O IMPACTO DA VELOCIDADE DE LEGENDAS E A PRESENÇA DE VÍDEO NA LEITURA DE LEGENDAS: UM ESTUDO DE REPLICAÇÃO COM RASTREAMENTO OCULAR[1]

Agnieszka Szarkowska; Valentina Ragni; David Orrego-Carmona; Sharon Black; Sonia Szkriba; Jan-Louis Kruger; Krzysztof Krejtz; Breno Silva

Tradução de Hector Webber Tech[2]

Resumo: Apresentamos os resultados de uma replicação direta do estudo de Liao *et al.* (2021) sobre como a velocidade das legendas e a presença de vídeo simultâneo impactam a leitura de legendas entre espectadores britânicos e poloneses. Nosso objetivo foi avaliar a generalização do resultado do estudo original em um grupo de participantes australianos falantes de inglês. O estudo explorou os efeitos tanto no nível da legenda quanto no nível da palavra, considerando a presença ou ausência de vídeo simultâneo e três velocidades de legenda: 12, 20 e 28 caracteres por segundo (cps). No geral, a maioria dos resultados originais foi replicada, confirmando que a presença de vídeo e a velocidade das legendas têm um impacto mensurável no processamento entre diferentes grupos de espectadores. Além disso, surgiram diferenças em como falantes nativos e não nativos processam as legendas, especialmente relacionadas a *wrap up* e efeito da frequência e comprimento das palavras. O presente artigo descreve a replicação e os resultados de maneira detalhada e discute algumas de suas implicações

Palavras-chave: Legendagem, Velocidade de leitura, Velocidade da legenda, Replicação, Vídeo simultâneo, Língua estrangeira, Processamento em L2.

THE IMPACT OF VIDEO AND SUBTITLE SPEED ON SUBTITLE READING

Abstract: We present the results of a direct replication of Liao *et al*.'s (2021) study on how subtitle speed and the presence of concurrent video impact subtitle reading among British and Polish viewers. Our goal was to assess the generalisability of the original study's findings on a cohort of Australian English participants. The study explored both subtitle-level and word-level effects, considering the presence or absence of concurrent video and three subtitle speeds: 12, 20, and 28 characters per second (cps). Overall, most of the original results were replicated, confirming that the presence of video and the speed of the subtitles have a measurable impact on processing across different viewer groups. Additionally, differences in how native and non-native speakers process subtitles emerged, particularly related to wrap-up, word frequency and word length effects. The paper describes the replication in detail, presents the findings, and discusses some of their implications.

Keywords: Subtitling, Reading speed, Subtitle speed, Replication, Concurrent video, Foreign language, L2 processing.

Introdução

[1] Texto em língua inglesas: https://mail.jatjournal.org/index.php/jat/article/view/283
[2] Revisão da tradução: Profa. Dra. Márcia Moura da Silva.

No cenário audiovisual atual, as legendas tornaram-se indispensáveis para milhões de pessoas em todo o mundo. Essa tendência tem sido impulsionada por plataformas de streaming como Netflix, Amazon Prime ou Disney+, além de regulamentos de acessibilidade, como a Lei dos Americanos com Deficiência (Americans with Disabilities Act) e a Diretiva de Serviços de Comunicação Social Audiovisual (AVMSD) da União Europeia. As legendas não apenas auxiliam pessoas surdas ou com deficiência auditiva, mas também ajudam os ouvintes que enfrentam dificuldades com áudio pouco claro, fala rápida ou sotaques desafiadores (Díaz Cintas; Remael, 2021; Zajechowski, 2022). Elas também são indispensáveis para que falantes não nativos entendam conteúdo estrangeiro e são frequentemente usadas no aprendizado de idiomas (Vanderplank, 2016; Wang; Pellicer-Sánchez, 2022).

Apesar da ubiquidade das legendas e do crescente corpo de pesquisa sobre o seu processamento, ainda sabemos pouco sobre como diferentes grupos de espectadores interagem com conteúdo legendado, sobretudo como a própria presença de imagens em movimento e o aumento das velocidades das legendas afetam o processamento e a compreensão dos espectadores. As pesquisas sobre o processamento de legendas em geral, e a velocidade das legendas em particular, remontam à década de 1980 e continuam até hoje (d'Ydewalle *et al*., 1985; d'Ydewalle *et al*., 1987; Jensema, 1998; Koolstra *et al*., 1999; Szarkowska; Gerber-Morón, 2018; Kruger *et al*., 2022; Szarkowska *et al*., 2021). No entanto, a generalização desses resultados é prejudicada por variações nas metodologias, nos gêneros audiovisuais, nas durações dos vídeos, nos perfis demográficos dos espectadores, nas métricas de movimento ocular, nos cálculos de velocidade ou nos tipos de legendas.

Nosso estudo aborda essa lacuna ao conduzir uma replicação direta, conforme definida por LeBel *et al.* (2018), da pesquisa de Liao *et al*. (2021). Nosso objetivo é determinar se os resultados obtidos no estudo original, que se concentrou em uma amostra de espectadores australianos, também se aplicam a falantes nativos britânicos e falantes poloneses que têm o inglês como segunda língua. Trabalhamos com a hipótese de que os participantes poloneses, por não serem falantes nativos de inglês, teriam mais dificuldade com a tarefa do que os participantes britânicos, que, por outro lado, exibiriam padrões de leitura mais consistentes e menor dificuldade na execução da tarefa.

O artigo descreve a replicação em detalhe, apresenta os resultados e discute algumas de suas implicações.

1. Replicação

Ao realizar uma replicação direta, empregando os mesmos métodos e materiais do estudo original, pretendemos avançar em nossa compreensão de como diferentes perfis de espectadores interagem com vídeos legendados em diversas regiões do mundo. Até recentemente, as replicações de estudos sobre tradução audiovisual (TAV) têm sido raras (Díaz Cintas; Szarkowska, 2020). Além disso, estudos experimentais em TAV enfrentaram problemas, como número pequeno de amostras, relatos seletivos, designs experimentais com baixo poder estatístico, proporções de efeito reduzidas e análises estatísticas simples, o que, muitas vezes, leva à falta de confiabilidade dos resultados. Diante desses desafios, embora o número de estudos experimentais de rastreamento ocular sobre legendagem, foco principal desta investigação, tenha aumentado nos últimos anos, não podemos afirmar com total confiança que os resultados relatados nesses estudos sejam definitivos e universalmente aplicáveis. Portanto, em consonância com a perspectiva da Open Science Collaboration, acreditamos que "ainda há mais trabalho a ser feito para verificar se sabemos o que achamos que sabemos" (Open Science Collaboration 2015, p. 943).

A replicação frequentemente enfrenta resistência de alguns periódicos, e editores rejeitam artigos com a justificativa de falta de inovação, acrescentando observações como "já sabemos disso" (Open Science Collaboration 2015, p. 943). No entanto, como destacam Koole e Lakens (2012), a replicação é essencial para sustentar o progresso científico e garantir a confiabilidade dos resultados de pesquisa. Segundo Schmidt (2009), a replicação é "uma das ferramentas mais importantes para a verificação de fatos nas ciências empíricas" (Schmidt 2009, p. 90). Um impulso significativo para o movimento de replicação veio do *Reproducibility Project* (Projeto de Reprodutibilidade) realizado pela Open Science Collaboration (2015), que conduziu replicações de 100 estudos publicados em importantes revistas de psicologia. O projeto identificou uma disparidade marcante: apenas 36% dos estudos replicados obtiveram resultados significativos, em comparação com os 97% relatados pelos estudos originais. No entanto, é importante notar que isso não implica, necessariamente, falhas nos estudos originais, mas pode ser atribuído a erros aleatórios ou sistemáticos (Open Science Collaboration, 2015).

Há pelo menos duas lições a serem tiradas disso: em primeiro lugar, a replicação deve ser integrada à estrutura da pesquisa científica, incluindo o campo da TAV. Em segundo lugar, para que os estudos sejam replicáveis, os pesquisadores devem aderir aos princípios da Open Science e tornar seus conjuntos de dados, protocolos e resultados acessíveis à comunidade

científica para análise minuciosa. Nosso estudo está alinhado com esses princípios, e todos os conjuntos de dados estão acessíveis para fins de transparência e pesquisas futuras.

2. O estudo original

No estudo original de Liao *et al*. (2021, p. 31), falantes nativos de inglês australiano assistiram a vídeos sem som, acompanhados de legendas em inglês exibidas em três velocidades: 12 caracteres por segundo (cps), 20 cps e 28 cps. Cada participante assistiu a seis vídeos: três com imagens em movimento e legendas (um em cada velocidade) e três com legendas em uma tela preta (um em cada velocidade). Os movimentos oculares foram monitorados e a compreensão foi avaliada por meio de oito perguntas de múltipla escolha para cada vídeo (para uma metodologia detalhada, consulte o artigo original).

Os autores descobriram que, em velocidades mais altas, especialmente na de 28 cps, houve um efeito prejudicial sobre a compreensão. No entanto, a presença de vídeo auxiliou a compreensão, independentemente da velocidade das legendas. Em termos de movimentos oculares, o aumento da velocidade das legendas resultou em fixações menos frequentes e mais curtas, sacadas mais longas e menos cruzamentos entre legendas e imagens, o que levou a uma leitura mais superficial. Além disso, o estudo observou que efeitos estabelecidos na leitura, como o comprimento das palavras e a frequência (em que palavras mais longas e menos constantes são mais propensas a serem fixadas com mais frequência e por mais tempo), embora ainda presentes, foram modulados pela velocidade das legendas e pela presença de vídeo simultâneo. Da mesma forma, o efeito de *wrap-up* (finalização), que envolve gastar mais tempo nas palavras finais de uma sentença ou cláusula para integrar seu significado (Just; Carpenter, 1980; Warren *et al*., 2009) – uma tendência tradicionalmente interpretada em estudos de leitura como refletindo processamento de nível superior e integração lexical (Rayner *et al*., 2000) – foi influenciado pela velocidade das legendas e pela presença de vídeo.

Liao *et al*. (2021) descobriram que, nas velocidades mais altas e com vídeo, os espectadores gastavam menos tempo nas palavras ao final das legendas do que nas do meio – um efeito que os autores chamaram de "efeito de *wrap-up* reverso", também resultante de mais palavras ao final das legendas serem puladas nas velocidades mais altas. Isso sugere que os movimentos oculares são afetados tanto por dificuldades de processamento local quanto por restrições globais decorrentes da redução do tempo de leitura e da presença de uma tarefa visual secundária (ou seja, imagens em movimento) durante a visualização.

Na velocidade média de 20 cps, Liao *et al*. (2021) descobriram que o tempo gasto fixando palavras individuais nas legendas diminuiu, enquanto o comprimento médio das sacadas permaneceu inalterado. No entanto, a 28 cps, o tempo de fixação nas palavras diminuiu ainda mais, o que implica um impacto na pré-visualização parafoveal, enquanto o comprimento das sacadas e o número de palavras puladas aumentaram. Isso sugeriu uma mudança de decisões locais, baseadas nas palavras, sobre o ato de pular palavras para heurísticas mais globais, como pular palavras curtas, mas fixar palavras mais longas à medida que a velocidade das legendas ultrapassava 20 cps.

Outras descobertas do experimento original descrito por Liao *et al*. (2021) são relatadas por Kruger *et al.* (2022). Eles observaram que, à medida que a velocidade das legendas aumentava, mais palavras eram puladas ou não lidas. Menos palavras foram fixadas mais de uma vez, e houve um aumento no número de legendas não lidas até o fim. Em conjunto, essas descobertas sugerem que a velocidade excessiva pode não permitir tempo suficiente para resolver ambiguidades ou lidar com desafios de processamento.

Os resultados acima foram utilizados pelos autores para formular uma estrutura de controle de movimentos oculares na leitura multimodal, denominada “estrutura integrada de linguagem multimodal”. Essa estrutura, baseada no modelo computacional de leitura de Reichle (2020), acomoda entradas visuais e auditivas não textuais. A estrutura postula que os espectadores podem rastrear um número limitado de objetos previamente identificados no vídeo enquanto leem legendas em uma fase de processamento visual pré-atencional. Isso permite que esses objetos sejam vinculados ao conteúdo proposicional armazenado na memória de trabalho, derivado das legendas, formando, assim, o modelo de situação em evolução. Em essência, a estrutura sugere que a limitação inerente da atenção, que requer alocação serial, pode ser contornada ao monitorar simultaneamente objetos previamente identificados enquanto se lê as legendas. Esse componente da estrutura, argumentam os autores, é compatível com a teoria da codificação dual de Paivio (1971), que postula que a apresentação de informações em mais de uma modalidade beneficia a compreensão e o aprendizado. Consequentemente, as informações das legendas e o conteúdo do vídeo se complementam, o que explica uma melhor compreensão, mesmo na velocidade de legenda mais rápida, quando o vídeo está presente. O presente estudo de replicação busca validar ainda mais essa estrutura, oferecendo evidências adicionais para avaliar sua aplicabilidade geral.

3. Método

Os mesmos materiais audiovisuais e as legendas do estudo original foram utilizados. A menos que seja indicado o contrário, os mesmos procedimentos experimentais e protocolos de análise de dados foram seguidos (ver Liao *et al*., 2021, para mais detalhes).

3.1 Design

O estudo apresenta, para cada participante, um design de 2 x 3, em que 2 = vídeo: presente ou ausente e 3 = velocidade das legendas: 12, 20 e 28 cps. As condições de vídeo e de velocidade foram compensadas, e a ordem da apresentação de vídeos foi randomizada.

3.2 Participantes

Os participantes foram 51 falantes nativos de inglês do Reino Unido (sexo feminino = 33, sexo masculino = 13, não binário = 4; média de idade = 21,49, DP = 2,84, faixa etária = 18–33) e 42 falantes nativos de polonês (sexo feminino = 34, sexo masculino = 8; média de idade = 23,12, DP = 3,52, faixa etária = 18–33). A proficiência média em inglês dos participantes poloneses, medida pelo teste on-line LexTALE (Lemhöfer; Broersma, 2012), foi de 79,08 (DP = 10,38), variando de 62,5 a 100, o que corresponde aproximadamente aos níveis C1 e C2 do CEFR (Conselho da Europa, 2018).

Quando questionados sobre a frequência com que assistiam a filmes em inglês com legendas em inglês, os participantes australianos do estudo original relataram uma média de 2,78 (DP = 1,96) em uma escala de 1 a 7, em que 1 representava a menor familiaridade e 7, a maior (Kruger *et al*., 2022, p. 221). Esses números indicam que os participantes tiveram exposição limitada às legendas. Por outro lado, no nosso estudo, a legendagem foi o tipo de tradução audiovisual preferido pelos participantes do Reino Unido e da Polônia (90,2% para os do Reino Unido e 92,9% para os da Polônia). Além disso, ao serem questionados sobre assistir a conteúdos em inglês com legendas em inglês, eles relataram médias de 5,75 (DP = 1,65) e 5,86 (DP = 1,64), respectivamente, o que demonstra um alto nível de familiaridade com a legendagem.

3.3 Estímulos

Foram utilizados seis vídeos de aproximadamente nove minutos cada, retirados do documentário *Planeta Terra* da BBC. Os vídeos tinham duração semelhante (9–10 minutos), número de legendas (80–84), número de palavras (607–681) e escores de legibilidade. Os escores de legibilidade foram calculados com o índice Flesch Reading Ease (Graesser *et al*., 2014), na ferramenta *Coh-Metrix* (McNamara *et al*., 2010). A trilha sonora foi removida de todos os clipes de vídeo para eliminar a influência do áudio no processamento e permitir a manipulação consistente da velocidade das legendas, o que, de outra forma, teria causado assincronias com o som, introduzindo vieses indesejados.

3.4 Equipamento

Os movimentos oculares foram coletados com um EyeLink Portable Duo no Reino Unido e com um *EyeLink 1000 Plus Eye Tracker* (SR-Research, Ontario, Canadá) na Polônia. Os dados foram registrados de forma binocular a uma taxa de amostragem de 1000 Hz, utilizando um adesivo de alvo para rastreamento remoto com o *Portable Duo* e um apoio de queixo para minimizar o movimento da cabeça com o *1000 Plus*.

3.5 Procedimento

Ao chegarem, os participantes do Reino Unido e da Polônia assinaram formulários de consentimento, preencheram um questionário sobre suas preferências e hábitos relacionados à legendagem[1] e realizaram o *Teste de Proficiência em Inglês LexTALE* (Lemhöfer; Broersma, 2012).

Após a calibração (9 pontos com o *EyeLink 1000 Plus* e 13 pontos com o *EyeLink Portable Duo*, conforme as recomendações do fabricante), os participantes assistiram aos vídeos. Ao final de cada vídeo, eles responderam às mesmas perguntas de compreensão do experimento original. O experimento durou aproximadamente duas horas, incluindo dois intervalos de 5 minutos após o segundo e o quarto vídeo. Antes de cada vídeo, a calibração foi verificada e o rastreador ocular foi recalibrado, quando necessário.

3.6 Análise

Seguindo o estudo original, foram realizadas análises em nível de legendas e de palavras. As análises no nível das legendas concentraram-se em duas áreas de interesse

macro: a região das legendas e a região do vídeo. As medidas de rastreamento ocular examinadas foram a duração média de fixação (DMF), o número total de fixações, o comprimento médio das sacadas e o número de sacadas de cruzamento entre legendas e vídeo. No nível das palavras, examinamos a duração do olhar e o tempo total de fixação para avaliar o impacto da frequência e do comprimento das palavras. A duração do olhar é a soma de todas as fixações em uma área de interesse durante a primeira leitura, refletindo o processamento lexical inicial, enquanto o tempo total reflete processos relativamente tardios (como a integração lexical), pois inclui fixações e regressões em todas as passagens de leitura. Para avaliar os efeitos de *wrap-up*, analisamos o comportamento ocular nas palavras finais das legendas, observando o tempo total e a probabilidade de pular palavras, que se refere à fixação ou não de uma palavra durante a primeira leitura.

Para possibilitar uma comparação direta com o estudo original, fixações menores que 60 ms ou maiores que 800 ms foram excluídas. Oito participantes poloneses e quatorze britânicos foram excluídos devido à qualidade dos dados (ver detalhes no Pacote de Replicação), resultando em dados de rastreamento ocular analisáveis de 34 participantes da Polônia e 37 do Reino Unido.

Os dados foram analisados por meio de modelos lineares de efeitos mistos (generalizados) no R (versão 4.1.2). Nas análises globais (nível de legendas), a velocidade das legendas e a condição de vídeo foram inseridas nos modelos como efeitos fixos, juntamente com suas interações. Nas análises no nível das palavras (cujos resultados estão disponíveis como material suplementar devido à limitação de espaço), a frequência e o comprimento das palavras e as interações a elas relacionadas também foram adicionados como preditores. Nas análises de *wrap-up*, a localização das palavras também foi considerada para avaliar o comportamento nas palavras finais das legendas. Efeitos aleatórios foram adicionados a ambos os participantes e aos itens; os itens são as legendas completas ou as palavras individuais. Estruturas máximas foram ajustadas inicialmente e refinadas de acordo com a abordagem parcimoniosa de modelos mistos (Bates *et al*., 2015). A codificação de contraste foi aplicada (função contr.sdif no R), variáveis foram transformadas em log quando necessário, e o pacote emmeans foi utilizado para comparar pares de médias após o ajuste dos modelos. Por fim, o nível de significância foi ajustado para p = 0,0125 (0,05/4), com a correção de Bonferroni para múltiplas comparações, a fim de evitar a inflação das taxas de falso positivo (von der Malsburg; Angele, 2017).

4. Resultados

4.1 Compreensão

No geral, a precisão de compreensão foi de 74% entre os participantes do Reino Unido e 73% entre os participantes poloneses (Tabela 1), um pouco superior ao estudo original (70%). A compreensão foi menor quando os participantes assistiram às legendas sem o vídeo do que com o vídeo em todas as velocidades, confirmando o impacto positivo das imagens em movimento na precisão da compreensão. A adição de vídeo teve um efeito positivo significativo na precisão da compreensão (Polônia: $z = 4{,}081$, $p < 0{,}001$; Reino Unido: $z = 4{,}276$, $p < 0{,}001$). Diferentemente do estudo original e contrário às nossas expectativas, a velocidade das legendas não teve impacto significativo na precisão da compreensão, tanto no grupo do Reino Unido quanto no grupo da Polônia (Figura 1). Ao contrário da análise original, nenhuma interação entre as condições de velocidade e vídeo foi encontrada.

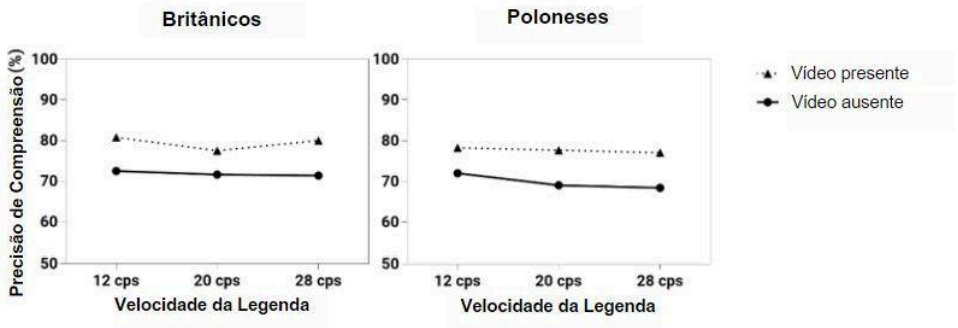
Tabela 1

Precisão de Compreensão por Velocidade e Condição de Vídeo

	12 cps		20 cps		28 cps	
	Vídeo presente	Vídeo ausente	Vídeo presente	Vídeo ausente	Vídeo presente	Vídeo ausente
Reino Unido	79.94 (40.09)	70.57 (45.63)	75.52 (43.05)	69.79 (45.97)	77.86 (41.57)	70.31 (45.74)
Polônia	78.27 (41.29)	72.02 (44.95)	77.67 (41.70)	69.04 (46.29)	77.08 (42.09)	68.45 (46.53)

Figura 1

Precisão de Compreensão por Velocidade de Legendas e Vídeo



4.2 Análises de rastreamento ocular no nível da legenda

Para possibilitar comparações com o estudo original, excluímos das análises de rastreamento ocular três participantes do Reino Unido e quatro da Polônia, cujas pontuações de compreensão em pelo menos dois vídeos foram inferiores a 40%. As Figuras 2–3 e as Tabelas 2–5 apresentam estatísticas descritivas e gráficos para análises de movimentos oculares nas regiões de legenda e de vídeo, bem como resumos dos modelos.

Tabela 2

Estatísticas Descritivas para Medidas de Movimentos Oculares no Nível de Legenda (Participantes do Reino Unido)

		Durações Médias de Fixação (ms)		Comprimento Médio de Sacadas (graus)		Número Total de Fixações		Número de Cruzamentos
Presença de Vídeo	Velocidade da Legenda	Região da Legenda (SD)	Região do Vídeo (SD)	Região da Legenda (SD)	Região do Vídeo (SD)	Região da Legenda (SD)	Região do Vídeo (SD)	Legenda/Vídeo (SD)
Ausente	12 cps	253 (57)	291 (160)	3.62 (2.60)	3.73 (2.71)	10.10 (3.68)	0.68 (1.53)	1.67 (1.16)
Ausente	20 cps	238 (62)	272 (150)	3.60 (2.24)	3.58 (2.55)	7.15 (2.41)	0.72 (1.31)	1.43 (0.84)
Ausente	28 cps	227 (58)	269 (153)	3.79 (2.29)	3.97 (2.54)	5.69 (1.92)	0.59 (1.14)	1.33 (0.70)
Presente	12 cps	204 (48)	334 (113)	3.73 (2.26)	4.11 (3.04)	7.97 (3.13)	3.55 (2.27)	2.21 (1.11)
Presente	20 cps	200 (49)	323 (135)	3.75 (2.16)	4.34 (3.09)	6.19 (2.34)	2.06 (1.44)	1.61 (0.72)
Presente	28 cps	195 (46)	307 (142)	3.81 (2.07)	4.82 (3.28)	5.50 (1.72)	1.44 (1.51)	1.30 (0.56)

Tabela 3

Estatísticas Descritivas para Medidas de Movimentos Oculares no Nível de Legenda (Participantes Poloneses)

		Durações Médias de Fixação (ms)		Comprimento Médio de Sacadas (graus)		Número Total de Fixações		Número de Cruzamentos
Presença de Vídeo	Velocidade da Legenda	Região da Legenda (SD)	Região do Vídeo (SD)	Região da Legenda (SD)	Região do Vídeo (SD)	Região da Legenda (SD)	Região do Vídeo (SD)	Legenda/Vídeo (SD)
Ausente	12 cps	234 (44)	259 (153)	3.09 (2.40)	2.82 (2.43)	11.40 (3.34)	0.64 (1.75)	1.32 (0.69)
Ausente	20 cps	219 (42)	234 (151)	2.97 (1.91)	3.37 (2.55)	8.14 (2.15)	0.31 (0.69)	1.15 (0.46)
Ausente	28 cps	214 (47)	250 (174)	3.18 (1.84)	3.32 (2.54)	6.29 (1.65)	0.33 (0.75)	1.12 (0.40)
Presente	12 cps	202 (39)	333 (122)	2.90 (1.94)	3.94 (2.83)	9.66 (3.23)	2.82 (1.97)	2.00 (1.00)
Presente	20 cps	196 (40)	310 (136)	2.85 (1.62)	4.59 (3.15)	7.40 (2.10)	1.67 (1.24)	1.42 (0.62)
Presente	28 cps	192 (41)	289 (141)	3.05 (1.63)	4.89 (3.27)	5.82 (1.56)	1.23 (0.93)	1.17 (0.43)

Figura 2

Comportamento Geral de Visualização dos Participantes do Reino Unido em Função da Presença/Ausência de Vídeo e Velocidade da Legenda

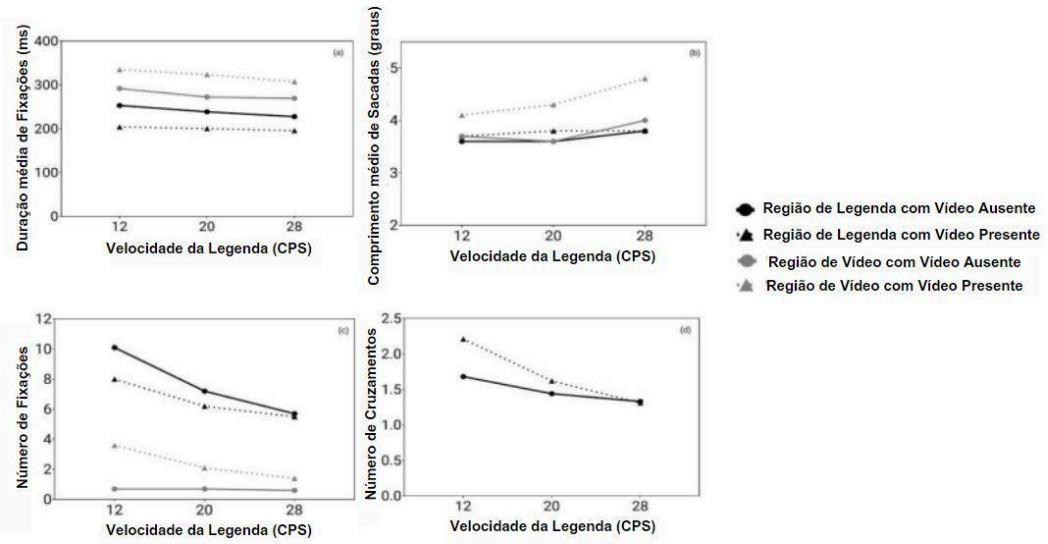


Figura 3

Comportamento Geral de Visualização dos Participantes Poloneses em Função da Presença/Ausência de Vídeo e Velocidade da Legenda

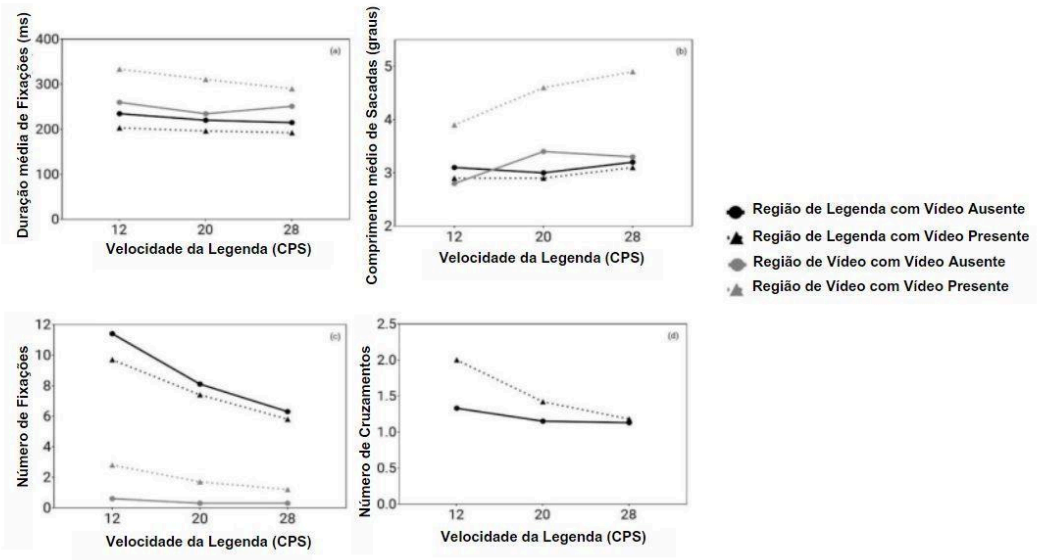


Tabela 4

Resultados dos Modelos Lineares Mistos (LMMs) para Análises no Nível de Legenda (Reino Unido)

		Região da Legenda				Região do Vídeo			
Medidas	**Contrastes**	**β**	**SE**	**t**	**p**	**β**	**SE**	**t**	**p**
Duração Média de Fixação	interseção	5.361065	0.016471	325.488	**<.0001*** **	5.58495	0.01453	384.467	**<.0001*** **
	Vídeo (Presente-Ausente)	-0.174216	0.010005	-17.413	**<.0001*** **	0.21248	0.02187	9.713	**<.0001*** **
	Velocidade(20-12cps)	-0.043129	0.005891	-7.322	**<.0001*** **	-0.05624	0.01431	-3.930	**<.0001*** **
	Velocidade(20-28cps)	-0.035041	0.005698	-6.150	**<.0001*** **	-0.05342	0.01707	-3.130	**0.003** *
	Vídeo x Velocidade(20-12cps)	0.035278	0.015595	2.262	0.030	-0.01202	0.02538	-0.474	0.635
	Vídeo x Velocidade(28-20cps)	0.038175	0.015262	2.501	0.017	-0.03592	0.02535	-1.417	0.156
Comprimento Médio de Sacadas	interseção	1.156e+00	2.102e-02	54.998	**<.0001*** **	1.135e+00	2.846e-02	39.873	**<.0001*** **
	Vídeo (Presente-Ausente)	5.763e-02	1.260e-02	4.573	**<.0001*** **	1.985e-01	4.471e-02	4.439	**<.0001*** **
	Velocidade(20-12cps)	2.394e-02	9.483e-03	2.524	0.016	4.262e-02	2.273e-02	1.875	0.060
	Velocidade(20-28cps)	4.109e-02	9.607e-03	4.277	**<.0002*** **	6.587e-02	2.409e-02	2.735	**0.006** *
	Vídeo x Velocidade(20-12cps)	-1.114e-02	8.865e-03	-1.256	0.209	9.089e-03	4.370e-02	0.208	0.835
	Vídeo x Velocidade(28-20cps)	-3.615e-02	1.007e-02	-3.591	**<.0004*** **	5.632e-02	4.609e-02	1.222	0.221
Número Total de Fixações	interseção	7.089e+00	1.754e-01	40.412	**<.0001*** **	1.530e+00	8.420e-02	18.166	**<.0001*** **
	Vídeo (Presente-Ausente)	-1.079e+00	1.814e-01	-5.948	**<.0001*** **	1.637e+00	1.535e-01	10.667	**<.0001*** **
	Velocidade(20-12cps)	-2.337e+00	1.663e-01	-14.053	**<.0001*** **	-7.475e-01	1.104e-01	-6.773	**<.0001*** **
	Velocidade(20-28cps)	-1.135e+00	9.650e-02	-11.762	**<.0001*** **	-3.477e-01	5.907e-02	-5.887	**<.0001*** **
	Vídeo x Velocidade(20-12cps)	1.200e+00	7.959e-02	15.082	**<.0001*** **	-1.476e+00	5.484e-02	-26.914	**<.0001*** **
	Vídeo x Velocidade(28-20cps)	7.793e-01	8.078e-02	9.648	**<.0001*** **	-5.993e-01	5.567e-02	-10.766	**<.0001*** **
Número de Cruzamentos	interseção	1.53657	0.04338	35.418	**<.0001*** **				
	Vídeo (Presente-Ausente)	0.32568	0.06404	5.085	**<.0001*** **				
	Velocidade(20-12cps)	-0.38002	0.04694	-8.096	**<.0001*** **				
	Velocidade(20-28cps)	-0.17810	0.03076	-5.791	**<.0001*** **				
	Vídeo x Velocidade(20-12cps)	-0.41939	0.04405	-9.520	**<.0001*** **				
	Vídeo x Velocidade(28-20cps)	-0.25939	0.04155	-6.244	**<.0001*** **				

Tabela 5

Resultados dos Modelos Lineares Mistos (LMMs) para Análises no Nível de Legenda (Polônia)

		Região da Legenda				Região do Vídeo			
Medidas	**Contrastes**	**β**	**SE**	**t**	**p**	**β**	**SE**	**t**	**p**
Duração Média de Fixação	interseção	5.326244	0.015318	347.708	**<.0001*** **	5.49576	0.02640	208.173	**<.0001*** **
	Vídeo (Presente-Ausente)	-0.123348	0.008063	-15.298	**<.0001*** **	0.30771	0.03286	9.363	**<.0001*** **
	Velocidade(20-12cps)	-0.049965	0.007325	-6.821	**<.0001*** **	-0.09994	0.01411	-7.083	**<.0001*** **
	Velocidade(20-28cps)	-0.025045	0.005254	-4.767	**<.0001*** **	-0.04541	0.01430	-3.175	0.001
	Vídeo x Velocidade(20-12cps)	0.029948	0.011830	2.531	0.016	0.01437	0.02824	0.509	0.610
	Vídeo x Velocidade(28-20cps)	0.006198	0.009168	0.676	0.503	-0.07951	0.02862	-2.778	0.005
Comprimento Médio de Sacadas	interseção	9.412e-01	2.067e-02	45.531	**<.0001*** **	1.077e+00	3.837e-02	28.080	**<.0001*** **
	Vídeo (Presente-Ausente)	-1.673e-02	8.555e-03	-1.956	0.059	3.130e-01	5.841e-02	5.359	**<.0001*** **
	Velocidade(20-12cps)	1.189e-02	8.775e-03	1.355	0.184	1.406e-01	3.091e-02	4.548	**<.0001*** **
	Velocidade(20-28cps)	7.407e-02	8.668e-03	8.546	**<.0001*** **	-1.607e-03	3.529e-02	-0.046	0.963
	Vídeo x Velocidade(20-12cps)	-1.135e-02	7.942e-03	-1.429	0.153	-8.996e-04	6.221e-02	-0.014	0.988
	Vídeo x Velocidade(28-20cps)	-6.483e-03	9.067e-03	-0.715	0.474	1.237e-01	7.073e-02	1.748	0.080
Número Total de Fixações	interseção	8.120e+00	1.544e-01	52.592	**<.0001*** **	1.171e+00	6.224e-02	18.816	**<.0001*** **
	Vídeo (Presente-Ausente)	-9.932e-01	1.098e-01	-9.050	**<.0001*** **	1.478e+00	1.015e-01	14.563	**<.0001*** **
	Velocidade(20-12cps)	-2.767e+00	1.450e-01	-19.085	**<.0001*** **	-7.385e-01	9.814e-02	-7.525	**<.0001*** **
	Velocidade(20-28cps)	-1.712e+00	1.097e-01	-15.602	**<.0001*** **	-2.098e-01	5.084e-02	-4.127	**<.0001*** **
	Vídeo x Velocidade(20-12cps)	1.025e+00	7.003e-02	14.641	**<.0001*** **	-8.213e-01	4.550e-02	-18.048	**<.0001*** **
	Vídeo x Velocidade(28-20cps)	2.150e-01	7.008e-02	3.069	**0.002** **	-4.526e-01	4.553e-02	-9.940	**<.0001*** **
Número de Cruzamentos	interseção	1.35181	0.02100	64.383	**<.0001*** **				
	Vídeo (Presente-Ausente)	0.34507	0.04151	8.314	**<.0001*** **				
	Velocidade(20-12cps)	-0.36372	0.03722	-9.771	**<.0001*** **				
	Velocidade(20-28cps)	-0.13051	0.02606	-5.007	**<.0001*** **				
	Vídeo x Velocidade(20-12cps)	-0.42612	0.04121	-10.341	**<.0001*** **				
	Vídeo x Velocidade(28-20cps)	-0.21926	0.04075	-5.381	**<.0001*** **				

4.2.1 Duração média de fixação (DMF)

Nas análises realizadas tanto para os participantes do Reino Unido quanto para os da Polônia, assim como no estudo original, a DMF foi maior na região do vídeo em comparação à da legenda. Identificamos um efeito principal tanto do vídeo quanto da velocidade na DMF (Tabela 4 e Tabela 5). Na região da legenda, a presença do vídeo resultou em uma redução significativa na DMF em comparação à condição sem vídeo. A DMF também diminuiu com o aumento da velocidade em ambas as condições de vídeo, em todas as três coortes. Não foram

encontradas interações entre velocidade e vídeo nas coortes do Reino Unido e da Polônia após a aplicação da correção de Bonferroni para múltiplas comparações. Isso contrasta com a coorte australiana, onde a interação foi significativa em ambos os intervalos de velocidade.

Os resultados na região do vídeo foram similares entre as três coortes: a adição do vídeo produziu um aumento significativo na DMF nessa região, e o aumento da velocidade levou a uma redução significativa na DMF. A única interação significativa foi registrada para a coorte polonesa nas velocidades mais altas, indicando que, para os espectadores poloneses, havia uma menor diferença na DMF na região do vídeo entre 20 e 28 cps quando o vídeo não estava em exibição.

4.2.2 Comprimento médio de sacadas

Assim como no estudo original, os participantes do Reino Unido e da Polônia realizaram sacadas mais longas ao observar a região do vídeo do que ao ler as legendas.

Nas análises da região de legendas, a principal diferença entre as três coortes foi o impacto do vídeo. Diferentemente da duração e do número de fixações, o comprimento das sacadas na coorte polonesa não foi significativamente afetado pela presença ou ausência de vídeo, assim como na coorte australiana. Por outro lado, na coorte do Reino Unido, o comprimento das sacadas aumentou nessa região.

Na região de legendas, a velocidade teve um efeito significativo no comprimento das sacadas apenas na configuração de csp mais alta, em todas as três coortes. Conforme a velocidade das legendas aumentava de 20 cps para 28 cps, o comprimento das sacadas também aumentava, pois os espectadores faziam sacadas mais longas para tentar ler as legendas que desapareciam rapidamente. Não houve interação significativa entre velocidade e vídeo na análise dessa região nas coortes australiana e polonesa. No entanto, na coorte do Reino Unido, observou-se uma interação significativa entre vídeo × velocidade na velocidade mais alta. Uma análise mais detalhada das médias indicou que, na coorte do Reino Unido, o efeito da velocidade sobre o comprimento das sacadas na região de legendas foi mais pronunciado na ausência de vídeo.

Na região do vídeo, a principal diferença entre as três coortes foi o efeito da velocidade. Enquanto a presença do vídeo afetou positivamente o comprimento médio das sacadas nessa região em todas as coortes, a velocidade das legendas levou a sacadas mais longas apenas na transição de 12 cps para 20 cps na coorte australiana e na coorte polonesa. Na coorte do Reino Unido, no entanto, o comprimento das sacadas na região do vídeo foi consideravelmente

maior apenas na transição de 20 cps para 28 cps. Não foi identificada nenhuma interação significativa entre velocidade e vídeo na análise dessa região em todas as coortes.

4.2.3 Número total de fixações

De modo geral, o comportamento ocular nas regiões de legendas e de vídeo foi muito semelhante entre as três coortes analisadas. Assim como os participantes australianos, no estudo original, os espectadores do Reino Unido e da Polônia realizaram mais fixações nas legendas do que no vídeo (Austrália: 7,4 vs. 1,3; Reino Unido: 7,1 vs. 1,54; Polônia: 8,12 vs. 1,17).

A presença de vídeo emergiu como um preditor significativo do número de fixações nas regiões de legendas e de vídeo, um padrão consistente entre os participantes do Reino Unido, Polônia e Austrália. Ao lidar com duas fontes de informações significativas (imagens e legendas), os espectadores realizavam menos fixações nas legendas na condição em que o vídeo estava presente, em comparação à condição sem vídeo.

Também foi identificado um efeito significativo da velocidade das legendas sobre o número de fixações na região de legendas: à medida que a velocidade aumentava, o número de fixações nas legendas diminuía, tanto na condição com vídeo quanto na condição sem vídeo. De forma semelhante, na região do vídeo, à medida que a velocidade das legendas aumentava, o número de fixações diminuía em ambas as condições de vídeo.

Em conformidade com as análises do estudo original, interações significativas entre as condições de vídeo e velocidade também foram observadas em ambas as condições de vídeo para os participantes do Reino Unido e da Polônia. Embora a velocidade das legendas tenha impactado o número de fixações em ambas as condições de vídeo, o efeito observado foi mais pronunciado na condição sem vídeo, que se assemelha mais à leitura de texto impresso.

4.2.4 Número de cruzamentos

Para todos os participantes, o número de sacadas de cruzamento entre as regiões de legendas e de vídeo foi significativamente maior quando o vídeo estava presente em comparação à condição sem vídeo, refletindo a saliência das imagens em movimento. Além disso, à medida que aumentava a velocidade das legendas, o número de cruzamentos diminuía significativamente. Em todas as três análises, foram encontradas interações significativas, embora houvesse algumas diferenças. Uma análise mais detalhada das médias do Reino

Unido e da Polônia revelou que o efeito da velocidade das legendas sobre o número de cruzamentos foi mais evidente na condição com vídeo, mas também se manifestou na condição sem vídeo em velocidades mais baixas (12 cps e 20 cps). No entanto, na análise original, o efeito significativo da velocidade das legendas sobre o número de cruzamentos foi observado apenas quando as imagens estavam sendo exibidas.

4.3. Análises de rastreamento ocular a nível de palavra

4.3.1 Efeitos da frequência e do comprimento das palavras: duração do olhar

Assim como no estudo original, as análises para a Polônia e o Reino Unido revelaram efeitos principais dos quatro fatores fixos (veja as Tabelas 6–9 nos anexos). Primeiro, a duração do olhar apresentava uma relação direta com o comprimento da palavra: à medida que o comprimento aumentava, a duração do olhar também aumentava. Segundo, a duração do olhar apresentou uma relação inversa com a velocidade das legendas e a frequência das palavras: à medida que aumentavam a velocidade e a frequência, a duração do olhar diminuía significativamente (Figuras 4 e 5). Por fim, a adição de vídeo resultou em uma diminuição considerável na duração do olhar sobre as palavras das legendas.

Assim como no estudo original, os resultados para o Reino Unido e para a Polônia foram modulados por relações. Identificou-se, por exemplo, uma relação entre comprimento e frequência das palavras em ambas as coortes com inglês nativo (Austrália e Reino Unido), em que palavras mais longas e de menor frequência atraíram uma duração de olhar maior, mas isso não foi observado na coorte polonesa após a correção de Bonferroni. Além disso, diferentemente da análise original, que identificou apenas uma relação entre três fatores (comprimento × frequência × vídeo), as análises para a Polônia e para o Reino Unido revelaram diferentes relações significativas. Uma relação comprimento × frequência × velocidade relevante foi encontrada apenas entre 12–20 cps (caracteres por segundo) nos dados da coorte do Reino Unido e da Polônia. Comparações pareadas mostraram que os efeitos de frequência e de comprimento foram particularmente acentuados na velocidade mais baixa (12 cps) e se tornaram atenuados à medida que a velocidade aumentava. Na coorte polonesa, também foi encontrada uma relação significativa entre comprimento × velocidade × vídeo, tanto entre 12–20 cps quanto entre 20–28 cps. Uma análise mais detalhada dos pares de médias revelou que, na velocidade mais alta, entre os participantes poloneses, os efeitos do

comprimento das palavras ainda estavam presentes (com os espectadores passando mais tempo lendo palavras mais longas), mas esses efeitos foram atenuados e menos influenciados pela presença do vídeo.

4.3.2 Efeitos da frequência e do comprimento das palavras: tempo total

De forma semelhante à duração do olhar, foram identificados efeitos principais em todos os quatro preditores. Primeiro, os tempos totais apresentaram uma relação direta com o comprimento das palavras: os espectadores passaram mais tempo fixando palavras mais longas. Segundo, os tempos totais foram inversamente proporcionais à velocidade e à frequência das palavras: à medida que aumentaram a velocidade e a frequência, os tempos totais diminuíram significativamente (veja as Figuras 4 e 5). Os tempos totais também diminuíram após a adição do vídeo.

A frequência das palavras estava relacionada à velocidade em todos os níveis na coorte da Polônia, mas não na coorte do Reino Unido. Em comparação, na coorte da Austrália, essa relação foi significativa apenas entre 12 e 20 cps, sugerindo que os efeitos da frequência das palavras podem ser mais perceptíveis para falantes não nativos, especialmente em velocidades mais altas.

Em todas as três coortes, também foram encontradas relações entre comprimento e vídeo, nas quais palavras mais longas levaram a tempos totais mais elevados quando o vídeo estava ausente. Uma relação comprimento × velocidade foi identificada nas três coortes entre 12–20 cps, indicando que os efeitos do aumento da velocidade nos tempos totais foram mais evidentes em palavras mais longas. Além disso, uma relação entre comprimento e frequência foi observada nas três coortes, com efeitos de frequência mais acentuados em palavras mais longas. Por fim, uma relação entre vídeo e velocidade também se manifestou nas três análises. A adição de vídeo resultou em uma diminuição constante dos tempos totais gastos lendo palavras em todas as três velocidades nas coortes nativas, enquanto na coorte não nativa polonesa, a diminuição nos tempos totais foi mais pronunciada entre 12 e 20 cps.

Figura 4

Efeitos de Comprimento de Palavra e Frequência de Palavra Ajustados por LMM para Participantes do Reino Unido em Função da Presença/Ausência de Vídeo e Velocidade da Legenda, com Durações de Olhar no Painel (A) e Tempos Totais no Painel (B).
A Frequência de Palavra é Baseada na Escala de Zipf do Corpus SUBTLEX-UK.

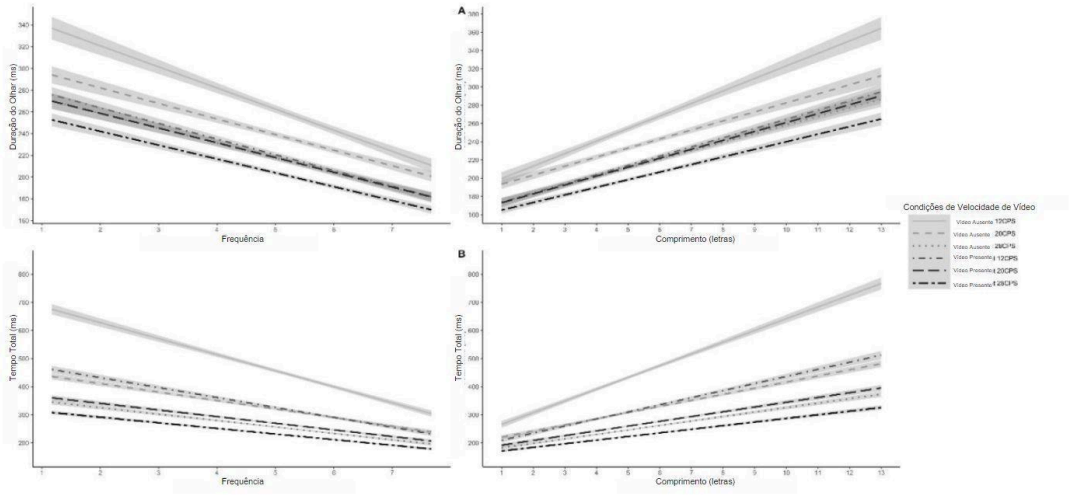
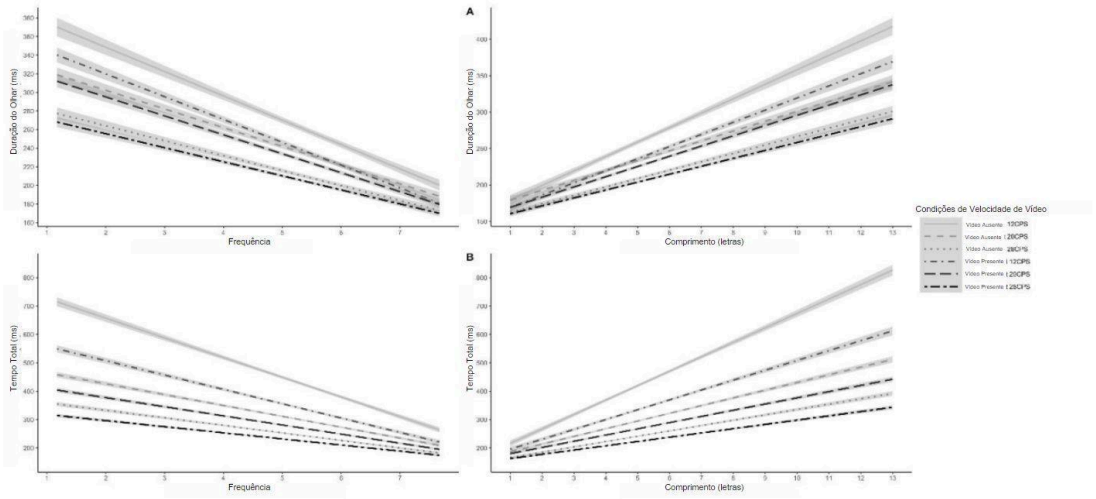


Figura 5

Efeitos de Comprimento de Palavra e Frequência de Palavra Ajustados por LMM para Participantes Poloneses em Função da Presença/Ausência de Vídeo e Velocidade da Legenda, com Durações de Olhar no Painel (A) e Tempos Totais no Painel (B).
A Frequência de Palavra é Baseada na Escala de Zipf do Corpus SUBTLEX-UK.



4.3.3 Efeitos de *wrap-up*: tempos totais

Para avaliar os efeitos de wrap-up, comparamos os tempos totais de fixação em palavras na posição final das legendas com os da posição intermediária. Similar a análises anteriores,

velocidade e vídeo tiveram efeitos principais significativos em todas as coortes. Nas análises para a Austrália e o Reino Unido, nenhum efeito principal significativo foi observado na localização das palavras. No entanto, na análise para a Polônia, foi identificado um efeito principal de localização, indicando que, diferentemente das duas coortes nativas, a polonesa exibiu uma redução significativa nos tempos de leitura de palavras no final das legendas.

Uma relação significativa entre vídeo e velocidade foi observada em todas as análises para as três coortes em todos os intervalos de velocidade. Uma relação entre localização e velocidade foi significativa entre 12–20 cps na coorte do Reino Unido, entre 20–28 cps na coorte da Polônia e em todas as velocidades na análise original da coorte da Austrália. Além disso, uma relação entre localização, velocidade e vídeo foi significativa entre 12–20 cps na análise do Reino Unido, entre 20–28 cps na análise da Polônia e em todas as velocidades na análise original. Comparações pareadas adicionais apontaram semelhanças entre todas as coortes: sem vídeo, observou-se um efeito de wrap-up "padrão" a 12 cps, mas não nas duas velocidades mais altas. Com vídeo, nenhum efeito de *wrap-up* foi identificado, pois os tempos totais para palavras no final das legendas se tornaram mais curtos, possivelmente porque as palavras desapareciam rápido demais para serem lidas.

4.3.4 Efeitos de *wrap-up*: probabilidade de pular palavras

No modelo do Reino Unido, o vídeo foi um preditor significativo para as taxas do fenômeno de pular palavras: quando o vídeo estava presente, o pulo de palavras aumentava significativamente, sugerindo que os espectadores do Reino Unido – assim como os da Austrália – dedicaram mais tempo às imagens quando essas estavam presentes, resultando em mais palavras puladas nas legendas. Por outro lado, na análise da Polônia, o vídeo não teve efeito principal significativo nas taxas de pular palavras; ou seja, independentemente da presença do vídeo, os espectadores tendem a pular palavras em proporções semelhantes. A velocidade das legendas teve efeitos mais claros em todas as coortes: à medida que a velocidade aumentava, as taxas de pular palavras também aumentavam significativamente.

Nesta análise, não houve efeitos principais da localização isoladamente nas taxas de pular palavras em nenhuma das coortes. Inicialmente, a localização apresentou um efeito na coorte polonesa – refletindo a análise dos tempos totais de wrap-up acima –, mas esse efeito tornou-se não significativo após a correção de Bonferroni.

Em todas as análises, foram encontradas relações significativas de localização × vídeo e de localização × velocidade. Além disso, apenas na coorte do Reino Unido, a relação entre

três fatores — localização × velocidade × vídeo — foi significativa em todas as velocidades. Comparações pareadas da relação entre localização e vídeo mostraram que palavras no final das legendas foram puladas com mais frequência quando o vídeo estava presente em todas as três coortes, sugerindo que os espectadores tiveram menos tempo para ler as legendas até o final enquanto acompanhavam as imagens em movimento. Em contraste, as palavras em posição intermediária foram puladas com menos frequência na presença de vídeo.

Comparações pareadas das relações entre localização e velocidade revelaram que, em todas as coortes, as palavras no final das legendas apresentaram efeitos mais claros relacionados à velocidade e eram puladas cada vez mais à medida que a velocidade aumentava.

5. Discussão geral

O objetivo deste estudo foi replicar os resultados de Liao *et al*. (2021) sobre o impacto da velocidade das legendas e da presença ou ausência de imagens em movimento no comportamento de leitura multimodal. Além disso, buscamos confirmar se esses resultados, originalmente observados em falantes nativos de inglês australiano, também se aplicam a falantes nativos de inglês no Reino Unido e a falantes poloneses que leem legendas em inglês como segunda língua (L2). De modo geral, os resultados do estudo original foram replicados com sucesso, confirmando que os efeitos da presença de vídeo e das variações na velocidade das legendas são robustos e consistentes tanto em populações nativas quanto em não nativas.

5.1 Compreensão

Em relação ao estudo original, esses resultados mostram que a inclusão de vídeo melhora a compreensão ao fornecer um contexto visual valioso à narrativa. No entanto, os autores do estudo original consideraram surpreendente que a presença de vídeo tenha melhorado a compreensão, dado o princípio da redundância descrito na teoria da carga cognitiva (Diao & Sweller, 2007), segundo o qual a atenção a duas fontes de entrada simultâneas (vídeo e legendas) poderia prejudicar a compreensão. Argumentamos, por outro lado,que esse resultado não contradiz o princípio da redundância, porque o vídeo não pode ser considerado verdadeiramente redundante em relação às legendas. Diferentemente dos diálogos falados e das legendas correspondentes – que podem ser idênticos ou quase idênticos e, assim, redundantes – as imagens, frequentemente, não mostram de maneira exclusiva ou

exata o que é transmitido pelas legendas. As legendas, por sua vez, podem fornecer informações adicionais relacionadas ao que é mostrado na tela. Portanto, sugerimos que as imagens sejam vistas como complementares, e não redundantes. Assim, não é surpreendente que os participantes tenham demonstrado maior compreensão ao visualizar as imagens, já que estas ajudam a ampliar o entendimento, em vez de duplicar informações.

Curiosamente, ao contrário do estudo original, velocidades mais altas das legendas não resultaram em redução significativa da compreensão entre nossos participantes. Esse resultado pode ser atribuído ao fato de que, ao contrário dos participantes australianos do estudo original, tanto os espectadores poloneses quanto os do Reino Unido estão habituados à leitura de legendas e à integração de informações em situações de leitura multimodal. Sabe-se que a experiência prévia com legendas impacta o processo de leitura, com espectadores menos experientes enfrentando mais dificuldades do que aqueles familiarizados com a prática (Jensema, 1998; Szarkowska & Gerber-Morón, 2018). Além disso, considerando o papel complementar das imagens, é possível que nossos grupos de participantes tenham explorado o vídeo de forma mais eficaz para apoiar sua compreensão em velocidades mais altas de legendas, embora isso ainda precise ser testado empiricamente.

Ademais, reconhecemos limitações relacionadas à construção das perguntas de compreensão no estudo original. Essas questões combinavam elementos de compreensão e de memória, ocasionalmente exigindo que os participantes recordassem detalhes específicos, números ou a formulação exata das legendas. Isso significa que conclusões sobre o impacto da velocidade na compreensão não podem ser estabelecidas de forma confiável, pois houve confusão com elementos de memória.

5.2 O impacto da velocidade de legendas e a presença de vídeo nos movimentos oculares

Tanto a presença do vídeo quanto a velocidade das legendas influenciaram significativamente os padrões de visualização de todos os participantes. Nas análises no nível das legendas, os espectadores, de forma consistente, priorizaram a leitura das legendas em vez de assistir às imagens em movimento, evidenciado por um maior número de fixações nas legendas em comparação ao vídeo. Quando o vídeo estava presente, os participantes interagiam com as imagens em movimento, conforme indicado pelas durações de fixação média (DFM) e sacadas na região do vídeo mais longas . Nossos resultados estão alinhados com pesquisas anteriores sobre diferenças na percepção de cenas e na leitura, em que a leitura geralmente envolve fixações mais frequentes e curtas e sacadas mais breves (Rayner, 2009).

O número e a duração das fixações nas legendas diminuíram com a presença de vídeo e com o aumento da velocidade. A 28 cps, a leitura das legendas tornou-se mais uniforme, com os espectadores fazendo números semelhantes de fixações em ambas as condições de vídeo e realizando menos cruzamentos entre legendas e imagens devido ao tempo limitado. Nessa velocidade, os espectadores podem não ter tempo suficiente para realizar processos cognitivos vitais durante a leitura, como resolver ambiguidades linguísticas, confirmar significados e integrar palavras ao contexto prévio.

A velocidade mais alta resultou em sacadas mais longas na região das legendas, enquanto os espectadores tentavam ler as legendas antes que desaparecessem. Isso confirmou o efeito consistente da velocidade sobre os movimentos oculares. No entanto, diferentemente da duração e do número de fixações, o comprimento das sacadas na região das legendas foi influenciado pela presença do vídeo apenas na coorte do Reino Unido, mas não nas coortes polonesa e australiana. Isso pode ser atribuído à maior proficiência em leitura multimodal dos participantes do Reino Unido, que lhes permitiu ajustar seu comportamento visual de forma mais eficiente, fazendo sacadas mais longas para concluir a leitura das legendas rapidamente e processar as imagens. No caso dos participantes poloneses, igualmente versados em processamento multimodal, a ausência de sacadas mais longas pode ser explicada pelo fato de estarem lendo em língua estrangeira, o que exige um processamento mais detalhado das palavras das legendas, independentemente da presença do vídeo.

Nas análises no nível das palavras, confirmamos – em conformidade com o estudo original – a existência de dois efeitos bem estabelecidos na pesquisa sobre a leitura de legendas: o comprimento e a frequência das palavras (Clifton *et al*., 2016). O comprimento e a frequência das palavras são dois dos "Três Grandes" no processamento lexical (o terceiro é a previsibilidade, que não foi testada neste estudo). Esses efeitos foram mais pronunciados na leitura de legendas sem vídeo à velocidade mais baixa, semelhante à leitura de texto estático. No entanto, a adição de imagens em movimento e o aumento da velocidade das legendas reduziram esses efeitos, confirmando o papel da velocidade de legendas e da presença de vídeo como restrições abrangentes no comportamento de leitura.

5.3 Efeitos de *wrap-up*

Diferenças interessantes entre a leitura de texto estático e a de legendas surgiram quanto ao efeito de *wrap-up* (Rayner *et al*., 2000). No contexto das legendas, este efeito foi observado em tempos de leitura prolongados nas palavras finais, na velocidade mais baixa (12

cps), quando o vídeo estava ausente. No entanto, na velocidade mais alta, com o vídeo presente, todos os espectadores passaram menos tempo nas palavras finais das legendas do que nas do meio. Isso sugere que as legendas desapareceram rápido demais para serem completamente processadas, o que potencialmente interrompeu a integração em níveis superiores.

Diferenças no efeito de *wrap-up* foram particularmente evidentes na coorte polonesa, na qual a posição da palavra (final vs. meio) influenciou os tempos totais de leitura dessas legendas. Os participantes poloneses exibiram uma redução notável nos tempos totais de leitura para as palavras finais das legendas, mesmo considerando a velocidade das legendas e a presença do vídeo. Esses tempos de leitura refletem processos de integração pós-lexical, indicando que os espectadores não nativos podem enfrentar desafios para processar legendas de forma abrangente em comparação com os nativos. Essas dificuldades podem ser atribuídas a processos como a conversão mental de representações linguísticas de orações e sentenças em representações proposicionais, que podem não estar tão automatizados em falantes não nativos (Segalowitz & Hulstijn, 2005). Isso sugere que espectadores de inglês como segunda língua (L2) podem ler a um ritmo mais lento, precisando de mais tempo para compreender o significado das palavras e ajustar suas representações mentais antes de concluir a leitura das legendas.

Uma observação sobre o efeito de *wrap-up* nas legendas é relevante. Na leitura tradicional, o efeito de *wrap-up* normalmente ocorre no final de frases ou orações e é influenciado pela pontuação (Warren *et al*., 2009). No entanto, as legendas nem sempre se alinham aos limites de orações ou sentenças, podendo não formar unidades autônomas que terminem com pontuação. Mesmo quando uma legenda constitui uma sentença completa, o efeito de *wrap-up* pode ser influenciado pela presença de outra legenda imediatamente a seguir, o que potencialmente interrompe a integração. Também é possível que a integração pós-lexical continue após uma legenda desaparecer, especialmente se nenhuma outra legenda surgir imediatamente. Consequentemente, uma exploração mais aprofundada dos efeitos de *wrap-up* na leitura multimodal, considerando a sobreposição entre sentenças e legendas, é necessária para compreender como o processamento e a integração de sentenças em níveis superiores ocorrem à medida que as legendas aparecem e desaparecem.

6. Conclusões

Este estudo replicou os principais resultados de Liao *et al*. (2021), confirmando o impacto da velocidade das legendas e da presença de vídeo nos movimentos oculares de diferentes grupos de participantes, reafirmando a robustez desses efeitos tanto no nível das legendas quanto no nível das palavras, em populações nativas e não nativas. Nossos resultados também destacaram o impacto negativo de velocidades excessivas de legendas (acima de 20 cps): a 28 cps, os espectadores tiveram dificuldades para acompanhar as legendas que desapareciam rápido, frequentemente perdendo as palavras finais, o que compromete a experiência de assistir a conteúdos audiovisuais legendados.

Agradecimentos

Este estudo foi realizado no âmbito do projeto WATCH-ME, financiado pelo *National Science Centre* (Centro Nacional de Ciência da Polônia), Programa OPUS 19, 2020/37/B/HS2/00304.

Referências

BATES, D.; KLIEGL, R.; VASISHTH, S.; BAAYEN, H. **Parsimonious mixed models.** Disponível em: https://doi.org/10.48550/ARXIV.1506.04967. Acesso em: 20 dez. 2024.

COUNCIL OF EUROPE. **Council for Cultural Co-operation. Education Committee. Modern Languages Division**. Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Cambridge: Cambridge University Press, 2018.

CLIFTON, C. *et al*. **Eye movements in reading and information processing:** Keith Rayner's 40 year legacy. Journal of Memory and Language, v. 86, p. 1–19, 2016. Disponível em: https://doi.org/10.1016/j.jml.2015.07.004. Acesso em: 20 dez. 2024.

DIAO, Y.; SWELLER, J. **Redundancy in foreign language reading comprehension instruction:** Concurrent written and spoken presentations. Learning and Instruction, v. 17, n. 1, p. 78–88, 2007. Disponível em: https://doi.org/10.1016/j.learninstruc.2006.11.007. Acesso em: 20 dez. 2024.

DÍAZ CINTAS, J.; REMAEL, A. **Subtitling:** Concepts and practices. Nova York: Routledge, 2021.

d'YDEWALLE, G.; MUYLLE, P.; RENSBERGEN, J. v. **Attention shifts in partially redundant information situations.** In: GRONER, R.; McCONKIE, G. W.; MENZ, C. (Eds.). Eye movements and human information processing. p. 375–384. Elsevier, 1985.

D'YDEWALLE, G.; 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.; LEVY-SCHOEN, A. (Eds.). Eye movements: From physiology to cognition. Amsterdam: Elsevier, 1987. p. 313–321.

GRAESSER, A. C.; McNAMARA, D. S.; CAI, Z.; CONLEY, M.; LI, H.; PENNEBAKER, J. **Coh-metrix measures text characteristics at multiple levels of language and discourse.** The Elementary School Journal, v. 115, p. 210–228, 2014.

JENSEMA, C. **Viewer reaction to different television captioning speeds.** American Annals of the Deaf, v. 143, n. 4, p. 318–324, 1998.

JUST, M. A.; CARPENTER, P. A. **A theory of reading:** From eye fixations to comprehension. Psychological Review, v. 87, n. 4, p. 329–354, 1980.

KOOLE, S. L.; LAKENS, D. **Rewarding replications:** A sure and simple way to improve psychological science. Perspectives on Psychological Science, v. 7, n. 6, p. 608–614, 2012. Disponível em: https://doi.org/10.1177/1745691612462586. Acesso em: 20 dez. 2024.

KOOLSTRA, C. M.; VAN DER VOORT, T. H. A.; D'YDEWALLE, G. **Lengthening the presentation time of subtitles on television:** Effects on children's reading time and recognition. Communications, v. 24, n. 4, p. 407–422, 1999. Disponível em: https://doi.org/10.1515/comm.1999.24.4.407. Acesso em: 20 dez. 2024.

KRUGER, J.-L.; WISNIEWSKA, N.; LIAO, S. **Why subtitle speed matters:** Evidence from word skipping and rereading. Applied Psycholinguistics, v. 43, n. 1, p. 211–236, 2022. Disponível em: https://doi.org/10.1017/S0142716421000503. Acesso em: 20 dez. 2024.

LEBEL, E. P. *et al*. **A unified framework to quantify the credibility of scientific findings.** Advances in Methods and Practices in Psychological Science, v. 1, n. 3, p. 389–402, 2018. Disponível em: https://doi.org/10.1177/2515245918787489. Acesso em: 20 dez. 2024.

LEMHÖFER, K.; BROERSMA, M. Introducing LexTALE: **A quick and valid lexical test for advanced learners of English.** Behavior Research Methods, v. 44, n. 2, p. 325–343, 2012. Disponível em: https://doi.org/10.3758/s13428-011-0146-0. Acesso em: 20 dez. 2024.

LIAO, S. *et al*. **Using eye movements to study the reading of subtitles in video.** Scientific Studies of Reading, v. 25, n. 5, p. 417–435, 2021. Disponível em: https://doi.org/10.1080/10888438.2020.1823986. Acesso em: 20 dez. 2024.

MCNAMARA, D. S.; LOUWERSE, M. M.; McCARTHY, P. M.; GRAESSER, A. C. **Coh-metrix:** Capturing linguistic features of cohesion. Discourse Processes: A Multidisciplinary Journal, v. 47, n. 4, p. 292–330, 2010.

OPEN SCIENCE COLLABORATION. **Estimating the reproducibility of psychological science.** Science, v. 349, n. 6251, 2015. Disponível em: https://doi.org/10.1126/science.aac4716. Acesso em: 20 dez. 2024.

PAIVIO, A. **Imagery and verbal processes.** Nova York: Holt, Rinehart & Winston, 1971.

RAYNER, K. **Eye movements and attention in reading, scene perception, and visual search.** The Quarterly Journal of Experimental Psychology, v. 62, n. 8, p. 1457–1506, 2009. Disponível em: https://doi.org/10.1080/17470210902816461. Acesso em: 20 dez. 2024.

RAYNER, K.; KAMBE, G.; DUFFY, S. A. **The effect of clause wrap-up on eye movements during reading.** The Quarterly Journal of Experimental Psychology: Section A, v. 53, n. 4, p. 1061–1080, 2000. Disponível em: https://doi.org/10.1080/713755934. Acesso em: 20 dez. 2024.

REICHLE, E. D. **Computational models of reading:** A handbook. Oxford University Press, 2020.

SCHMIDT, S. **Shall we really do it again? The powerful concept of replication is neglected in the social sciences.** Review of General Psychology, v. 13, n. 2, p. 90–100, 2009. Disponível em: https://doi.org/10.1037/a0015108. Acesso em: 20 dez. 2024.

SEGALOWITZ, N.; HULSTIJN, J. **Automaticity in bilingualism and second language learning.** In: KROLL, J. F.; DE GROOT, A. M. B. (Eds.). Handbook of bilingualism: Psycholinguistics approaches. Oxford: Oxford University Press, 2005. p. 371–388.

SZARKOWSKA, A.; GERBER-MORÓN, O. **Viewers can keep up with fast subtitles:** Evidence from eye movements. PLoS ONE, v. 13, n. 6, 2018. Disponível em: https://doi.org/10.1371/journal.pone.0199331. Acesso em: 20 dez. 2024.

SZARKOWSKA, A.; SILVA, B.; ORREGO-CARMONA, D. **Effects of subtitle speed on proportional reading time.** Translation, Cognition & Behavior, v. 4, n. 2, p. 305–330, 2021. Disponível em: https://doi.org/10.1075/tcb.00057.sza. Acesso em: 20 dez. 2024.

VANDERPLANK, R. **Captioned media in foreign language learning and teaching:** Subtitles for the deaf and hard-of-hearing as tools for language learning. Nova York: Palgrave Macmillan, 2016.

VON DER MALSBURG, T.; ANGELE, B. **False positives and other statistical errors in standard analyses of eye movements in reading.** Journal of Memory and Language, v. 94, p. 119–133, 2017. Disponível em: http://dx.doi.org/10.1016/j.jml.2016.10.003. Acesso em: 20 dez. 2024.

WANG, A.; PELLICER‐SÁNCHEZ, A. **Incidental vocabulary learning from bilingual subtitled viewing:** An eye‐tracking study. Language Learning, v. 72, n. 3, p. 765–805, 2022. Disponível em: https://doi.org/10.1111/lang.12495. Acesso em: 20 dez. 2024.

WARREN, T.; WHITE, S. J.; REICHLE, E. D. **Investigating the causes of wrap-up effects:** Evidence from eye movements and E-Z Reader. Cognition, v. 111, n. 1, p. 132–137, 2009. Disponível em: https://doi.org/10.1016/j.cognition.2008.12.011. Acesso em: 20 dez. 2024.

ZAJECHOWSKI, M. **Survey:** Why America is obsessed with subtitles. 2022.

Recebido em: 15/06/2025
Aprovado em: 01/12/2025