive, apresentaram uma queda acentuada no número de regressões nas legendas rápidas.

A partir das variáveis de Número de Linhas e Velocidade, significativas na análise estatística, rodamos o Modelo de Regressão Logística Misto de Regressão, que apontou que a maior probabilidade de ocorrência de regressão está nas condições experimentais de 2 linhas, 2L145 e 2L180 (51,84% e 42,34%, respectivamente). Essa probabilidade cai bastante para as condições experimentais de 1 linha – 1L145 e 1L80 (24,78% e 18,35%) –, o que era esperado, dado que para cada inserção de legenda de 2 linhas, há 2 inserções de legenda de 1 linha. Proporcionalmente, ao dividir a probabilidade das condições experimentais de 2 linhas por 2, as probabilidades de ocorrência ficaram bem próximas entre 1L145 e 2L145 (24,78% e 25,92%, respectivamente) e 2L145 e 2L180 (18,35% e 21,17%, respectivamente). As duas médias maiores em 145 ppm, contudo, fazem-nos acreditar que, na verdade, o tempo de exposição em tela é a principal variável envolvida na ocorrência de regressão em legendagem. Apesar de ser apontado como uma medida de custo de processamento, ou seja, quanto maior a quantidade de movimentos regressivos, mais tempo o leitor/telespectador leva para processar o texto/legenda

(RAYNER; LIVERSEDGE, 2004). Contudo, há um fato que acreditamos merecer mais estudos, principalmente dentro do contexto de recepção de produtos audiovisuais legendados. Em nosso estudo, a condição experimental que teve menos ocorrências de regressões foi a 1L180; no entanto, essa condição é a que tem legendas que passam menos tempo na tela. Em contrapartida, a condição 2L145 foi a que apresentou maior probabilidade de ocorrência de regressão. As legendas que estavam nessa condição passaram mais tempo na tela em comparação às legendas na condição 1L180 por terem 2 linhas e serem mais lentas, o que reforça a opinião de Kruger e Steyn (2014) sobre a regressão para o estudo do processamento de legendas. Segundo os autores, pelo fato de os espectadores não terem controle sobre a velocidade de apresentação das legendas, é possível que eles não consigam retornar à legenda para um esclarecimento ou desambiguação de sentido.

Sendo assim, parece que quanto mais tempo a legenda está em tela, mais é permitido ao leitor de legendas que ele faça uma regressão para resolver alguma ambiguidade. Isso faz com que a ocorrência de regressão tenha um peso maior para as nossas variáveis que lidam com aspectos de tempo de exposição da legenda (Número de Linhas e Velocidade), como também o caso da ocorrência de releitura a ser apresentada a seguir.

Sobre a variável de releitura, também a analisamos em duas perspectivas: i) duração média de releitura; e ii) ocorrência de releitura. Primeiramente, sobre a duração média de releitura, na variável independente de Grupo, os surdos tiveram uma média maior do que os ouvintes (320,91 ms vs. 298,78 ms), em uma direção, conforme esperávamos. Além disso, a ANOVA identificou significância estatística na diferença das médias, apontando que, assim como havia ocorrido na medida de duração média de fixação, os surdos têm um custo maior de processamento enquanto leitores de português como L2 (RAYNER; BARBARA; POLLATSEK, 2013; RAYNER; LIVERSEDGE, 2004), aconteceu também nas fixações de releitura, que são consideradas um esforço cognitivo extra. Esses resultados estão alinhados com os encontrados por Vieira (2016), que identificou um menor tempo médio de releitura em ouvintes do que surdos quando expostos a legendas bem segmentadas de documentário (223,77 ms vs. 265,28 ms).

Na variável Número de Linhas, por sua vez, a estatística descritiva também esteve em consonância com nossa hipótese alternativa, visto que as legendas de 1 linha tiveram uma duração média de releitura maior do que as de 2 linhas (316,5 ms vs. 304,6 ms). A Análise de Variância (ANOVA) sobre essa medida na variável de Número de Linhas, contudo, não apontou significância estatística. É possível que a baixa quantidade de movimentos de releitura identificados (225) tenha feito que os 11,9 ms de diferença entre os níveis da variável não

tenham sido o suficiente para apontar uma significância estatística, o que não foi um impedimento para a variável de Grupo, dada a grande discrepância entre ouvintes e surdos (22,13 ms). A duração média de releitura não foi uma medida utilizada pelos estudos de rastreamento ocular sobre legendagem para a investigação do processamento de 1 e 2 linhas, então não é possível contextualizar com outros resultados.

Em relação à variável Velocidade, a diferença entre os dois níveis foi pequena, sendo a velocidade de 180 ppm ligeiramente maior do que a média de 145 ppm (311,41 ms vs. 308,47 ms). Sendo assim, essa diferença não se mostrou significativa na Análise de Variância (ANOVA). Comparando com os resultados de Vieira (2016), a autora identificou um tempo médio maior de releitura para as legendas bem segmentadas de 145 ppm do que 180 ppm (284,14 ms vs. 204,91 ms). Mesmo sem a autora ter feito análise estatística sobre essa medida, dada a grande diferença entre médias nos níveis da variável, é possível que houvesse significância estatística.

Apesar da significância na variável de Grupo na Análise de Variância (ANOVA), o Modelo de Regressão Linear Misto de Duração Média de Releitura não identificou nenhuma significância nas variáveis do trabalho (Grupo, Número de Linhas e Velocidade), o que impossibilitou estimar valores dessa medida que pudessem extrapolar para a população além dos participantes do experimento. Acreditamos que um maior número de dados sobre a duração de releitura seja necessário para que reforcemos ou rejeitemos os resultados aqui encontrados.

A outra parte de análise de releitura foi sobre a ocorrência dessa medida em relação às variáveis independentes do estudo. Na relação dessa medida com a variável de Grupo, apesar de uma proporção maior dos casos ter sido realizada por ouvintes do que por surdos (52,9% vs. 47,1%), não identificamos uma significância estatística a partir de um Modelo de Regressão Logística Misto de ouvintes em comparação com os surdos. A proporção entre o total de fixações válidas<sup>86</sup> realizadas por ambos os grupos, porém, mostrou que a proporção de movimentos de releitura entre surdos e ouvintes foi bem parecida (5,54% vs. 5,51%). Ou seja, o número maior de casos de releitura entre os ouvintes parece sinalizar que eles terminaram de ler as legendas primeiro, dado o alto número de fixações e menor duração média de fixação, e por isso realizaram mais fixações de releitura. Isso, contudo, não foi significativo na análise estatística, o que pode ter sido motivado pelo baixo número de releituras (225), situação semelhante à de ocorrência de regressão. Mesmo assim, o número absoluto maior de ocorrências de releitura converge com o que foi observado por Vieira (2016) em relação a um

---

<sup>86</sup> Fixações válidas para a medida de duração média de fixação de até 600 milissegundos e fixações de releitura de até 800 milissegundos.

maior número de releituras por ouvintes, assim como em De Linde e Kay (1999). Neste estudo, inclusive, os ouvintes tiveram duas vezes mais casos de releitura do que os surdos.

Sobre a ocorrência de releituras na variável de Número de Linhas, houve um número bem maior de releituras nas legendas de 2 linhas do que de 1 linha (61,3% vs. 38,7%). Fazendo uma correlação com o total de fixações realizadas nos dois níveis da variável, também houve uma proporção maior de fixações de releitura nas legendas de 2 linhas do que de 1 linha (6,54% vs. 4,42%). Essa diferença foi estatisticamente significativa e foi apontada pelo Modelo de Regressão Logística Misto. Os outros estudos sobre a influência do número de linhas no processamento de legendas não analisaram a ocorrência de releitura, então não há como contextualizar com esses outros estudos. Mas, assim como aconteceu com a ocorrência de regressão discutida anteriormente, o tempo maior de exposição em tela parece estimular a ocorrência de releituras, como no caso da LSE de 2 linhas.

Já em relação à variável de Velocidade, a proporção maior de casos de releitura nas legendas de 145 ppm do que de 180 ppm (75,1% vs. 24,9%) apresentou significância estatística no Modelo de Regressão Logística Misto, apontando que quanto mais tempo em tela, maior a ocorrência de releituras, algo que já era esperado. Além disso, a proporção de fixações de releitura no total de fixações realizadas na velocidade de 145 ppm foi também maior do que em 180 ppm (7,85% vs. 2,91%). Esses resultados reforçam os dados de Vieira (2016), que também observou um número maior de casos de releitura nas legendas bem segmentadas de 145 ppm, sem, contudo, ter feito análise estatística sobre essa diferença. Da mesma forma, também condiz com os resultados de De Linde e Kay (1999), mesmo em velocidades distintas das aplicadas neste estudo. Os autores, por sua vez, observaram a ocorrência de releitura em velocidades distintas – 74 e 139 ppm –, tendo identificado, também, um número maior de releituras na velocidade mais lenta (74 ppm), quase o dobro de ocorrências.

O Modelo de Regressão Logística Misto mais ajustado, dessa forma, foi o que apresentou significância nas variáveis de Número de Linhas e Velocidade. Além disso, houve interação entre *Número de Linhas:Velocidade* na velocidade de 145 ppm, ou seja, a mudança de 1 para 2 linhas aumentou bastante a probabilidade de ocorrência de releitura. As condições experimentais de 145 ppm (1L145 e 2L145) apresentaram as maiores probabilidades de ocorrência de releitura (12,34% e 40,19%, respectivamente), sendo seguidas, em ordem decrescente, por 2L180 e 1L180 (8,62% e 4,28%). Enquanto que a velocidade de 180 ppm mostrou que a diferença de probabilidade de ocorrência de releitura entre os níveis de 1 e 2 linhas foi bem próximo (4,28% e 4,31%, respectivamente), a grande diferença da condição experimental de 2L145 para as outras parece apontar para o fato de que quanto maior o tempo

de exposição em tela, maior a probabilidade de ocorrência de releitura, enfatizando a velocidade como principal responsável pelo fenômeno da releitura em textos dinâmicos como o da legendagem, ainda maior do que o Número de Linhas. Sendo assim, a ocorrência de releitura, assim como a ocorrência de regressão discutida anteriormente, no contexto do processamento de textos dinâmicos, pode estar mais relacionada ao tempo de exposição em tela do que a uma dificuldade de processamento, como ocorre em textos estáticos (RAYNER; LIVERSEDGE, 2004).

A outra variável que analisamos foi a duração média de encapsulamento, a última fixação realizada em cada legenda. Sobre variável de Grupo nessa medida, as 816 fixações de encapsulamento de até 800 milissegundos sinalizaram uma discrepância grande entre ouvintes e surdos (303,49 ms vs. 378,5 ms), havendo diferença estatística entre os níveis mesmo após a filtragem de *outliers* na Análise de Variância (ANOVA) e no Modelo de Regressão Linear Misto de Encapsulamento ajustado. Inclusive, mesmo após identificar um possível efeito perturbador de um dos vídeos utilizados no experimento, o efeito significativo dessa variável se confirmou sem e com o vídeo em questão. Esses resultados convergem com o que foi observado por Vieira (2016), que também identificou uma média maior de encapsulamento por surdos em comparação com ouvintes lendo legendas bem segmentadas em português (383,25 ms vs. 358,26 ms), apesar de a autora não ter realizado nenhum teste estatístico sobre essa medida. As médias dos dois grupos foi maior nessa variável resposta do que as identificadas no nosso estudo, especialmente a de ouvintes, o que também pode estar relacionado à maior densidade de informação no texto de legendas de um documentário do que de um filme. Assim como colocado por Vieira (2016), é esperado que os surdos tenham um encapsulamento maior, ou seja, um tempo médio maior na última fixação da legenda (JUST; CARPENTER, 1980), por serem usuários de português como L2, assim como tem sido apontado ao longo dessa subseção. Nesse sentido, o tempo médio dessa medida, relacionada à associação de referências do leitor/telespectador ao final da sentença/legenda, é esperado que seja maior em públicos como o surdo brasileiro lendo português.

É interessante observar que as médias próximas de encapsulamento do nosso estudo e de Vieira (2016) ajudam a validar a manipulação dos vídeos legendados utilizados no nosso experimento. Vídeos mais curtos – 8 a 15 segundos – não mostraram uma maior dificuldade de processamento em termos de encapsulamento do que vídeos maiores de 1 minuto (VIEIRA, 2016).

No que diz respeito à variável Número de Linhas, as legendas de 1 linha apresentaram um tempo médio maior de encapsulamento do que as legendas de 2 linhas



(343,92 ms vs. 326,68 ms). Essa diferença, contudo, caiu para 9,32 milissegundos após a filtragem de *outliers* e não apresentou significância estatística nem na Análise de Variância (ANOVA) nem no Modelo ajustado. Assim sendo, a expectativa de que a duração média de encapsulamento seria maior nas legendas de 1 linha do que de 2 linhas foi em parte confirmada. Dada essa não significância, entretanto, acreditamos que um número maior de dados sobre o encapsulamento possa confirmar ou refutar com mais propriedade esse resultado. Outros estudos sobre o processamento de legendas de 1 e 2 linhas também não utilizaram essa medida, que consideramos relevante para observar o custo de processamento a partir do processamento de cada legenda e completudes sintáticas distintas.

Sobre a variável de Velocidade, as legendas de 145 ppm tiveram uma duração média de encapsulamento ligeiramente maior do que as de 180 (340,13 ms vs. 336,33 ms); mesmo após a filtragem de *outliers*, a Análise de Variância e o Modelo de Regressão Linear Misto de Encapsulamento não apresentaram significância estatística para essa variável, corroborando a expectativa de que não haveria diferença significativa entre os níveis dessa variável. Vieira (2016), ao observar as médias de encapsulamento em legendas bem segmentadas de documentário, identificou uma média bem maior nas legendas de 145 ppm do que nas de 180 ppm (386,66 ms vs. 354,85 ms), mas não foi realizado teste estatístico sobre essa medida para aferição da significância dessa diferença.

Dessa forma, a partir dos dados coletados, conseguimos convergir um Modelo de Regressão Linear Misto de Encapsulamento com a única variável que se mostrou significativa: Grupo. Além disso, dada a significância de um dos vídeos utilizados como item experimental, rodamos esse modelo com e sem o vídeo em questão, o que gerou dois valores estimados – duração média de encapsulamento – distintos para surdos e ouvintes. O modelo sem o vídeo apresentou como valor estimado para ouvintes 288,92 ms, enquanto para surdos, o valor de 361,5 ms. Este modelo apresentou um valor de  $R^2$  condicional maior do que o modelo com todos os vídeos, indicando que a variável Grupo explica em 27,7% do fenômeno de duração média de encapsulamento em legendas, um valor explicativo não tão alto. Isso significa que há outros fatores além de Grupo – ser ouvinte ou surdo – que explicam a duração média de encapsulamento em legendas.

Sobre a variável de deflexão, analisamos, neste trabalho, em relação apenas a sua ocorrência, não seu tempo médio. Além disso, consideramos apenas os movimentos deflexivos – movimento de ida e volta do texto legendado para a imagem – durante a leitura da legenda, e não após a finalização da sua leitura. De acordo com os 139 movimentos deflexivos identificados, ouvintes realizaram um número maior de movimentos deflexivos do que os

surdos (59% vs. 41%), contudo, a partir do Modelo ajustado de Regressão Logística Misto de Deflexão, não houve significância entre os níveis. É possível que esse resultado tenha sido causado pelo número baixo de observações dessa medida. Assim, nossos resultados não corroboram os achados de Vieira<sup>87</sup> (2016), que observou um número maior de movimentos deflexivos de surdos em relação a ouvintes quando expostos a um documentário legendado, especialmente em legendas rápidas (180 ppm) bem segmentadas. Souza<sup>88</sup> (2020), por sua vez, observou que os participantes surdos e ouvintes realizaram mais deflexões nas legendas lentas tanto em legendas com boa segmentação quanto com má segmentação, divergindo de Vieira e se assemelhando aos nossos resultados no que tange à variável de velocidade. Essa distinção no trabalho de Souza (2020) entre os níveis da variável de Velocidade foi mais acentuada no grupo de surdos do que de ouvintes; ou seja, os surdos fizeram bem menos deflexões nas legendas rápidas (180 ppm) do que nas lentas (145 ppm).

Ademais, nossos resultados não convergem com os resultados de Szarkowska e Gerber-Morón (2018) sobre o maior número de deflexões de surdos do que ouvintes, especialmente em legendas de 3 linhas quando comparadas com legendas de 2 linhas. Entretanto, estão alinhados com a perspectiva de Kruger, Szarkowska e Krejtz (2015) de que surdos e ouvintes têm padrão de processamento semelhante, de não realizar muitas deflexões durante a leitura; ou seja, ensurdecidos teriam uma probabilidade maior de realizar deflexões durante a leitura de legendas do que surdos e ouvintes (KRUGER; SZARKOWSKA; KREJTZ, 2015).

Sobre a variável de Número de Linhas, houve um número bem maior de deflexões em legendas de 2 linhas do que de 1 linha (66,9% vs. 33,1%). Essa diferença também se mostrou significativa no Modelo Ajustado de Regressão Logística de Deflexão que será apresentado mais adiante. Esses resultados convergem com os resultados de d'Ydewalle e De Bruycker (2007), que também identificaram um número maior de deflexões em legendas de 2 linhas processadas por crianças e adultos, especialmente na legendagem padrão. Apesar de reconhecerem que essa significância da deflexão pode indicar um padrão de leitura menos fluido nas legendas de 2 linhas, os autores apontam que essa conclusão precisa ser qualificada, visto que, dada a regra dos 6 segundos, há tempo suficiente para que se vá à imagem após uma primeira leitura do texto. Essa consideração indica que os autores consideraram o movimento de deflexão também como aquele realizado após a leitura completa da legenda, diferente do

---

<sup>87</sup> A autora, entretanto, analisou os dois tipos de movimentos de deflexão citados anteriormente, diferentemente deste trabalho.

<sup>88</sup> Assim como Vieira (2020), Souza (2020) utilizou os dois tipos de deflexão na sua análise.

que fizemos aqui e semelhante a um dos tipos de deflexão considerados por Vieira (2016) e Souza (2020). Ainda sobre isso, Szarkowska e Gerber-Morón (2018) observaram as deflexões de ouvintes, surdos e ensurdecidos britânicos expostos a legendas de 2 e 3 linhas. As autoras identificaram uma maior quantidade de revisitas (*revisits*), outro nome utilizado para deflexão, nas legendas de 3 linhas do que nas de 2 linhas, sugerindo que um número maior de linhas ocasiona um número maior de deflexões. Contudo, do mesmo modo que d'Ydewalle e De Bruycker (2007), Szarkowska e Gerber-Morón (2018) indicaram que isso seria motivado pelo maior tempo em tela, que permitiria o movimento de deflexão após a leitura da legenda. Esse ponto mostra que os autores levaram em conta somente ou principalmente – isso não é claro no conceito de deflexão utilizado por elas – movimentação de deflexão após a finalização da leitura da legenda, e não durante, como é feita aqui.

Em relação à variável de Velocidade, houve um número bem maior de movimentos deflexivos em 145 ppm do que em 180 ppm (77% vs. 23%), corroborando os resultados de Souza (2020). Essa diferença foi significativa no nosso modelo para a variável Velocidade. Esses resultados são semelhantes aos de Vieira (2016), que também observou um número maior de deflexões nas legendas de 145 ppm do que 180 ppm, mas houve interação em suas variáveis: legendas de 145 ppm (lentas) propiciaram mais deflexões para os ouvintes, enquanto as legendas de 180 ppm (rápidas) para os surdos. Essa interação, porém, não foi identificada em nosso trabalho. As variáveis de Número de Linhas e Velocidade, dessa forma, foram utilizadas para ajustar um Modelo de Regressão Logística Misto de Deflexão de modo que pudessemos estimar a probabilidade de ocorrência de deflexão a partir dessas variáveis independentes e dos efeitos mistos (vídeos e participantes). A partir desse modelo, identificamos uma probabilidade maior de ocorrência nas condições experimentais de 145 ppm: 1L145 e 2L145 (7,21% e 30,88%, respectivamente), especialmente nesta última. Além disso, quando comparados os pares de 1 linha (1L145 e 1L180) e de 2 linhas (2L145 e 2L180), percebemos uma maior probabilidade de ocorrência de deflexão nas legendas de 2 linhas. Essa diferença, inclusive, pode ser vista de forma proporcional, dividindo a probabilidade de ocorrência de deflexão das legendas de 2 linhas por 2, em 1L145 e 2L145 (7,21% vs. 15,44%, respectivamente) e 1L180 e 2L180 (1,17% vs. 3,19%, respectivamente).

Dada a maior probabilidade de ocorrências nas legendas de 145 ppm, assim como o maior efeito estatístico dessa variável (Velocidade) no modelo em comparação com Número de Linhas e a grande discrepância de 2L145 para as outras condições experimentais, podemos correlacionar os nossos dados com d'Ydewalle e De Bruycker (2007) e Szarkowska e Gerber-Morón (2018) e apontar que o maior tempo de exposição em tela parece proporcionar um

número maior de deflexões também durante a leitura das legendas, não só após a finalização destas. Nesse sentido, podemos também questionar a visão de deflexão como algo que perturba o processamento de legendas (DE LINDE; KAY, 1999; D'YDEWALLE; DE BRUYCKER, 2007; KRUGER; SZARKOWSKA; KREJTZ, 2015; SZARKOWSKA; GERBER-MORÓN, 2018) e refletir sobre o que coloca Vieira (2016) sobre esse fenômeno. A autora defende que a deflexão é um aspecto comum do processamento de vídeos legendados e faz parte do processamento de dois níveis semióticos distintos – imagético e verbal – que ocorrem em um gênero multimodal como é a legendagem, sendo inclusive algo que facilita o processamento de surdos brasileiros quando expostos a legendas rápidas (180 ppm) e bem segmentadas de documentários. Além disso, no caso de um gênero audiovisual em que há ação ocorrendo em tela, quanto mais tempo de exposição da legenda, mais chance de ocorrer alguma movimentação ou fala de personagens e, por isso, a atenção do telespectador se voltar para a imagem e ele realizar esse processo de idas e vindas entre áreas de interesse, como observaram De Linde e Kay (1999). No caso dos surdos, há ainda a movimentação de lábios dos personagens, que pode ocasionar esse movimento de deflexão (SZARKOWSKA; GERBER-MORÓN, 2018), especialmente em legendas que ficam mais tempo em tela.

A variável de latência, diferentemente das outras analisadas até aqui, não se trata de uma medida de análise do processamento *per se*, mas de alocação de atenção entre imagem e legenda no caso da legendagem. No que diz respeito à variável Grupo, os participantes surdos tiveram uma duração média de latência maior do que os ouvintes (395,75 ms vs. 369,49 ms), mas, após a filtragem de *outliers*, a Análise de Variância (ANOVA) não identificou diferença significativa entre os níveis dessa variável, assim como não convergiu em nosso Modelo de Regressão Linear Misto de Latência. Isso nos permite dizer que os dados apontam para um maior tempo de latência dos surdos do que dos ouvintes, mas não uma diferença estatisticamente significativa. Contudo, isso pode ter sido motivado pelas 652 observações utilizadas para a análise estatística, número não tão alto. Isso porque a direção apontada pelos dados descritivos é bem similar à direção dos de Vieira (2016), que também identificou um tempo maior de latência – que a autora denominou como “atrasos de fixação” – nos participantes surdos em comparação com os ouvintes expostos a legendas bem segmentadas de um documentário. A autora, entretanto, trabalhou com uma medida de tempo total em segundos, juntando todos os tempos de latência, e não por inserção de legenda como fizemos aqui, o que faz com que não possamos comparar as médias dos seus grupos diretamente com os resultados aqui observados.

Sobre a Variável Número de Linhas, os participantes tiveram um tempo maior nas legendas de 1 linha no que nas legendas de 2 linhas (348,58 ms vs. 367,58 ms, respectivamente). Essa média maior, inclusive, foi além do que esperávamos. D'Ydewalle *et al.* (1987) observou um tempo maior de latência nas legendas de 2 linhas em seu trabalho em duas condições experimentais (1 – sem conhecimento da língua da legenda e com som e 2 – com conhecimento da língua da legenda e com som), mas uma diferença proporcional e não absoluta. Em outras palavras, comparando o tempo de latência de 2 inserções de 1 linha com o de 2 linhas, o tempo de latência nas legendas de 2 linhas seria menor, apontando uma possibilidade menor de perda de conteúdo linguístico nas legendas de 2 linhas a partir de uma demora para se começar a processar a legenda e ela desaparecer antes do término desse processo. Na outra condição experimental do estudo de d'Ydewalle *et al.* (1987), assistindo ao vídeo legendado sem som, os participantes tiveram um tempo maior de latência nas legendas de 1 linha do que nas legendas de 2 linhas (382 ms vs. 348 ms), sem significância estatística. Esse resultado pode ser correlacionado com o nosso na medida em que também manipulamos o vídeo legendado do experimento sem som, dado que isso poderia influenciar os participantes ouvintes ao escutarem o áudio na sua L1. Assim, também identificamos um tempo maior absoluto na latência de legendas de 1 linha do que nas legendas de 2 linhas.

Os resultados de d'Ydewalle e De Bruycker (2007) são similares aos das duas primeiras condições experimentais de d'Ydewalle *et al.* (1987), que também observaram um tempo maior proporcional de latência de crianças e adultos em legendas de 2 linhas quando comparadas às legendas de 1 linha (501 ms vs. 375 ms), sendo este efeito significativo. Os resultados dessa medida, entretanto, apresentaram interação entre Número de Linhas e Legendagem, apontando uma significância na legendagem reversa (1 linha = 416 ms, 2 linhas = 610 ms), mas não na legendagem padrão (1 linha = 333 ms, 2 linhas = 391 ms).

De qualquer forma, tanto os resultados de d'Ydewalle *et al.* (1987) quanto os de d'Ydewalle e De Bruycker (2007) indicaram que, proporcionalmente, as legendas de 2 linhas permitiriam um maior tempo de leitura aos telespectadores, assim como no nosso trabalho. Nesse sentido, o fato de nossos resultados indicarem um tempo absoluto maior de latências nas legendas de 1 linha do que de 2 linhas para realizar a latência da imagem nas condições experimentais de 1 linha coloca ainda mais ênfase nos resultados já encontrados por d'Ydewalle *et al.* (1987) e d'Ydewalle e De Bruycker. (2007). O fato de termos isolado as legendas de 1 linha e as de 2 linhas em condições experimentais distintas pode ser o causador desse fato, o que não foi feito por d'Ydewalle *et al.* (1987) e d'Ydewalle e De Bruycker (2007). Ou seja, legendas de 1 linha seguidas uma da outra podem gerar um efeito sobre duração média de

latência ainda maior do que isoladas. Outra explicação seria o fato de termos omitido o som do vídeo, algo incomum para ouvintes, mas não para surdos, o que pode fazer com que as legendas de 1 linha tenham um tempo maior de latência, assim como foi observado em uma das condições experimentais de d'Ydewalle *et al.* (1987). De qualquer forma, o fato de demorarem mais tempo para acessar as legendas de 1 linha é ainda mais problemático no contexto das legendas de 180 ppm, pois é possível que a legenda suma sem que o telespectador consiga lê-la por completo.

Em relação à Velocidade, as legendas de 145 ppm tiveram um tempo médio de latência maior do que as de 180 ppm (367,48 ms vs. 348,58 ms), efeito significativo tanto na Análise de Variância (ANOVA) quanto no modelo ajustado. No entanto, este efeito apresentou interação com Número de Linhas, o que falaremos mais à frente. Esses resultados divergem de d'Ydewalle *et al.* (1987), que não observaram diferença significativa no tempo médio de latência utilizando as regras de 2 e 4 segundos, velocidades mais rápidas do que a de 6 segundos. Além disso, divergem dos resultados de Vieira (2016), que observou um tempo maior de “atraso” nas fixações em legendas rápidas (180 ppm) bem segmentadas de documentários. A autora, contudo, não fez teste estatístico dessa medida em questão.

A partir de Número de Linhas e Velocidade, ajustamos um Modelo de Regressão Linear Misto de Latência, que apresentou interação *Número de Linhas:Velocidade*. Essa interação mostrou que não há diferença significativa entre legendas de 1 e 2 linhas na velocidade rápida (180 ppm), mas na velocidade lenta (145 ppm), o tempo de latência em 2 linhas é significativamente menor do que nas legendas de 1 linha. Esse modelo, então, previu que as duas condições experimentais de 2 linhas (2L145 e 2L180) foram os menores valores dessa medida (340,2 ms e 344,86 ms, respectivamente), enquanto as condições experimentais de 1 linha (1L145 e 1L180) tiveram os maiores valores da medida (386,56 ms e 351,43 ms, respectivamente). Apesar de não ser uma medida de processamento, mas sim de alocação de atenção, o tempo menor nas legendas de 2 linhas proporciona um tempo maior de leitura. Dessa forma, as legendas de 2 linhas também oferecem uma menor possibilidade de perder conteúdo da legenda, especialmente em legendas mais rápidas.

Assim, ao fazermos uma avaliação geral dos resultados das medidas significativas, podemos apontar que as legendas de 2 linhas ofereceram um padrão de leitura mais regular do que as de 1 linha para os ouvintes (menor duração média de fixação, maior número de fixações por legenda, duração média de latência), sendo a medida de duração de fixação a mais relevante para o estudo do processamento. Outras medidas que tiveram uma menor média para os ouvintes nas condições experimentais de 2 linhas, como duração média de releitura, duração

média de regressão, não foram estatisticamente significativas, o que não nos permite afirmar categoricamente a diferença entre 1 e 2 linhas em relação a essas medidas. Nas legendas de 1 linha, a duração média de encapsulamento foi menor, mas uma diferença bem pequena e que não apontou significância estatística. A ocorrência significativa de regressão, deflexão e releitura (2L145) pareceu mais algo associado ao tempo de exposição em tela do que ao número de linhas e maior completude sintática que ela proporciona.

Ainda em relação aos ouvintes, os resultados das variáveis não nos permitem afirmar categoricamente se as legendas de 145 ou de 180 ppm foram melhor processadas, visto que a maior parte das variáveis dependentes numéricas não apresentou significância estatística; contudo, elas apontam marginalmente para um custo menor de processamento nas legendas de 180 ppm. Dentre as que apresentaram significância estatística, o tempo de latência foi menor nas de 180 ppm, especialmente em relação às legendas na condição experimental 1L145, enquanto a modelagem de número de fixações por legenda previu um número maior de fixações em legendas de 145 ppm (esse resultado, entretanto, parece estar associado ao número maior de releituras e regressões que os ouvintes fizeram, não só ao fato de naturalmente fazerem mais fixações e mais curtas do que os surdos como falantes de português como L1). Dentre as medidas que não apresentaram significância, as legendas de 180 ppm tiveram uma menor duração média de fixação e menor duração de encapsulamento, enquanto as legendas de 145 ppm tiveram uma menor duração média de regressão e menor duração média de releitura. As variáveis categóricas, entretanto, de ocorrência de regressão, releitura (2L145, principalmente) e deflexão apontaram para uma ocorrência maior em legendas de 145 ppm, o que parece apontar para uma maior digressão no processamento de legendas de 145 ppm por parte dos ouvintes.

No que diz respeito aos surdos, podemos apontar que as legendas de 2 linhas também ofereceram um padrão de leitura mais regular do que as de 1 linha (menor duração média de fixação e duração média de latência) de acordo com o modelo ajustado. Proporcionalmente, os surdos tiveram uma média maior de fixações em legendas de 1 linha, mas a diferença foi bem pequena. Dentre as variáveis que não foram significativas, os surdos tiveram, nas legendas de 2 linhas, uma duração média de regressão e duração média de encapsulamento menor do que nas legendas de 1 linha. Dentre as legendas de 1 linha, os surdos tiveram uma duração média de releitura menor, mas também sem significância. Assim como aconteceu com os ouvintes, a significância da ocorrência de regressão, deflexão e releitura (2L145, principalmente) em legendas de 2 linhas pareceu mais algo associado ao tempo de

exposição em tela, maior nas legendas de 2 linhas, do que algo relacionado à sintaxe dessas legendas.

De maneira análoga ao grupo de ouvintes, a falta de significância em diversas medidas e a circunstância das significâncias de algumas destas, conforme já colocado anteriormente sobre as variáveis categóricas analisadas, dificultam a tarefa de apontar categoricamente se as legendas de 145 ou 180 ppm proporcionam um melhor processamento para os surdos. Contudo, de acordo com as medidas analisadas e sua concepção para os estudos de processamento de legendas, as legendas de 180 ppm parecem propiciar um processamento menos custoso e mais fluido. Esse tipo de contextualização dos dados é necessário porque trabalhamos com várias medidas de processamento, além do fato de algumas dessas medidas – as quais são historicamente utilizadas nos estudos sobre o processamento de textos estáticos – não se mostrarem tão significantes para a análise de textos dinâmicos como a legendagem, vide a ocorrência de regressão e releitura (KRUGER; STEYN, 2014). Dentre as medidas que apresentaram significância estatística, a duração média de latência foi menor nas legendas de 180 ppm, especialmente em relação à 1L145. As legendas de 145 e 180 ppm processadas pelos participantes surdos não diferiram em termos de significância estatística na medida de duração média de fixação. As legendas de 145 ppm, por um lado, tiveram um maior número de fixações por legenda, sendo essa diferença significativa. Além disso, as legendas de 145 ppm processadas por surdos tiveram uma menor duração média de encapsulamento e de regressão, mas sem diferença estatística significativa. As legendas de 180 ppm, por outro lado, tiveram uma menor duração média de releitura, também sem significância estatística. Do mesmo modo que os ouvintes, as variáveis categóricas de regressão, releitura (2L145) e deflexão foram significativamente maiores nas legendas de 145 ppm, mas esses dados precisam ser mais qualificados sobre a sua relação com o custo de processamento dos participantes, dado o que já foi discutido sobre a concepção dessas medidas para o estudo do processamento de legendas.

Resumindo, a partir dessa discussão e da análise do resultado de cada medida à luz da teoria exposta, percebemos que, como esperado, os participantes surdos, enquanto usuários de português como L2, tiveram um maior custo de processamento dos vídeos legendados. Isso, entretanto, já era algo previsto em teoria. Sobre a variável Número de Linhas, as legendas de 2 linhas se mostraram mais confortáveis do que as legendas de 1 linha não só para ouvintes, mas também para surdos, conforme estipulava a hipótese deste trabalho e grande parte da literatura da área (D'YDEWALLE *et al.*, 1987; D'YDEWALLE *et al.*, 1991; D'YDEWALLE; DE BRUYCKER, 2007; D'YDEWALLE; GIELEN, 1992; PRAET *et al.*, 1990), com exceção de Kruger e Steyn (2014). A variável de Velocidade, por sua vez, dada a complexidade de algumas



das medidas e a falta de significância de outras, apontou apenas parcialmente para um menor custo de processamento nas legendas de 180 ppm do que nas legendas de 145 ppm, refutando, inicialmente, a nossa hipótese de que não haveria diferença em termos de processamento de legendas nessas duas velocidades. Os resultados sobre Velocidade, entretanto, precisam ser mais qualificados, principalmente no que diz respeito à influência da ocorrência de deflexão, regressão e releitura em textos dinâmicos como a legendagem.

## 5 CONCLUSÃO

O presente trabalho investigou a influência do número de linhas e da velocidade no processamento de LSE por ouvintes e surdos brasileiros. Para tal, utilizamos uma metodologia experimental baseada na utilização de rastreamento ocular seguindo o Princípio da Ligação Olho-Mente (JUST; CARPENTER, 1980; RAYNER, 1998), segundo o qual a nossa atenção está na nossa fixação do olhar. A partir de estudos sobre processamento de textos estáticos (JUST; CARPENTER, 1980; RAYNER, 1998; RAYNER; BARBARA; POLLATSEK, 2013; RAYNER; LIVSERDGE, 2004; REICHLER; RAYNER; POLLATSEK, 2003), do processamento de legendas (D'YDEWALLE *et al.*, 1987; D'YDEWALLE *et al.*, 1991; D'YDEWALLE; DE BRUYCKER, 2007; DE LINDE; KAY, 1999; KRUGER; SZARKOWSKA; KREJTZ, 2015; KRUGER; STEYN, 2004; PRAET *et al.*, 1990; SZARKOWSKA; GERBER-MORÓN, 2018) e, mais especificamente para o nosso contexto, a recepção de LSE (ARAÚJO, 2008; ARAÚJO, 2012; ARAÚJO; NASCIMENTO, 2011; FRANCO; ARAÚJO, 2003) e o estudo sobre o processamento de legendas por surdos brasileiros através de rastreamento ocular (MONTEIRO, 2016; SOUZA, 2020; VIEIRA, 2016), a presente pesquisa teve os seguintes objetivos: a) investigar o custo de processamento de legendas de 1 linha e 2 linhas em surdos e ouvintes; b) investigar o custo de processamento de legendas rápidas (180 palavras por minuto) e lentas (145 palavras por minuto) em surdos e ouvintes.

Para responder às seguintes perguntas de pesquisa associadas a esses objetivos específicos: a) Até que ponto legendas de 1 linha ou 2 linhas influenciam o processamento de legendas por surdos e ouvintes?; b) Até que ponto legendas rápidas (180 palavras por minuto) ou lentas (145 palavras por minuto) influenciam o processamento de surdos e ouvintes?, formulamos as seguintes hipóteses baseadas na literatura da área: a) Legendas de 1 linha induzem um processamento menos linear e, dessa forma, mais custoso do que legendas de 2 linhas; b) Legendas rápidas (180 palavras por minuto) e lentas (145 palavras por minuto) não diferem significativamente no processamento de surdos e ouvintes.

Em suma, a nossa tese era de que legendas de 2 linhas facilitariam o processamento de legendas também por surdos, além de ouvintes, algo que já era previsto em teoria (D'YDEWALLE *et al.*, 1987; D'YDEWALLE *et al.*, 1991; D'YDEWALLE; DE BRUYCKER, 2007; DE LINDE; KAY, 1999; PRAET *et al.*, 1990). Além disso, a outra parte da nossa tese era de que as legendas com velocidades consideradas lentas (145 ppm) e rápidas (180 ppm) não apresentariam diferença significativa para surdos e ouvintes; os ouvintes por

estarem lendo em sua L1 (português), e os surdos por apresentarem uma recepção confortável de legendas tanto na velocidade de 145 ppm quanto 180 ppm (ARAÚJO, 2012; NASCIMENTO; ARAÚJO, 2011), apesar de Vieira (2016) ter percebido um processamento mais fluido por surdos em legendas rápidas. Entretanto, como é algo que ainda é controverso na literatura, decidimos testar essa hipótese sobre velocidade no nosso trabalho junto à hipótese sobre o número de linhas.

Após análise dos dados, os resultados das medidas de processamento indicaram um processamento mais confortável por ouvintes e surdos em legendas de 2 linhas, visto que estas ofereceram um padrão de processamento mais regular. A partir das medidas de processamento, observamos que as legendas de 2 linhas tiveram uma menor duração média de fixação e uma menor duração média de latência em ambos os grupos, medidas significativas estatisticamente. Outras medidas que apontaram um processamento mais confortável nas legendas de 2 linhas, entretanto, não foram estatisticamente significativas, como duração média de releitura e duração média de regressão. A duração média de encapsulamento também foi menor nas legendas de 2 linhas processadas por participantes surdos, enquanto os ouvintes tiveram uma duração média de encapsulamento praticamente igual em legendas de 1 e 2 linhas. Essa variável também não apresentou significância estatística.

A variável de número de fixações por legenda apresentou um maior número de fixações proporcionalmente em legendas de 2 linhas por ouvintes, enquanto surdos realizaram um número maior de fixações por legenda nas legendas de 1 linha, especialmente na condição de 145 ppm. Segundo a nossa perspectiva inicial sobre essa variável, os ouvintes realizariam um número maior de fixações por serem leitores de português como L1 (RAYNER; BARBARA; POLLATSEK, 2013), mas fixações mais curtas. Isso em comparação com os surdos do experimento, que processam as legendas enquanto leitores de português como L2. Segundo a nossa hipótese, as legendas de 2 linhas também proporcionariam um número maior de fixações, mas mais curtas da mesma forma. Contudo, apesar de os ouvintes terem realizado mais fixações do que os surdos e as legendas de 2 linhas terem proporcionado um número maior de fixações para esse grupo, os ouvintes também realizaram um número maior de regressões e de releitura do que os surdos, assim como as legendas de 2 linhas também apresentaram uma maior ocorrência dessas medidas. Assim, o papel dessa medida da forma que ela foi analisada no estudo ficou mais associada ao tempo de exposição em tela – especialmente para os ouvintes, que realizaram mais fixações e leram as legendas com mais frequência – do que a um possível indício de mais conforto. Uma possibilidade para estudos futuros é a análise apenas de fixações únicas – desconsiderando regressões e releituras – para essa medida. Entretanto, no contexto de

textos dinâmicos como a legendagem, parece-nos difícil isolar essa medida do tempo de exposição, o que dificulta a sua compreensão como índice de processamento leitor de legendas.

Nesse sentido, é interessante ressaltar, conforme já frisado no parágrafo anterior, o caso da regressão e da releitura enquanto indícios do custo de processamento de legendas por surdos e ouvintes. A ocorrência significativa de regressão e releitura em legendas de 2 linhas tanto por surdos quanto por ouvintes coloca em xeque o papel dessas medidas como índices do custo de processamento de legendas. Se a ocorrência de regressão e releitura, no âmbito do processamento de textos estáticos, representa um sinal de desconforto, no contexto do processamento de legendas parece ser algo mais associado ao tempo de exposição em tela. Isso se alinha ao pensamento de Kruger e Steyn (2014), os quais questionam a análise de regressão para análise do custo de processamento de legendas porque o telespectador, diferentemente de um texto impresso, não pode retornar à legenda no momento que deseja para resolver alguma dificuldade durante a leitura do texto, sendo o tempo de exposição da legenda controlado pelo legendista.

Além das medidas de releitura e regressão, a medida de deflexão, vista inicialmente como indicador de custo de processamento, também foi significativa nas legendas de 2 linhas. É válido ressaltar, primeiramente, as concepções distintas de deflexão – ou sob outros termos igualmente utilizados para descrever esse movimento – na literatura de Tradução Audiovisual. Ao contrário de estudos que consideram um dos dois tipos de deflexão, tanto a ida à imagem e volta à legenda depois da finalização da leitura desta (D'YDEWALLE *et al.*, 2007; SZARKOWSKA; GERBER-MORÓN, 2018) quanto a ida à imagem e volta à legenda antes da finalização da leitura desta (KRUGER; SZARKOWSKA; KREJTZ, 2015), ou ainda esses dois tipos de deflexão (SOUZA, 2020; VIEIRA, 2016), o presente trabalho se debruçou apenas sobre o segundo tipo. Isso porque nos pareceu, inicialmente, que o ato de ir à imagem e voltar à legenda antes da finalização desta era mais um indício de ruptura no processamento do texto (KRUGER; SZARKOWSKA; KREJTZ, 2015) por surdos e ouvintes<sup>89</sup> do que o ato de voltar à legenda após a finalização desta. Este tipo de deflexão, associado ao ato de releitura da legenda, já era vislumbrado como algo relacionado a um tempo maior de exposição de uma legenda na tela; ou seja, a legendas mais lentas. Isso porque, de acordo com a hipótese da automaticidade (D'YDEWALLE; GIELEN, 1992), a leitura de legendas é mais ou menos automática. Dessa forma, o fato de as legendas ainda estarem presentes em tela por mais tempo que o devido

---

<sup>89</sup> Isso porque para os ensurdecidos, os quais não foram selecionados para esta pesquisa, esse tipo de movimento deflexivo seria uma característica do seu padrão de processamento, não um indicativo de custo, conforme colocado anteriormente (KRUGER; SZARKOWSKA; KREJTZ, 2015).

causaria esse tipo de deflexão chamado por Szarkowska e Gerber-Morón (2018) de *revisit* (*revisit*).

Assim como no caso de regressão e releitura, a maior probabilidade de ocorrência de deflexão nas legendas de 2 linhas do que nas legendas de 1 linha por surdos e ouvintes pareceu algo mais relacionado ao tempo de exposição das legendas do que a uma dificuldade de processamento causado pela sintaxe das legendas. Os ouvintes, inclusive, realizaram um número maior de deflexões do que os surdos. Caso fosse um indicativo de dificuldade de processamento, os ouvintes não teriam tido um desempenho condizente com o seu lugar de usuários de português como L1 processando legendas intralinguísticas, conforme foi possível perceber ao comparar os resultados desse grupo controle com outros estudos semelhantes realizados no Brasil (MONTEIRO, 2016; SOUZA, 2020; VIEIRA, 2016).

Por conseguinte, enfatizamos o que foi colocado por Vieira (2016) sobre o fenômeno de deflexão, mesmo a autora tendo utilizado para análise os dois tipos de deflexão mencionados anteriormente. Segundo Vieira (2016), a deflexão é uma característica comum do processamento de vídeos legendados, principalmente dos surdos, visto que ele denota a integração entre os meios semióticos verbal – apenas na modalidade escrita, no caso da LSE – e imagético. Assim, é uma medida que mostra conforto ao invés de custo de processamento, conforme preconiza a literatura em Tradução Audiovisual (DE LINDE; KAY, 1999; D'YDEWALLE *et al.*, 2007; KRUGER; SZARKOWSKA; KREJTZ, 2015; SZARKOWSKA; GERBER-MORÓN, 2018). Assim, quanto mais tempo da legenda na tela, conforme proporcionado por legendas de 2 linhas na condição lenta (145 ppm), condição experimental com maior probabilidade de ocorrência de deflexão segundo nosso modelo, maior a chance do tipo de deflexão analisado nesta pesquisa. Isso também pode estar relacionado à imagem: quanto mais tempo em tela, maior a chance de movimentação de cena e, dessa forma, o telespectador mudar sua atenção para a imagem e voltar para a leitura da legenda. Isso seria ainda mais relevante para os telespectadores surdos, visto que a movimentação de lábios é outra questão que chama a atenção dos surdos para a imagem durante o processamento de vídeos legendados (SZARKOWSKA; GERBER-MORÓN, 2018).

Em relação à variável Velocidade, especificamente, a maior parte dos resultados das variáveis numéricas não apresentou significância estatística, como foi o caso da duração média de fixação, em que os ouvintes tiveram um desempenho ligeiramente melhor nas legendas rápidas (180 ppm) e os surdos, ligeiramente melhor nas lentas (145 ppm). Do mesmo modo, as medidas de duração média de encapsulamento, duração média de releitura e regressão também não apresentaram significância estatística, e a direção dos seus resultados variou entre

as legendas de 145 ppm e 180 ppm em ambos os grupos de participantes. Por um lado, os ouvintes tiveram uma melhor performance nas legendas de 180 ppm na medida de duração média de encapsulamento, mas uma melhor performance nas legendas de 145 ppm em relação à duração média de regressão e duração média de releitura. Os surdos, por sua vez, tiveram uma melhor performance nas legendas de 180 ppm na medida de duração média de releitura, mas uma melhor performance nas legendas de 145 ppm nas medidas de duração média de encapsulamento e de regressão.

A única variável dependente numérica que apresentou significância estatística foi a de duração média de latência, que apontou um tempo menor de latência para legendas de 180 ppm para ambos os grupos. Isso, inclusive, é especialmente relevante em legendas mais rápidas, visto que elas desaparecem mais rápido da tela, e os telespectadores podem perder conteúdo linguístico das legendas caso não façam o movimento de latência rápido o suficiente. Assim, há uma clara relação dessa medida com o tempo de exposição em tela. A interação *Velocidade:Número de Linhas* nessa medida, inclusive, é algo que coloca mais ênfase no conforto do processamento de legendas de 2 linhas e de 180 ppm. Na velocidade de 180 ppm, não houve diferença significativa entre legendas de 1 e 2 linhas. Já na velocidade de 145 ppm, as legendas de 1 linha apresentaram um tempo significativamente maior de latência do que as legendas de 2 linhas. Visto que as legendas de 1 linha ficam menos tempo em tela do que as legendas de 2 linhas, esse tempo maior de latência pode ser especialmente problemático nesse contexto.

A variável de número de fixações por legenda foi maior nas legendas de 145 ppm do que nas legendas de 180 ppm por surdos e ouvintes. Houve, contudo, uma interação *Grupo:Número de Linhas:Velocidade*: há pouca mudança entre o número de fixações e ouvintes nas legendas de 145 e 180 ppm de 1 linha. Quando em legendas de 2 linhas, entretanto, há uma queda acentuada por parte dos surdos nas legendas de 2 linhas e 145 ppm quando comparados com os ouvintes. Como já colocado anteriormente, essa medida não favoreceu a nossa análise porque esse número maior de fixações por parte dos ouvintes está fortemente relacionado ao maior número de regressões e releituras que eles fizeram.

Sobre as variáveis categóricas do estudo – ocorrência de regressão, ocorrência de releitura e ocorrência de deflexão – e sua relação com a variável independente *Velocidade*, todas foram significativas para uma maior probabilidade de ocorrência em legendas de 145 ppm do que em legendas de 180 ppm. Contudo, como já discutido anteriormente, a ocorrência de regressão não parece ser uma boa medida de índice de desconforto no processamento de textos legendados (KRUGER; SZARKOWSKA; KREJTZ, 2015), semelhantemente a ocorrência de

releitura, dado que o telespectador não possui controle sobre o tempo de exposição do texto legendado. Desse modo, são medidas que parecem estar mais relacionadas simplesmente a um tempo maior de exposição em tela do que a um desconforto. Um ponto que fortalece isso é a ocorrência de releitura, que apresentou interação *Número de Linhas: Velocidade* para ambos os grupos em nosso modelo; ou seja, a velocidade de 145 ppm proporciona um maior número de releituras significativamente no nível de 2 linhas do que em 1 linha, isto é, em legendas que naturalmente possuem um maior tempo de exibição (2 linhas). No caso da deflexão, contrário ao que inicialmente pensávamos, aparenta ser um fenômeno inerente ao processamento de textos legendados, especialmente no caso de legendas que passam mais tempo em tela (2 linhas e 145 ppm, no nosso caso), visto que, como já foi colocado, há mais chance de ocorrer algo em tela e, dessa forma, o olhar do participante se dividir entre texto e legenda.

A argumentação em torno da variável *Velocidade* e sobre se há diferença significativa no processamento de legendas de 145 ppm e 180 ppm foi mais extensa dada a não significância da maior parte das variáveis numéricas e a circunstância da significância das variáveis categóricas discutidas acima. Contudo, a partir do que foi colocado ao longo do trabalho, é possível afirmar que as legendas de 180 ppm proporcionam marginalmente um melhor processamento para ouvintes e surdos. Dessa forma, a hipótese de que não haveria diferença significativa no processamento dessas duas velocidades de legenda foi parcialmente refutada. No caso dos surdos, isso reforça os dados das pesquisas exploratórias (ARAÚJO; NASCIMENTO, 2011; ARAÚJO, 2012) e as pesquisas experimentais com rastreamento ocular realizadas com surdos no Brasil (SOUZA, 2020; VIEIRA, 2016), as quais apontaram que os surdos dos seus experimentos tiveram um melhor processamento em legendas rápidas bem segmentadas. Esse bom processamento de legendas rápidas (180 ppm) por surdos, inclusive, influencia outros parâmetros linguísticos que dizem respeito a uma legenda confortável, como o da redução linguística. Havendo menor necessidade de se reduzir o texto, menor a possibilidade de se criar um texto pouco coeso e coerente através de estratégias de edição linguística. Além disso, torna mais fácil a tarefa de legendar programas audiovisuais com uma velocidade de fala alta, visto que o número de caracteres por segundo – medida utilizada, normalmente, nesse tipo de tradução audiovisual – que pode ser utilizado é maior.

Um aspecto da metodologia do presente trabalho que se mostrou positiva foi a utilização de vídeos legendados curtos – de 8 a 15 segundos – enquanto estímulos experimentais, a despeito do que tradicionalmente os estudiosos da área preferem a partir da noção de que vídeos mais longos oferecem mais pistas contextuais e um maior período de aclimatação para o processamento das legendas. Nesse sentido, a partir do que vislumbramos

seguindo o postulado de Orero *et al.* (2018) de que a utilização de vídeos mais curtos necessita de um maior número de vídeos por condição experimental, assim como o fizemos, o uso de vídeos mais curtos na nossa metodologia não comprometeu os nossos resultados, visto que alcançamos resultados semelhantes ao que a literatura da área diz sobre o processamento de legendas de 1 linha e 2 linhas nas velocidades de 145 ppm e 180 ppm, assim como em relação ao processamento de surdos e ouvintes. Com inspiração no tipo de manipulação de estímulos da Psicolinguística, a qual também faz uso de textos curtos, nossos resultados fortalecem esse tipo de manipulação experimental com vídeos legendados. Isso é especialmente relevante a partir de questões relacionadas à pouca informação que ainda temos sobre a influência da imagem no processamento de vídeos legendados. A utilização de vídeos mais curtos fez com que fosse possível um controle experimental mais rígido sobre o tipo de cena e os aspectos cinematográficos inerentes a uma produção audiovisual como a de um filme, gênero utilizado no nosso experimento.

Isso, contudo, não elimina a chance de que o vídeo utilizado como estímulo experimental influencie o processamento das legendas além do que está sendo verdadeiramente testado a partir das variáveis independentes, como foi possível observar em nosso estudo especialmente na medida de duração média de fixação. Essa questão é pouco abordada ainda nos estudos da Tradução Audiovisual, e o fato de termos considerado o vídeo – juntando o aspecto textual da legenda de cada vídeo e sua questão imagética – como um efeito misto a ser considerado nos Modelos de Regressão Linear e Logístico Mistos é algo que, segundo nossa compreensão, deve ser adotado com mais frequência nos estudos sobre o processamento de vídeos legendados. Ademais, para além da consideração dos aspectos visuais no processamento de vídeos legendados, há também que se considerar como características imagéticas específicas agem em interação com os parâmetros linguísticos, técnicos e paralinguísticos da legendagem. Isso, entretanto, sabemos ser tarefa bem complexa, mas que deve ser uma preocupação dos estudos de processamento de legendas, visto que i) há essa lacuna no conhecimento da área e ii) essa investigação pode resultar em melhorias na recepção desse tipo de tradução audiovisual, especialmente quando consideramos o contexto da acessibilidade do nosso estudo através da LSE como ferramenta de inserção do surdo na sociedade.

Mas essa questão é ampla e necessita, primeiramente, de um modelo mais robusto de compreensão dos aspectos do processamento de textos dinâmicos e multimodais como a legendagem, semelhante à iniciativa de Just e Carpenter (1980) para o estudo de textos estáticos nos anos 1980. Apesar da relevância dos estudos de processamento de textos estáticos (JUST; CARPENTER, 1980; RAYNER, 1998; RAYNER; BARBARA; POLLATSEK, 2013;



RAYNER; LIVERSEDGE, 2004; REICHLE; RAYNER; POLLATSEK, 2003) para este e outros trabalhos semelhantes, observamos que algumas noções historicamente vistas como índice de custo de processamento não se provaram eficazes para a investigação do processamento de vídeos legendados, dada a especificidade desse tipo de texto, como foi o caso da ocorrência de regressão, releitura e deflexão. A criação desse modelo deve considerar, como já colocado, as características linguísticas, técnicas e paralinguísticas da legendagem, assim como os aspectos imagéticos supracitados. Além da interação entre os meios semióticos, esse modelo deve abarcar aspectos da cognição do público que usufrui desse tipo de texto, especialmente no caso dos surdos, que possuem uma língua visual e, na falta do *input* sonoro, as informações imagéticas, atreladas ao verbal da legendagem, podem representar benefícios ainda maiores para esse espectador do que para os ouvintes. Esse tipo de consideração sobre o sujeito surdo já foi apontado em Vieira (2016) e Vieira, Teixeira e Chaves (2017) e é algo que poderia ter sido trabalhado nesta pesquisa. Fica, portanto, a possibilidade de desdobramento para futuras pesquisas.

Outra questão que pode ser desenvolvida a partir da presente pesquisa é o estudo específico acerca da influência da ocorrência de regressão, releitura e deflexão no processamento de textos legendados por ouvintes e surdos. Isso porque, apesar das considerações que fizemos acerca de como regressão e releitura não parecem ser bons índices para estudar o custo de processamento de vídeos legendados, assim como a deflexão parece ser um aspecto inerente ao processamento desses vídeos, ainda há a necessidade de experimentos mais controlados para o estudo dessas medidas específicas. É válido ressaltar que não manipulamos questões textuais da legenda para a ocorrência dessas medidas e, por isso, houve uma pequena quantidade de dados sobre elas. Dessa forma, pesquisas que sejam mais específicas sobre essas medidas e que consigam um maior número de dados podem confirmar ou refutar as considerações que aqui fizemos sobre o seu papel no processamento de vídeos legendados.

Essa necessidade de um maior número de dados, inclusive, pode ser colocada em relação ao estudo de várias das medidas que não apresentaram significância estatística e que poderiam apontar de maneira mais categórica para a confirmação ou refutação das hipóteses do trabalho, vide duração média de encapsulamento, duração média de regressão e duração média de releitura.

Uma outra possibilidade de desdobramento dessa pesquisa é o estudo de como número de linhas e velocidade (145 e 180 ppm) influenciam o processamento de outros gêneros audiovisuais legendados. É possível que a análise de outros gêneros audiovisuais que

estabeleçam relações distintas entre imagem e texto legendado apresentem resultados diferentes, principalmente em relação ao número de linhas. Historicamente, por exemplo, a LSE de propagandas políticas tem sido confeccionada com legendas de 1 linha no Brasil e tem se mostrado efetiva (ARAÚJO, 2009). Nesse caso, é possível que legendas de 1 linha com maior redundância em relação ao que acontece na imagem sejam processadas de forma mais eficiente do que legendas de 2 linhas, conforme Praet *et al.* (1990) vislumbram? Este trabalho não conseguiu vislumbrar esse aspecto por questões de objetivos e procedimentos de análise, mas é uma possibilidade que se configura como viável para o futuro. Inclusive, diferentemente do que foi colocado por Praet *et al.* (1990), em um contexto com imagens – vídeos legendados do experimento –, as legendas de 2 linhas proporcionaram um melhor desempenho aos participantes do estudo em termos de processamento.

Em relação à utilização de Modelos de Regressão Linear e Logística Mistos no trabalho, esse tipo de tratamento estatístico se mostrou efetivo e permitiu que estimássemos valores para várias das medidas analisadas, como a duração média de fixação. Esse tipo de modelagem estatística, inclusive, estima, através do seu  $R^2$  condicional, o quanto um fenômeno (medida) é explicado a partir de uma variável independente e de efeitos mistos (participante e vídeo, no nosso caso). Esse dado fez com que descobríssemos que o número de linhas (1 ou 2) explica 9,4% do fenômeno de duração média de fixação, um número relativamente baixo; ou seja, há outros fatores que influenciam a duração média de fixação além do número de linhas de uma legenda. A investigação de como outros fatores podem influenciar uma medida como a duração média de fixação no processamento de vídeos legendados pode levar à criação de um modelo teórico para o estudo desse tipo de processamento citado anteriormente. Isso passa pela aplicação desse tipo de modelagem em outros trabalhos experimentais com legendagem, algo ainda pouco realizado.

Um outro desdobramento dessa pesquisa passa pela análise dos relatos retrospectivos e questionários pós-coleta que os participantes deste estudo responderam após assistirem aos vídeos do experimento. Apesar da ênfase na análise da movimentação ocular dos participantes e de acreditarmos que esse dado quantitativo é o principal a ser levado em consideração para o estudo do processamento de legendas, é possível que a análise desses relatos sobre a compreensão dos participantes nos auxilie a responder às questões desta pesquisa com mais propriedade, mesmo que sejam dados mais complementares do que essenciais. Isso, principalmente, considerando uma análise estatística desses dados qualitativos.

Acreditamos que os resultados deste trabalho fortaleçam a área de conhecimento sobre o processamento de vídeos legendados por ouvintes e, principalmente, por surdos,

público-alvo principal desta pesquisa. Dessa forma, os achados aqui dispostos podem melhorar ainda mais a recepção de LSE por surdos brasileiros, dando continuidade à pesquisa com Tradução Audiovisual Acessível realizada na Universidade Estadual do Ceará desde 2002 e contribuindo ainda mais para a inclusão social através da informação dos surdos do Brasil.

## REFERÊNCIAS

- A REDE social. Direção: David Fincher. Produção: Scott Rudin, Dana Brunetti, Michael de Luca e Ceán Chaffin. Roteiro: Aaron Sorkin e Ben Mezrich. Música: Trent Reznor e Atticus Ross. Columbia Pictures, 2010. 1 DVD (121 min.), son., color.
- ARAÚJO, V. L. S. Por um modelo de legendagem para Surdos no Brasil. **Tradução e Comunicação**, Revista Brasileira de Tradutores, São Paulo: UN BERO, n. 17, p. 59–76, 2008.
- ARAÚJO, V. L. S. In Search of SDH Parameters for Brazilian Party Political Broadcasts. **The Sign Language Translator and Interpreter**, Manchester: St. Jerome Publishing Company, v. 3, n. 2, p. 157-167, 2009.
- ARAÚJO, V. L. S.; NASCIMENTO, A. K. P. Investigando parâmetros de legendas para Surdos e Ensurdecidos no Brasil. **Tradução em Revista**, v. 2, p. 1- 18, 2011. Disponível em: <https://www.maxwell.vrac.puc-rio.br/18862/18862.PDF>. Acesso em: 10 jan. 2018.
- ARAÚJO, V. L. S. **Legendagem para surdos**: em busca de um modelo para o Brasil. Relatório Técnico n. 306948/2008-7. Fortaleza: CNPq, Fev./2012.
- ARAÚJO, V. L. S. **A segmentação na legendagem para surdos e ensurdecidos (LSE)**: um estudo baseado em corpus. Relatório Técnico n. 306441/2011-0. Fortaleza: CNPq, Fev./2015.
- ARAÚJO, V. L. S.; ASSIS, Í. A. P. A segmentação linguística na legendagem para surdos e ensurdecidos (LSE) de ‘Amor Eterno Amor’: uma análise baseada em corpus. **Letras e Letras**, Uberlândia, v. 30, n. 2, p. 156-184, jul./dez. 2014.
- ASSIS, Í. A. P. **A segmentação na LSE de amor eterno amor**: uma análise baseada em corpus. 2013. 71f. Monografia (Bacharelado em Letras Inglês) – Universidade Estadual do Ceará, Fortaleza- CE, 2013.
- ASSIS, Í. A. P. **Legendagem para Surdos e Ensurdecidos (LSE)**: análise baseada em corpus da segmentação linguística em Amor Eterno Amor. 2016. 107f. Dissertação (Mestrado em Linguística Aplicada) – Programa de Pós-graduação em Linguística Aplicada, Universidade Estadual do Ceará, Fortaleza, 2016.
- CARVALHO, F.S.F.; SEOANE, A.F. Relação entre unidades de velocidade de leitura na legendagem para surdos e ensurdecidos. **Tradução Audiovisual Acessível**. Transversal – Revista em Tradução, Fortaleza, v. 5, n. 9, p.137-153, 2019. Disponível em: <http://periodicos.ufc.br/transversal/article/view/41951>. Acesso em: 4 jan. 2020.
- CASTILHO, A. de. **Nova Gramática do Português Brasileiro**. São Paulo: Contexto, 2012. ISBN: 978-85-7244-462-0.
- CHAVES, É. G. **Legendagem para Surdos e Ensurdecidos**: um estudo baseado em corpus da segmentação nas legendas de filmes brasileiros em DVD. 2012. 126 f. Dissertação (Mestrado em Linguística Aplicada) – Programa de Pós-graduação em Linguística Aplicada, Universidade Estadual do Ceará, Fortaleza-CE, 2012.

CHAVES, É. G. **Legendagem de filmes em língua inglesa e portuguesa: análise dos parâmetros de segmentação e velocidade de legendas para ouvintes**. 2017. 218f. Tese (Doutorado em Linguística Aplicada) – Programa de Pós-graduação em Linguística Aplicada, Universidade Estadual do Ceará, Fortaleza-CE, 2017.

DE LINDE, Z.; KAY, N. **The semiotics of subtitling**. Manchester: St. Jerome Publishing, 1999.

DINIZ, N. S. L. **A Segmentação em Legendagem para Surdos e Ensurdidos: um Estudo Baseado em Corpus**. 2012. 149 f. Dissertação (Mestrado em Linguística) – Universidade Federal de Minas Gerais, Belo Horizonte-MG, 2012.

DÍAZ CINTAS, J; MUÑOZ-SÁNCHEZ, P. Fansubs: audiovisual translation in an amateur environment. **The Journal of Specialised Translation**, Londres, n. 6, jul. 2006.

DÍAZ CINTAS, J; REMAEL, A. **Audiovisual Translation: Subtitling**. Manchester, UK, Kinderhook, NY, UK: St. Jerome Publishing, 2007. ISBN: 978-1900650-95-3/1-900650-95-9.

DÍAZ CINTAS, J. Teaching and learning to subtitle in an academic environment. In: DÍAZ CINTAS, J. (ed.). **The Didactics of Audiovisual Translation**. Amsterdam and Philadelphia: John Benjamins, 2008. p. 89-103.

D'YDEWALLE, G.; MUYLLE, P.; VAN RENSBERGEN, J. Attention shifts in partially redundant information situation. In: GRONER, R.; McCONKIE, G. W.; MENZ, C. (Eds.). **Eye movements and human information processing**. Amsterdam: North-Holland, 1985.

D'YDEWALLE, G.; VAN RENSBERGEN, J.; V.; POLLET, J. Reading a message when the same message is available auditorily in another language: the case of subtitling. In: O'REGAN, J.K.; LÉVY-SCHOEN, A. (org.). **Eye Movements: From Physiology to Cognition**. Amsterdã; Nova York: Elsevier Science Publishers, 1987. p. 313-321.

D'YDEWALLE, G.; WARLOP, L.; VAN RENSBERGEN, J. Differences between young and older adults in the division of attention over different sources of TV information. **Medienpsychologie: Zeitschrift für Individual- und Massenkommunikation**, v. 1, p. 42-57, 1989.

D'YDEWALLE, G.; PRAET, C.; VERFAILLIE, K.; VAN RENSBERGEN, J. Choosing between redundant information channels: Speech and text. **Communication Research**, v. 18, p. 650-666, 1991.

D'YDEWALLE, G.; GIELEN, I. Attention allocation with overlapping sound, image, and text. In: RAYNER, K. (Ed.) **Eye movements and visual cognition: Scene perception and reading**. New York: Springer-Verlag, 1992. p. 415-427.

D'YDEWALLE, G.; DE BRUYCKER, W. Eye movements of children and adults while reading television subtitles. **European Psychologist**, v. 12, n. 3, p. 196-205, 2007.

DUCHOWSKI, A. **Eye-tracking methodology**. Londres: Springer, 2007.

EMMOREY, K. *et al.* Eye gaze during comprehension of American sign language by native and beginning signers. **Journal of Deaf and Deaf Education**, Oxford University Press, p. 237-243, 2008. Disponível em: <http://jdsde.oxfordjournals.org/>. Acesso em: 20 jan. 2017.

FRANCO, E.; ARAÚJO, V. L. S. Reading Television: Checking deaf people's Reactions to Closed Subtitling in Fortaleza, Brazil. **The Translator**, v. 9, n. 2, p. 249-267, 2003.

GABRIEL, M. H. C. **Problemas de segmentação linguística na legendagem para surdos e ensurdecidos (LSE) de “Cheias de Charme”**: uma análise baseada em corpus. 2015. 108 f. Dissertação (Mestrado em Linguística Aplicada) – Programa de Pós-graduação em Linguística Aplicada, Universidade Estadual do Ceará, Fortaleza-CE, 2015.

GHIA, E. The Impact of Translation Strategies on Subtitle Reading. In: PEREGO, E. (ed.) **Eye Tracking in Audiovisual Translation**. Aracne: Roma, 2012. p. 157–182.

GODOY, M. C. Introdução aos modelos lineares mistos para os estudos da linguagem. **PsyArXiv**. <https://doi.org/10.17605/OSF.IO/9T8UR>. Disponível em: <https://osf.io/rbgnq/>. Acesso em: 24 nov. 2020.

IVARSSON, J.; CARROLL, M.; **Subtitling**. Simrishamm, Suécia: TransEditHB, 1998.

JAKOBSON, R. On linguistic aspects of translation. In: BROWER, R. A. (Ed.). **On Translation**. Cambridge, MA: Harvard University Press, 1959. p. 232-239.

JENSEMA, C. Viewer reaction to different television captioning speeds. **American Annals of the Deaf**, v. 143, n. 4, p. 318–324, 1998.

JENSEMA, C.; SHARKAWY, S.; DANTURTHI, R.; BURCH, R.; HSU, D. Eye movement patterns of captioned television viewers. **American Annals of the Deaf**, v. 145, n. 3, p. 275–285, 2000.

JUST, M. C.; CARPENTER, P. A. A Theory of Reading: from eye fixations to comprehension. **Psychological Review**, v. 87, n. 4, p. 329-354, 1980.

KARAMITROGLOU, F. A Proposed Set of Subtitling Standards in Europe. **Translation Journal**, v. 2, n. 2, p. 1-15, 1998. Disponível em: <http://translationjournal.net/journal//04stndrd.htm>. Acesso em: 15 mar. 2011.

KRUGER, J. L.; STEYN, F. Subtitles and eye tracking: reading and performance. **Reading Research Quarterly**, v. 49, n. 1, p. 105-120, 2014. <https://doi.org/10.1002/rrq.59>. Disponível em: [https://www.researchgate.net/publication/259542714\\_Subtitles\\_and\\_Eye\\_Tracking\\_Reading\\_and\\_Performance](https://www.researchgate.net/publication/259542714_Subtitles_and_Eye_Tracking_Reading_and_Performance). Acesso em: 5 mar. 2020.

KRUGER, J. L.; SZARKOWSKA, A.; KREJTZ, I. Subtitles on the moving image: an overview of eye tracking studies. **Refractory: a journal of entertainment media**, v. 25, p. 1-14, 2015. Disponível em: <http://refractory.unimelb.edu.au/2015/02/07/kruger-szarkowska-krejtz/>. Acesso em: 28 fev. 2020.

MARQUER, P. La méthode expérimentale: quelques points de repère. In: NICOLAS, S. (Ed.). **La psychologie cognitive**. Paris: Armand Colin, 2003, p. 185-206.

MONTEIRO, S. V. M. **Legendagem para surdos e ensurdecidos (LSE):** um estudo sobre a segmentação na campanha política de 2010. 230f. Tese (Doutorado em Linguística Aplicada) – Programa de Pós-graduação em Linguística Aplicada, Universidade Estadual do Ceará, Fortaleza-CE, 2016.

ORERO, P.; DOHERTY, S.; MATAMALA, A.; KRUGER, J. L.; PEDERSEN, J.; PEREGO, E.; ROMERO-FRESCO, P.; ROVIRA-ESTEVA, S.; SOLER, O.; SZARKOWSKA, A. Conducting experimental research in audiovisual translation (AVT): a position paper. **Journal of specialised translation**, v. 30, p. 105-126, 2018. Disponível em: [https://www.jostrans.org/issue30/art\\_orero\\_et\\_al.pdf](https://www.jostrans.org/issue30/art_orero_et_al.pdf). Acesso em: 1 mar. 2020.

PEREGO, E.; DEL MISSIER, F.; PORTA, M.; MOSCONI, M. The cognitive effectiveness of subtitle processing. **Media Psychology**. Philadelphia, PA: Routledge, 2010. p. 243-272. Disponível em: <http://www2.units.it/delmisfa/papers/SubtitlesProcessing2010.pdf>. Acesso em: 24 out. 2016.

PEREGO, E. Introduction. In: PEREGO, E. (ed.). **Eye Tracking in Audiovisual Translation**. Aracne: Roma, 2012. p. 7-11.

PRAET, C.; VERFAILLIE, K.; DE GRAEF, P.; VAN RENSBERGEN, J.; D'YDEWALLE, G. A one line text is not half a two line text. In: GRONER, R.; D'YDEWALLE, G.; PARHAM, R. (Ed.). **From eye to mind: Information acquisition in perception, search, and reading**. Amsterdam: Elsevier Science Publishers B.V. (North-Holland), 1990. p. 205-213.

QUARESMA, P.; PINHO, A. Análise de frequências da língua portuguesa. In: CONFERÊNCIA IBERO-AMERICANA INTERTIC, 2007. **Livro de Actas da Conferência Ibero-Americana InterTIC 2007**. Porto: IASK, 2007. p. 267-272.

RAJENDRAN, D.J.; DUCHOWSKI, A. T.; ORERO, P.; MARTINEZ, J., ROMERO-FRESCO, P. Effects Of text chunking on subtitling: A quantitative and qualitative examination. **Perspectives: studies in translatology**, v. 21, n. 1, p. 5-21, 2013.

RAYNER, K. Eye movements in reading and information processing: 20 years of research. **Psychological Bulletin**, v. 124, n. 3, p. 372-422, 1998.

RAYNER, K.; LIVERSEDGE, S. P. Visual and linguistic processing during eye fixation in reading. In: HENDERSON, J. M.; FERREIRA, F. (Org.). **Language, Vision and Action**. New York: Psychology, 2004. p. 56-104.

RAYNER, K.; BARBARA, J.; POLLATSEK, A. Movimentos oculares durante a leitura. In: SNOWLING, M. J.; HULME, C. (Org.). **A ciência da leitura**. Tradução de Ronaldo Cataldo Costa. Porta Alegre: Penso, 2013. p. 97-116.

REICHLER, E. D.; RAYNER, K.; POLLATSEK, A. The E-Z Reader model of eye-movement control in reading: comparisons to other models. **Behavioral and Brain Sciences**, Cambridge University Press, v. 26, p. 445-526, 2003.

ROMERO-FRESCO, P. More haste than speed: Edited versus verbatim respoken subtitles. **Vigo International Journal of Applied Linguistics**, v. 6, p. 109–133, 2009. Disponível em: [https://www.researchgate.net/publication/228502664\\_More\\_haste\\_less\\_speed\\_Edited\\_versus\\_verbatim\\_respoken\\_subtitles](https://www.researchgate.net/publication/228502664_More_haste_less_speed_Edited_versus_verbatim_respoken_subtitles). Acesso em: 15 mar. 2020.

SOUZA, G. T. 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**. 82f. 2020. Dissertação (Mestrado em Estudos da Tradução) – Programa de Pós-Graduação em Estudos da Tradução, Universidade Federal do Ceará, Fortaleza-CE, 2020.

SZARKOWSKA, A.; GERBER-MORÓN, O. Two or three lines: a mixed-methods study on subtitle processing and preferences. **Perspectives: studies in translation theory and practice**, v. 27, n. 1, p. 144-164, 2018.

TOBII STUDIO. **An introduction to eye tracking and Tobii Eye Trackers**. Tobii Technologi AB, 2010. Disponível em: [http://www.tobii.com/Global/Analysis/Training/WhitePapers/Tobii\\_EyeTracking\\_Introduction\\_WhitePaper.pdf](http://www.tobii.com/Global/Analysis/Training/WhitePapers/Tobii_EyeTracking_Introduction_WhitePaper.pdf). Acesso em: 10 jun. 2014.

VIEIRA, P. **A influência da segmentação e da velocidade na recepção de legendas para surdos e ensurdecidos**. 248f. 2016. Tese (Doutorado em Linguística Aplicada) – Programa de Pós-Graduação em Linguística Aplicada, Universidade Estadual do Ceará, Fortaleza-CE, 2016.

VIEIRA, P.; TEIXEIRA, E. N.; CHAVES, E. G. Caminhos do olhar: a movimentação ocular de espectadores surdos durante a exibição de vídeos legendados. **Trabalhos em Linguística Aplicada**. Dossiê: Tradução Audiovisual Acessível, v. 56, n. 2, p. 493-526, 2017. Disponível em: <http://www.scielo.br/pdf/tla/v56n2/2175-764X-tla-56-02-00493.pdf>. Acesso em: 15 jan. 2019.



## APÊNDICE A – PERFIL PRÉ-COLETA DOS PARTICIPANTES OUVINTES

**FICHA DO PARTICIPANTE de nº: \_\_\_\_\_**

**Nome:** \_\_\_\_\_

**E-mail:** \_\_\_\_\_

**Sexo:** ( ) feminino ( ) masculino

**Idade:** \_\_\_\_ anos.

**Data de nascimento:** \_\_\_\_ / \_\_\_\_ / \_\_\_\_:

**Estado de nascimento:** Ceará ( ) Outro ( ) \_\_\_\_\_

**Se marcou outro, desde quando (que ano) mora no Ceará?** \_\_\_\_\_.

**Formação:**

( ) Ensino superior completo.

( ) Ensino superior incompleto.

( ) Ensino Médio \_\_\_\_ ano.

**1. Que nota você daria para sua leitura em Língua Portuguesa? Escolha:**

1 2 3 4 5 6 7 8 9 10

**2. Você tem alguma dificuldade com a visão? ( ) não ( ) sim**

**3. 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) \_\_\_\_\_

**4. Você usa óculos ou lentes de contato para a dificuldade com a visão? ( ) sim ( ) não**

**5. Sua lateralidade é:**

- ( ) Destra (escreve com a mão direita)  
( ) Canhota (escreve com a mão esquerda)

**6. Você tem algum problema neurológico? ( ) sim ( ) não**

**7. Se sim, qual?** \_\_\_\_\_

**8. Você tem algum problema psiquiátrico, como por exemplo, depressão? ( ) sim ( ) não**

**9. Se sim, qual?** \_\_\_\_\_

**10. Você faz uso de remédios controlados ou drogas que afetam o sistema nervoso central? ( ) sim ( ) não**

**11. Em caso de resposta afirmativa à questão 10, qual?**  
\_\_\_\_\_

**12. Costuma ver filmes ou programas de TV legendados?**  
\_\_\_\_\_**13. Em caso de resposta afirmativa à questão 12, com que frequência?**

( ) Todos os dias ( ) 3 vezes por semana ( ) 2 vezes por semana ( ) raramente

**14. Você costuma ir ao cinema? ( ) sim ( ) não**

**15. Em caso de resposta afirmativa à questão 14, que tipo de filme você prefere ver?**  
\_\_\_\_\_

**16. Você costuma alugar filmes na locadora ou assistir no computador? ( ) sim ( ) não**

**17. Em caso de resposta afirmativa à questão 16, que tipo de filme você prefere ver?**  
\_\_\_\_\_

**18. Você já assistiu ao filme 'A Rede Social'? ( ) sim ( ) não**

**19. Em caso de resposta afirmativa à questão 18, fale um pouco sobre o filme:**  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

**APÊNDICE B – PERFIL PRÉ-COLETA DOS PARTICIPANTES SURDOS****FICHA DO PARTICIPANTE de nº: \_\_\_\_\_**

Nome: \_\_\_\_\_

E-mail: \_\_\_\_\_

Sexo: ( ) feminino ( ) masculino

Idade: \_\_\_\_ anos.

Data de nascimento: \_\_\_\_/\_\_\_\_/\_\_\_\_:

Estado de nascimento: Ceará ( ) Outro ( ) \_\_\_\_\_

Se marcou outro, desde quando (que ano) mora no Ceará? \_\_\_\_\_.

**Formação:**

( ) Ensino superior completo.

( ) Ensino superior incompleto.

( ) Ensino Médio \_\_\_\_ ano.

**1. Quando você ficou surdo?**

( ) Eu nasci surdo.

( ) \_\_\_\_ meses.

( ) \_\_\_\_ anos.

**2. Qual seu grau de surdez?**

( ) surdez leve

( ) surdez moderada

( ) surdez profunda

( ) \_\_\_\_\_

**3. Onde você aprendeu Libras pela primeira vez?**

( ) Na família

( ) Na escola

( ) Na igreja

( ) \_\_\_\_\_

**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? Escolha:**

1 2 3 4 5 6 7 8 9 10

**6. Que nota você daria para sua leitura em Língua Portuguesa? Escolha:**

1 2 3 4 5 6 7 8 9 10

**7. Você tem alguma dificuldade com a visão? ( ) sim ( ) não**

**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:**

( ) Destra (escreve com a mão direita)

( ) Canhota (escreve com a mão esquerda)

**17. Costuma ver filmes ou programas de TV legendados?**

( ) sim ( ) não

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

( ) Todos os dias ( ) 3 vezes por semana ( ) 2 vezes por semana ( ) raramente

**19. Você costuma ir ao cinema? ( ) sim ( ) não**

20. Em caso de resposta afirmativa à questão 19, que tipo de filme você prefere ver?

---

21. Você costuma alugar filmes na locadora ou assistir no computador? ( ) sim ( ) não

22. Em caso de resposta afirmativa à questão 21, que tipo de filme você prefere ver?

---

23. Você já assistiu ao filme 'A Rede Social'? ( ) sim ( ) não

24. Em caso de resposta afirmativa à questão 23, fale um pouco sobre o filme:

---

---

---

## APÊNDICE C – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO PARA OUVINTES

### TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Pretendemos desenvolver uma pesquisa intitulada **A Influência do número de linhas e velocidade no processamento de legendas de filmes por surdos e ouvintes: Um estudo experimental com rastreador ocular**. A pesquisa tem como objetivo analisar experimentalmente os parâmetros da Legendagem para Surdos e Ensurdidos (LSE) produzida no Brasil, a fim de orientar futuros legendistas na prática da legendagem. Entendemos que essa pesquisa permite a elaboração de padrões sistemáticos que possam fornecer subsídios tanto ao instrutor, quanto ao legendista em formação para uma melhoria na teoria e prática da legendagem. Estamos contribuindo tanto para a pesquisa em LSE quanto à formação de legendistas em cursos de graduação e pós-graduação. Por isso, solicitamos a sua contribuição como ouvinte, que assistirá às produções audiovisuais com legendas, tendo o seu movimento ocular rastreado pelo *Eye Tracker*. O conteúdo será gravado, no intuito de garantir a integridade das informações prestadas em relação ao objeto investigado. Os momentos para as aplicações dos instrumentos serão acordados previamente, contando com todos os esclarecimentos necessários à realização exitosa das mesmas. Os benefícios em participar da pesquisa consistem na oportunidade de pensar criteriosamente sobre o processo do qual está participando, exercitando a reflexão e a criticidade, bem como de poder contribuir diretamente com a melhoria da pesquisa e da prática em legendagem. Os possíveis desconfortos de sua participação podem ser gerados pela possibilidade do não entendimento de algumas das perguntas do relato guiado. Buscaremos de todas as formas minimizar esses desconfortos explicando melhor o conteúdo e o objetivo da pergunta. Esclarecemos que a pesquisa não lhe trará nenhum ônus e que você tem a liberdade para participar ou não da mesma e, ainda, de se recusar a responder às perguntas, bem como de retirar o seu consentimento a qualquer momento. Declaramos que não haverá divulgação personalizada das informações, que você não receberá qualquer reembolso ou gratificação devido à participação neste estudo e terá o direito a uma via deste Termo. Se necessário, você poderá entrar em contato a profissional responsável pela pesquisa para esclarecimento de eventuais dúvidas, a professora **Vera Lúcia Santiago Araújo** que pode ser encontrada pelos telefone(s) (85) 3101-2032, (85) 99995-1185, email [vera.santiago@uece.br](mailto:vera.santiago@uece.br) e/ou [verainnerlight@uol.com.br](mailto:verainnerlight@uol.com.br). Se você tiver alguma consideração ou dúvida sobre a ética da pesquisa, entre em contato com o Comitê de Ética em Pesquisa – CEP/UECE, telefone: (85) 3101 – 9890 ou email [CEP@uece.br](mailto:CEP@uece.br).

---

Assinatura do Responsável pela Pesquisa

---

De posse das informações sobre a pesquisa **A Influência do número de linhas e velocidade no processamento de legendas de filmes por surdos e ouvintes: Um estudo experimental com rastreador ocular**, concordo voluntariamente em participar da mesma, de forma livre e esclarecida.

Nome: \_\_\_\_\_

Assinatura: \_\_\_\_\_

Fortaleza (CE), \_\_\_\_\_ de \_\_\_\_\_ de \_\_\_\_\_.

## APÊNDICE D – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO PARA SURDOS

### TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Pretendemos desenvolver uma pesquisa intitulada **A Influência do número de linhas e velocidade no processamento de legendas de filmes por surdos e ouvintes: Um estudo experimental com rastreador ocular**. A pesquisa tem como objetivo analisar experimentalmente os parâmetros da Legendagem para Surdos e Ensurdidos (LSE) produzida no Brasil, a fim de orientar futuros legendistas na prática da legendagem. Entendemos que essa pesquisa permite a elaboração de padrões sistemáticos que possam fornecer subsídios tanto ao instrutor, quanto ao legendista em formação para uma melhoria na teoria e prática da legendagem. Estamos contribuindo, tanto para a pesquisa em LSE, quanto à formação de legendistas em cursos de graduação e pós-graduação. Por isso, solicitamos a sua contribuição como surdo(a), que assistirá às produções audiovisuais com legendas, tendo o seu movimento ocular rastreado pelo *Eye Tracker*. O conteúdo será gravado, no intuito de garantir a integridade das informações prestadas em relação ao objeto investigado. Os momentos para as aplicações dos instrumentos serão acordados previamente, contando com todos os esclarecimentos necessários à realização exitosa das mesmas. Os benefícios em participar da pesquisa consistem na oportunidade de pensar criteriosamente sobre o processo do qual está participando, exercitando a reflexão e a criticidade, bem como de poder contribuir diretamente com a melhoria da pesquisa e da prática em legendagem. Os possíveis desconfortos de sua participação podem ser gerados pela possibilidade do não entendimento de algumas das perguntas do relato guiado. Buscaremos de todas as formas minimizar esses desconfortos explicando melhor o conteúdo e o objetivo da pergunta. Esclarecemos que, a pesquisa não lhe trará nenhum ônus e que você tem a liberdade para participar ou não da mesma e, ainda, de se recusar a responder às perguntas, bem como de retirar o seu consentimento a qualquer momento. Declaramos que não haverá divulgação personalizada das informações, que você não receberá qualquer reembolso ou gratificação devido à participação neste estudo e terá o direito a uma via deste Termo. Se necessário, você poderá entrar em contato a profissional responsável pela pesquisa para esclarecimento de eventuais dúvidas, a professora **Vera Lúcia Santiago Araújo** que pode ser encontrada pelos telefone(s) (85) 3101-2032, 9995-1185, email [vera.santiago@uece.br](mailto:vera.santiago@uece.br) e/ou [verainnerlight@uol.com.br](mailto:verainnerlight@uol.com.br). Se você tiver alguma consideração ou dúvida sobre a ética da pesquisa, entre em contato com o Comitê de Ética em Pesquisa – CEP/UECE, telefone: (85) 3101 – 9890 ou email [CEP@uece.br](mailto:CEP@uece.br).

---

Assinatura do Responsável pela Pesquisa

---

De posse das informações sobre a pesquisa **A Influência do número de linhas e velocidade no processamento de legendas de filmes por surdos e ouvintes: Um estudo experimental com rastreador ocular**, concordo voluntariamente em participar da mesma, de forma livre e esclarecida.

Nome: \_\_\_\_\_

Assinatura: \_\_\_\_\_

Fortaleza (CE), \_\_\_\_\_ de \_\_\_\_\_ de \_\_\_\_\_.