

# Investigating the Influence of Subtitles Synchronization in the Viewer's Quality of Experience

Rodrigo Laiola Guimarães

Universidade Federal do Espírito Santo  
Av. Fernando Ferrari, S/N  
29075-910 Vitória - ES, Brazil  
rlaiola@inf.ufes.br

Jessica Oliveira Brito

Universidade Federal do Espírito Santo  
Av. Fernando Ferrari, S/N  
29075-910 Vitória - ES, Brazil  
jeubrito@gmail.com

Celso A. S. Santos

Universidade Federal do Espírito Santo  
Av. Fernando Ferrari, S/N  
29075-910 Vitória - ES, Brazil  
saibel@inf.ufes.br

## ABSTRACT

Watching a movie in which audio and video are out of sync is certainly an unpleasant experience, even with a big bucket of fresh popcorn. In a similar way, viewers can experience a negative experience when subtitles are not well synchronized with the media content. The point is that subtitles synchronization (or the lack thereof) plays a key role, positively or negatively, in the perception of quality that viewers have about the content. In this context, this work focuses on investigating how slightly more “relaxed” synchronization directives may affect significantly the perception of Quality of Experience (QoE) that users have when watching audiovisual content in English with subtitles in Brazilian Portuguese. Our observations are based on the analysis of data collected in an experiment with 24 users and the results suggest that subtitles beginning with delay have most influence in the viewers' QoE.

## CCS CONCEPTS

• Information systems → Information systems applications → Multimedia information systems • Human-centered computing → Human computer interaction (HCI) → HCI design and evaluation methods → User studies.

## KEYWORDS

Captions; Subtitles; Synchronization; QoE; SRT.

## ACM Reference format:

Rodrigo L. Guimarães, Jessica O. Brito and Celso A. S. Saibel. 2018. Investigating the Influence of Subtitles Synchronization in the Viewer's Quality of Experience. In *Proceedings of the XVII Brazilian Symposium on Human Factors in Computing Systems (IHC 2018)*. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3274192.3274222>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](mailto:Permissions@acm.org).

IHC 2018, October 22–26, 2018, Belém, Brazil

© 2018 Copyright is held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-6601-4/18/10...\$15.00

<https://doi.org/10.1145/3274192.3274222>

## 1 INTRODUÇÃO

Dados recentes indicam que cerca de 10% dos espectadores de transmissões televisivas usam legendas regularmente, e que este percentual sobe para 35% no caso de alguns vídeos distribuídos *on-line* [1]. O mais curioso é que a maioria desses usuários de legendas não possui qualquer problema auditivo; isso evidencia que esse tipo de mídia, complementar ao conteúdo audiovisual principal, muitas vezes desempenha um papel-chave para a Qualidade de Experiência (QoE) [41] do telespectador.

Entretanto, “nem tudo são flores no reino da Dinamarca”<sup>1</sup>. Legendas de baixa qualidade (por exemplo, com sincronia ruim e/ou tradução pobre) podem levar à insatisfação do usuário e, muitas vezes, fazer com que esse (ou essa) simplesmente desista de consumir o conteúdo audiovisual antes mesmo da “diversão” começar [12][44].

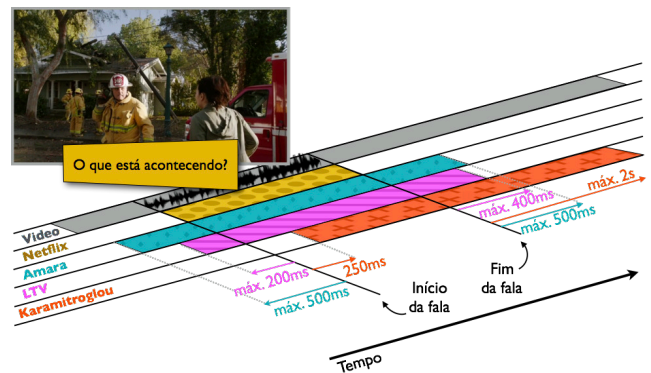


Figura 1. *Playback* de conteúdo audiovisual com legenda, onde o tempo de início e de fim de exibição do texto de cada legenda pode variar de acordo com o guia de legendagem adotado. Screenshot extraído de episódio da série *Scorpion* (<http://www.imdb.com/title/tt3514324/>).

<sup>1</sup> A tradução correta seria “Há algo de podre no Estado da Dinamarca”, Hamlet (1.4), dito por Marcelo a Horácio. A frase foi propositalmente alterada para ilustrar um problema decorrente do processo de legendagem.

Na tentativa de minimizar alguns dos problemas de QoE em se tratando de vídeos legendados, legendadores geralmente seguem recomendações especificadas em guias de legendagem [1][20][30][36]. Esses guias tratam de diferentes aspectos das legendas, tais como o nível da tradução para o idioma alvo [31][40], a quantidade de linhas sendo exibidas por vez, a quantidade de caracteres exibidos por linha (CPL) e por segundo (CPS), o tempo mínimo/máximo de duração da legenda para uma dada quantidade de texto e a sincronia (ajuste do ponto de início e de fim entre texto e fala) [10].

Neste contexto, este trabalho investiga o impacto de legendas na QoE dos usuários quando “o tempo está fora dos eixos”<sup>2</sup> (vide Figura 1). Em particular, a questão de pesquisa que pretendemos responder é:

*As diferentes diretivas de sincronia propostas em guias de legendagem influenciam, de forma significativa, a QoE reportada pelos usuários? Em caso positivo, como?*

A metodologia de pesquisa utilizada nesse processo foi baseada em uma abordagem multi-método que integra várias técnicas da área de Interação Humano-Computador (IHC). Nosso trabalho se baseia em resultados preliminares obtidos em estudos exploratórios com legendadores amadores da Web, onde observamos que o processo de sincronia é essencialmente manual, altamente iterativo e dispendioso [3]. A partir destes resultados, fizemos um levantamento das diretivas de sincronia especificadas em alguns guias de legendagem representativos encontrados na literatura e na indústria. Então, conduzimos um experimento com um grupo de usuários e analisamos estatisticamente o impacto de tais critérios de sincronia de legendas na QoE reportada pelos usuários.

Nossos resultados evidenciam que a QoE dos usuários<sup>3</sup> foi significativamente menor quando as legendas começavam depois do início da fala se comparadas com as experiências em que as mesmas legendas<sup>4</sup> iniciavam desde alguns milissegundos antes até o exato momento do início da fala. Esses resultados são relevantes não só do ponto de vista do usuário, mas também para a indústria cinematográfica (em particular, para serviços de *streaming* como Netflix<sup>5</sup>). Afinal, este último é um dos maiores

interessados em prover “boas” experiências aos seus telespectadores para, obviamente, mantê-los fidelizados! [17]

O restante do texto está organizado da seguinte forma: Na próxima seção contextualizamos e posicionamos nossa contribuição perante trabalhos relacionados. Em seguida, apresentamos alguns guias de legendagem encontrados, destacando particularmente as recomendações dirigidas à sincronia das legendas com o áudio. Adiante, detalhamos o procedimento experimental e análise dos dados coletados com o objetivo de responder nossa pergunta de pesquisa. Finalmente, refletimos sobre as lições aprendidas, e discutimos os desafios e oportunidades de trabalhos futuros, antes de recapitular as contribuições do trabalho na conclusão.

## 2 TRABALHOS RELACIONADOS

Na área de IHC, é possível listar inúmeros esforços dedicados ao estudo de aspectos da experiência do usuário na interação com artefatos tecnológicos [21][29][33][35]. No caso de artefatos audiovisuais, por exemplo, de Santana *et al.* [37] propõem uma abordagem para anotar vídeos de testes de usabilidade, visando assim apoiar profissionais de IHC em fases pós-teste. Para isso, são oferecidos meios de navegação e anotação sobre aspectos extraídos de *logs* de eventos capturados durante uma sessão de teste de usabilidade.

No contexto de sistemas multimídia [5], também há vários trabalhos de pesquisa dedicados à anotação para enriquecimento de conteúdo audiovisual [15][19][22][23]. Por exemplo, Cunha *et al.* [7] discutem o projeto e avaliação de uma ferramenta para dispositivos móveis que permite a criação de anotações temporais em um vídeo de interesse na forma de texto, áudio ou marcação de tinta. Por sua vez, Santos *et al.* [38] propõem uma estratégia de anotação de vídeo que visa fornecer aos usuários acesso personalizável aos eventos relevantes de acordo com uma ontologia pré-definida do domínio de conteúdo de vídeo. Já Souza *et al.* [39] discutem uma abordagem para apoiar o desenvolvimento sistemático de aplicações multimídia e multissensoriais de acordo com o padrão MPEG-V.

Alguns trabalhos enfatizam a questão da acessibilidade dos conteúdos audiovisuais. Por exemplo, Lima *et al.* [25] apresentam uma solução que incorpora aspectos sintáticos e semânticos na tradução do VLibras, um serviço de geração de máquinas do conteúdo de Língua Brasileira de Sinais (LIBRAS) para TIC (TV Digital, Web, Cinema Digital e dispositivos móveis). Por outro lado, Domingues *et al.* [8] mostram e validam uma solução capaz de gerar e distribuir automaticamente em dispositivos móveis faixas de áudio acessíveis que descrevem a experiência do cinema digital.

Do ponto de vista de engenharia de documentos, Fagá Jr. *et al.* [9] apresentam uma proposta de vocabulário para aplicativos de terceiros que permite aos usuários adicionar anotações multimídia mais genéricas ao conteúdo de vídeo gerado pelo usuário, enquanto Guimarães *et al.* [13][14] propõem um conjunto de transformações para documentos multimídia que

<sup>2</sup> Referenciamos a frase de Hamlet (1.5) para indicar que alguns guias de legendagem definem que a sincronia de início e de fim das legendas não precisa estar rigorosamente alinhada, respectivamente, com o início e o fim da fala.

<sup>3</sup> Os termos “usuário”, “espectador” e “telespectador” são usados de forma intercambiável neste artigo para se referir a pessoas comuns que consomem vídeos legendados.

<sup>4</sup> No nosso estudo, mantivemos o mesmo texto das legendas e aplicamos diretivas de sincronia diferentes.

<sup>5</sup> Algumas tecnologias mencionadas, se desconhecidas, podem ser facilmente identificadas através de uma pesquisa *on-line*; Portanto, elas não serão referenciadas.

possibilita aos usuários criarem e compartilharem comentários textuais personalizados que são exibidos em sincronia com vídeos de terceiros.

Na literatura, encontramos ainda trabalhos relacionados ao compartilhamento maciço de arquivos de legendas *on-line*. Por exemplo, Tiedemann [42] discute uma abordagem baseada em dicionário que usa alinhamento automático de palavras para melhorar a qualidade da sincronia ao construir um corpus paralelo a partir de legendas traduzidas. Já Britto *et al.* [4] formalizam as práticas de colaboração na criação e compartilhamento de legendas utilizando uma metodologia de gerenciamento de processos de negócio e comparam esse processo a modelos *crowdsourcing* e de colaboração entre voluntários anônimos. Os autores ainda destacam alguns desafios e oportunidades relevantes no contexto de aplicações e sistemas multimídia.

Na área de tradução, Spolidório [40] avalia as diferentes possibilidades de tradução audiovisual para uma temporada de série de TV. A autora observa aspectos de tradução associados à série, como o contexto e cultura de cunho humorístico apresentados na visão de profissionais e criadores de legendas amadoras na Internet. Nobre [31] aponta que normas e limitações técnicas associadas influenciam o processo de legendagem. É o caso, por exemplo, de traduções literais, que nem sempre são possíveis devido ao contexto e às limitações entre o idioma de origem e de destino. Ajustes também podem acontecer por conta de restrição de espaço e tempo de exibição dos textos das legendas. A autora observa que telespectadores têm se tornado cada vez mais exigentes e ativos, apontando eventuais erros cometidos pelos tradutores.

Peloi [32] faz uma reflexão das práticas de confecção e distribuição gratuita de cópias de legendas de filmes, séries, etc., sob uma perspectiva de cidadania e resistência. Feitosa [10], por sua vez, realizou um estudo comparativo entre a legendagem comercial e *fansubbing* (legendas de fãs) em um corpus composto por filmes legendados. Em relação aos aspectos de linguagem, ambas abordagens apresentaram reduções, condensações e omissões; porém, a legenda comercial costuma fazer mais uso desses recursos. No que diz respeito ao tempo de exibição e CPL, as legendas de fãs geralmente traduzem o máximo possível mesmo que não haja tempo suficiente de leitura; já as legendas comerciais mantêm a quantidade máxima de CPL igual a 34 com tempo de leitura médio de 2s para cada linha da legenda.

Apesar de nossa pesquisa ter pontos de similaridade com os trabalhos discutidos nesta seção, neste artigo de pesquisa nos concentramos em investigar como as diretivas de sincronia especificadas em guias de legendagem influenciam a QoE dos usuários ao assistirem conteúdos legendados em português. Na próxima seção, apresentamos alguns guias de legendagem que identificamos a partir de extensa revisão de trabalhos na área. Esses guias definem várias características de apresentação de uma legenda no que diz respeito aos critérios de tradução e sincronia em relação ao conteúdo audiovisual.

**Tabela 1. Recomendações de sincronia de início e fim de legenda em alguns guias de legendagem existentes.**

Guia de Legendagem	Início da legenda (máx.)	Fim da legenda (máx.)	Língua Portuguesa?
Karamitroglou	+250ms <sup>(1)</sup>	+2s	Sim <sup>(2)</sup>
BBC	-1500ms	+1500ms	Não
Trindade	0s	0s	Sim
Amara	-500ms	+500ms	Sim
Netflix <sup>(3)</sup>	0s	0s	Sim
LTV	-200ms	+400ms	Sim

Nota: O sinal indica se o tempo de início ou fim da legenda ocorre antes (-) ou depois (+) do início ou fim do evento de fala.

<sup>(1)</sup> Esse guia, diferentemente dos demais, especifica que a exibição da legenda deve começar sempre 250ms depois do início da fala.

<sup>(2)</sup> Guia direcionado aos países da Europa, inclusive Portugal.

<sup>(3)</sup> Esse guia também propõe um método de sincronia alternativo, baseado em frames do vídeo e/ou mudanças de cena.

### 3 GUIAS DE LEGENDAGEM

De modo geral, guias de legendagem fornecem um conjunto de recomendações para criação de legendas de “boa qualidade”. A Tabela 1 sumariza as recomendações de sincronia texto-fala para cada um dos guias de legendagem apresentados nesta seção. Vale ressaltar que guias também fornecem outras recomendações, como por exemplo, CPL e CPS, que apesar de importantes, estão fora do escopo deste artigo.

Na área de linguística, Karamitroglou [20] afirma que o tempo de exibição de uma legenda está diretamente relacionado ao tempo de leitura dos usuários. Segundo o mesmo autor, jovens e adultos do continente europeu com um nível de instrução médio e idade entre 14 e 65 anos leem em média de 150 a 180 palavras por minuto; já crianças e adolescentes com idade entre 6 e 14 anos, leem de 90 a 120 palavras por minuto. Assim, suas recomendações de sincronia de legendas são baseadas nos índices de leitura.

A rede BBC possui um guia para criação/formatação de legendas, especificamente para *Closed Captions* [1], que podem ser exibidas ou não dependendo da escolha do telespectador. Esse guia afirma que não é aconselhável um conjunto rígido de regras para todas as situações, mas sim instruções e conselhos práticos, já que a sua implementação depende do conteúdo, gênero e da experiência do profissional que cria a legenda.

Já Trindade [43] realizou um estudo com 13 filmes e séries de diferentes canais de TV a cabo comparando a transcrição do texto original com o texto das legendas em Português Brasileiro. Ela chega a conclusão que não é possível identificar uma padronização e nem controle no que diz respeito ao processo técnico da legendagem feito pelos canais de TV a cabo no Brasil.

**Algoritmo 1** Algoritmo de ajuste da sincronia

```

01. function adjust(subs_ref, max_begin, max_end):
02.     subs = copy(subs_ref) /* create a copy */
03.
04.     for i = 1 to length(subs)
05.         /* add delta to start and end times */
06.         subs[i].begin += random(max_begin)
07.         subs[i].end += random(max_end)
08.
09.     for i = 1 to length(subs) - 1
10.         if subs[i].end >= subs[i+1].begin
11.             /* end time of curr entry overlaps the
12.              begin of the next one */
13.             if subs[i+1].begin <= subs_ref[i].end
14.                 /* begin of next entry comes earlier than
15.                  ref end time of the previous one */
16.                 subs[i].end = subs_ref[i].end
17.                 subs[i+1].begin = subs[i].end + lms
18.             else:
19.                 subs[i].end = subs[i+1].begin - lms
20.     return subs

```

A comunidade Amara<sup>6</sup>, que legenda TED Talks e vídeos da plataforma Udacity, fornece um conjunto de recomendações para duração da legenda, quantidade de caracteres e número de linhas, para citar alguns. Essas diretivas não se limitam a um idioma específico. Por sua vez, o Netflix fornece guias de legendagem para diversos idiomas, além de um guia geral criado especificamente para a sua plataforma, seja para conteúdos originais ou não [30].

No Brasil, a comunidade Legendas.tv<sup>7</sup> (LTV) congrega entusiastas que se auto organizam em equipes para criar legendas no idioma Português Brasileiro para séries e filmes distribuídos *on-line*. O LTV, que se baseia em uma abordagem colaborativa na qual os membros exercem papéis específicos [4], também disponibiliza uma cartilha com recomendações de legendagem para os seus parceiros [36].

Nas próximas seções, descrevemos a metodologia que seguimos para realização do experimento e análise dos dados coletados.

## 4 PROCEDIMENTO EXPERIMENTAL

Projetamos um experimento com o intuito de responder a nossa pergunta de pesquisa e especificamos que a única variável independente seria o guia de legendagem escolhido como base do processo de sincronia das legendas. Já como variáveis dependentes, sobre as quais queríamos medir o efeito do guia de legendagem utilizado (variável independente), consideramos a experiência de assistir o vídeo legendado, a qualidade da sincronia da legenda, a sua utilidade para compreensão do vídeo e os níveis de satisfação, motivação e controle a cada sessão que o vídeo legendado fosse apresentado.

Para viabilizar a realização do estudo com um número menor de participantes, optamos por um *design* de experimento dentro do grupo (*within-group design* [24]), no qual cada participante seria exposto a todas as condições (ou seja, assistir o vídeo legendado com todas as variações de sincronia resultantes da utilização de cada um dos guias escolhidos). Essa abordagem nos permitiu isolar diferenças individuais e diminuir o nível de ruído entre participantes; em contrapartida, tem como desvantagem possíveis efeitos de aprendizagem e fadiga. Para minimizar o primeiro efeito, determinamos aleatoriamente a ordem de apresentação das legendas. Já para mitigar o segundo efeito, selecionamos um *clip* curto [41] (de aproximadamente 3min e 15s) da série de TV *Scorpion*<sup>8</sup>. De acordo com os legendadores do grupo LTV entrevistados em um estudo preliminar [3], vídeos com muito diálogo, como no caso do vídeo escolhido para esse experimento, demandam mais tempo e esforço nos ajustes de sincronia das legendas.

No nosso estudo, decidimos utilizar os guias de legendagem do Karamitroglou, do Netflix, do Amara e do LTV. A seleção foi feita tomando como critério a diversidade de diretivas de sincronia e os guias de legendagem que podem ser aplicados para a criação de legendas em Português. Nesse processo, utilizamos como base uma legenda do vídeo escolhido, a qual era distribuída no site Legendas.tv. Esta legenda, codificada em formato SRT [44], foi, em seguida, modificada por um legendador experiente para que refletisse as recomendações de sincronia do guia de legendagem do Netflix (início e término da exibição das legendas coincidem, respectivamente, com início e fim da fala). A partir dessa versão, as três outras foram geradas seguindo as diretivas de sincronia dos guias de legendagem do Amara, do LTV e do Karamitroglou (vide Tabela 1). O Algoritmo 1 traz o pseudocódigo do procedimento utilizado para ajustar a sincronia das legendas para cada um desses três guias utilizando como base a versão criada pelo legendador do LTV.

Delineamos que, após assistir o vídeo com uma das legendas, isto é com os diferentes arquivos SRT seguindo os diferentes guias, cada participante preencheria um questionário com escala *Likert* de 5 pontos [24] para as variáveis dependentes de interesse que queríamos medir (vale ressaltar que para satisfação, motivação e controle utilizamos o *Self-Assessment Manikin* ou SAM [2][18]). Para apresentar o vídeo legendado, utilizamos o reprodutor de mídia Media Player Classic<sup>9</sup> em tela cheia em um *laptop* de 14 polegadas, usando o Sistema Operacional Microsoft Windows. Os caracteres do texto das legendas tinham fonte Arial, cor branca e bordas pretas. Por fim, estabelecemos que os participantes deveriam se sentar próximos ao *laptop* e utilizar fone de ouvido durante as exibições do vídeo com os diferentes arquivos de legendas.

Primeiramente, realizamos um teste piloto para refinar o procedimento. Em seguida, recrutamos 24 participantes (3 mulheres) para nosso estudo utilizando uma amostragem

<sup>6</sup> <https://amara.org/pt/subtitling-platform>

<sup>7</sup> <http://legendas.tv>

<sup>8</sup> <http://www.imdb.com/title/tt3514324/>

<sup>9</sup> <https://mpc-hc.org>

aleatória simples. Todos participantes eram alunos da nossa universidade. Treze (13) se autodeclararam brancos, 8 pardos e os outros grupos étnicos mencionados foram negro (2) e indígena (1). Vinte e um (21) alunos disseram ter entre 18-24 anos, enquanto outros 2 estão na faixa de 25-34 anos e 1 entre 35-49 anos. Entre outras informações coletadas sobre o perfil de cada participante, a mais relevante para o estudo reportado neste artigo foi o nível de conhecimento da língua inglesa. Neste quesito, 9 declararam estar no nível avançado, 10 no intermediário e 5 no básico. Os alunos não receberam qualquer compensação e obtivemos o consentimento de todos para colaborar com este trabalho de pesquisa científica.

O estudo foi conduzido por um facilitador em nosso laboratório ao longo de 2 semanas. Em média, cada sessão durou por volta de 30min e contou com uma parte de instruções (*briefing*), questionários pré-experimento, apresentação do vídeo com os 4 arquivos de legendas (Amara, Netflix, Karamitroglou e LTV), seguido de questionários após cada apresentação e entrevista ao final. Esse processo nos possibilitou obter dados e várias anotações manuscritas. Na próxima seção, nos concentramos na apresentação da análise estatística sobre os dados coletados.

## 5 ANÁLISE DOS RESULTADOS

Antes de iniciar o processo de análise de significância estatística, os dados dos questionários foram cuidadosamente processados, codificados e consolidados no formato CSV (*Comma-separated values*). A partir daí, foi possível formular e testar estatisticamente algumas hipóteses em relação ao impacto dos diferentes guias de legendagem na QoE dos participantes. Para isso utilizamos extensivamente alguns pacotes estatísticos especializados, como RStudio<sup>10</sup> v1.1.383 e G\*Power<sup>11</sup> v3.1.9.3.

Na análise dos dados consideramos inicialmente o método estatístico de análise de variância (ANOVA) com medidas repetidas. A análise de variância é utilizada para comparar se há diferenças significativas nas médias de uma variável dependente em função de uma única variável independente (chamada de fator) com 2 ou mais categorias ou níveis. No nosso caso, as variáveis dependentes são a experiência de assistir o vídeo legendado, a qualidade da sincronia da legenda, a sua utilidade para compreensão do vídeo e os níveis de satisfação, motivação e controle a cada sessão que o vídeo e suas legendas eram apresentados. Já a única variável independente é o guia utilizado no processo de legendagem, que poderia ser Amara, Netflix, Karamitroglou ou LTV.

Lançamos mão do software G\*Power para investigar se nossa amostra seria suficiente para detectar um efeito com um grau de confiança satisfatório. Utilizamos como referência o teste de análise de variâncias com medidas repetidas e especificamos o tamanho do efeito = 0,253 (efeito médio),  $p = 0,05$  (nível de

significância  $\alpha$  ou erro Tipo I<sup>12</sup>), potência<sup>13</sup> = 0,95, número de grupos igual a 1 (análise dentro do grupo) e número de medidas por participante igual a 4 (condições de legendagem). Constatamos que um total de 35 amostras seriam necessárias para obter um resultado estatisticamente significativo com os parâmetros fornecidos. Constatamos que no nosso caso ( $N = 24$ ), a chance de corretamente rejeitar as hipóteses nulas é de aproximadamente 83%.

Para que a interpretação dos resultados obtidos na análise de variância com medidas repetidas seja feita corretamente, presume-se que os dados tenham certas características, como, por exemplo, que as variáveis dependentes sejam contínuas. Como nossos questionários utilizaram a escala *Likert*, essa conclusão não é trivial e o debate sobre ela ser considerada contínua ou ordinal é recorrente na literatura [16]. Para evitar que a validade dos nossos resultados fosse questionada com base nesse critério, optamos então por utilizar o teste de Friedman, considerado o equivalente não paramétrico da análise de variância com medidas repetidas.

Para o teste de Friedman, os pré-requisitos podem ser considerados como mais relaxados (em contrapartida, alguns críticos consideram que o método de análise de variância com medidas repetidas apesar de mais restritivo, é mais poderoso para detectar diferenças entre grupos [27]). Por exemplo, um grupo ou categoria precisa ser medido 3 ou mais vezes, as medidas precisam ser independentes umas das outras, as variáveis dependentes devem ser medidas no nível contínuo ou ordinal e as amostras não precisam ser normalmente distribuídas. Vale destacar que todos esses pré-requisitos foram plenamente satisfeitos pelo nosso conjunto de dados.

### 5.1 Experiência de Assistir Vídeo Legendado

$H_0$ : O guia de legendagem utilizado não teve efeito significativo na experiência de assistir o vídeo legendado (ou seja,  $\mu_{\text{experiência c/ Amara}} = \mu_{\text{experiência c/ Netflix}} = \mu_{\text{experiência c/ Karamitroglou}} = \mu_{\text{experiência c/ LTV}}$ ).

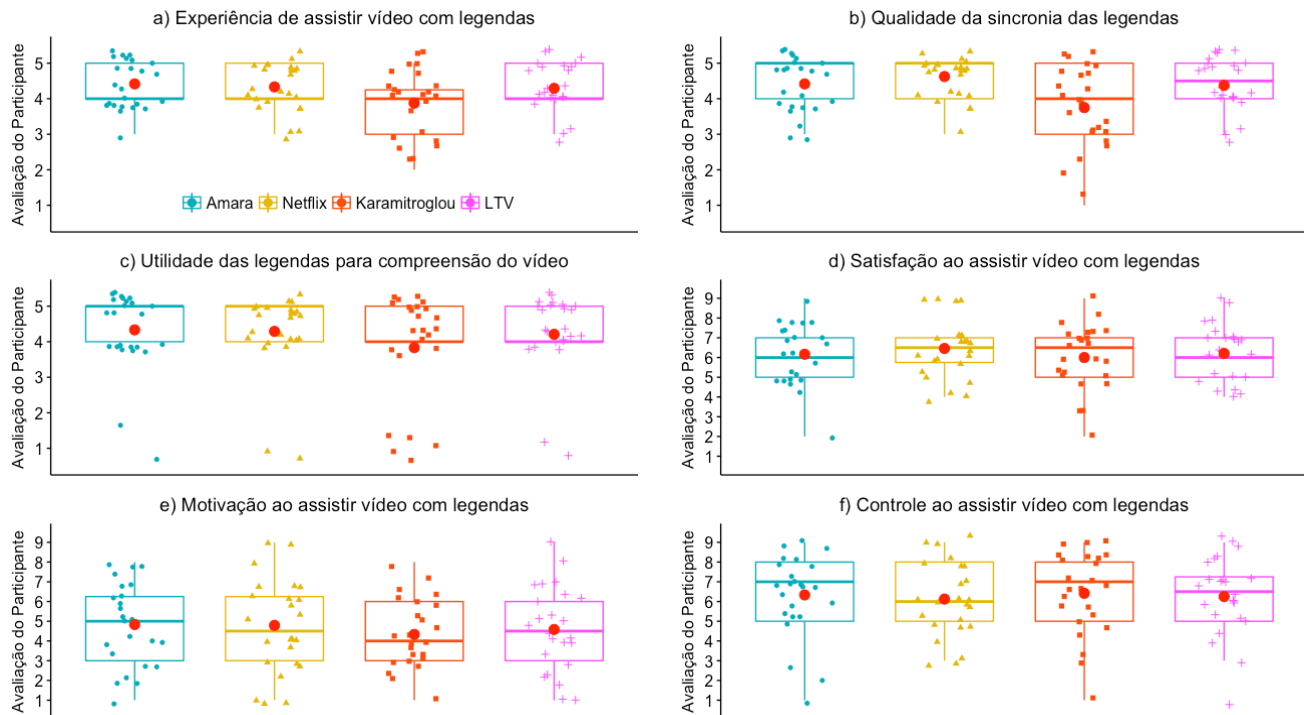
Na avaliação da hipótese nula  $H_0$  de que não houve diferença significativa na experiência dos participantes ao assistir o vídeo com legendas que seguiram diferentes diretrizes de sincronia, a variável independente, guia de legendagem, incluía 4 categorias: guia do Amara ( $\mu_{\text{experiência c/ Amara}} = 4,42$ , SD = 0,58, Mdn = 4 e  $n = 24$ ), guia do Netflix ( $\mu_{\text{experiência c/ Netflix}} = 4,33$ , SD = 0,70, Mdn = 4 e  $n = 24$ ), guia do Karamitroglou ( $\mu_{\text{experiência c/ Karamitroglou}} = 3,88$ ,

<sup>12</sup> Nível de significância para a estatística: uma determinação é feita para rejeitar ( $p < \alpha$ ) ou manter ( $p > \alpha$ ) a hipótese nula. Um valor de  $p$  de 1 significa que aceitamos a hipótese nula como verdadeira; já um valor de  $p$  de 0 significa que aceitamos a hipótese nula como falsa.

<sup>13</sup> A potência estatística permite determinar a probabilidade de cometer um erro de Tipo II (denotado por  $\beta$  [6]). Em outras palavras, é possível calcular a probabilidade de ter um falso negativo (quando falhamos em rejeitar  $H_0$  e  $H_0$  é realmente falsa) pela fórmula potência =  $1 - \beta$ .

<sup>10</sup> <https://www.rstudio.com>

<sup>11</sup> <http://www.gpower.hhu.de>



**Figura 2.** Diagrama de caixa (*boxplot*) que sumariza observações feitas durante experimento. Um ruído (*jitter*) foi adicionado aos dados para dar uma visão da dispersão das avaliações dos participantes. Os círculos vermelhos ao centro das caixas representam os valores médios das variáveis observadas para os diferentes guias de legendagem.

SD = 0,90, Mdn = 4 e  $n = 24$ ) e guia do LTV ( $\mu_{\text{experiência c/ LTV}} = 4,29$ , SD = 0,69, Mdn = 4 e  $n = 24$ ). A Figura 2.a) apresenta um resumo sobre a amostra e as observações que foram feitas. Executamos o teste de Friedman no RStudio e descobrimos que o fator guia de legendagem teve efeito estatisticamente significativo na experiência reportada pelos participantes ao assistir o vídeo legendado,  $\chi^2(3) = 12,28$  e  $p = 0,0065$ .

Em seguida, realizamos uma análise *post-hoc* utilizando o teste de postos sinalizados de Wilcoxon com correção de Bonferroni, resultando em um nível de significância definido em  $p < 0,008$ . Constatamos que não houve diferenças estatisticamente significantes entre os guias de legendagem do Amara-Netflix ( $Z = 0,46$  e  $p = 0,73$ ), do Amara-LTV ( $Z = 0,93$  e  $p = 0,45$ ), do Netflix-Karamitroglou ( $Z = 2,51$  e  $p = 0,01$ ), do Netflix-LTV ( $Z = 0,36$  e  $p = 0,86$ ) e do Karamitroglou-LTV ( $Z = -1,89$  e  $p = 0,06$ ). Porém, encontramos uma diferença estatisticamente significativa entre os guias do Amara-Karamitroglou ( $Z = 2,66$  e  $p = 0,007$ ). Portanto, neste caso, rejeitamos a hipótese nula  $H_0$  ( $\mu_{\text{experiência c/ Amara}} \neq \mu_{\text{experiência c/ Karamitroglou}}$ ).

## 5.2 Qualidade da Sincronia das Legendas

$H_0$ : O guia de legendagem utilizado não teve efeito significativo na avaliação dos participantes sobre a qualidade da sincronia das legendas ( $\mu_{\text{sincronia c/ Amara}} = \mu_{\text{sincronia c/ Netflix}} = \mu_{\text{sincronia c/ Karamitroglou}} = \mu_{\text{sincronia c/ LTV}}$ ).

Seguimos uma abordagem semelhante para avaliar a hipótese nula  $H_0$  de que não houve diferença significativa na qualidade da sincronia das legendas reportada pelos participantes. Para as amostras nas quais a sincronia das legendas seguiu o guia do Amara obtivemos uma qualidade média  $\mu_{\text{sincronia c/ Amara}} = 4,42$  (SD = 0,72, Mdn = 5 e  $n = 24$ ), enquanto que para os demais guias de legendagem a qualidade média foi  $\mu_{\text{sincronia c/ Netflix}} = 4,63$  (SD = 0,58, Mdn = 5 e  $n = 24$ ),  $\mu_{\text{sincronia c/ Karamitroglou}} = 3,75$  (SD = 1,15, Mdn = 5 e  $n = 24$ ) e  $\mu_{\text{sincronia c/ LTV}} = 4,38$  (SD = 0,71, Mdn = 4,5 e  $n = 24$ ). A Figura 2.b) descreve essas medidas. Já o teste de Friedman indicou que o fator guia de legendagem teve efeito estatisticamente significativo na avaliação dos participantes quanto à sincronia das legendas em relação ao vídeo,  $\chi^2(3) = 13,96$  e  $p = 0,003$ .

Também realizamos a análise *post-hoc* utilizando o teste de postos sinalizados de Wilcoxon com correção de Bonferroni ( $p < 0,008$ ) e observamos que não houve diferenças estatisticamente significantes entre os guias de legendagem do Amara-Netflix ( $Z = -1,07$  e  $p = 0,35$ ), do Amara-Karamitroglou ( $Z = 2,16$  e  $p = 0,03$ ), do Amara-LTV ( $Z = 0,30$  e  $p = 0,87$ ), do Netflix-LTV ( $Z = 1,32$  e  $p = 0,23$ ) e do Karamitroglou-LTV ( $Z = -2,27$  e  $p = 0,02$ ). Entretanto, constatamos que houve uma diferença estatisticamente significativa entre os guias do Netflix-Karamitroglou ( $Z = 3,17$  e  $p = 0,0008$ ). Assim, também rejeitamos esta hipótese nula  $H_0$  ( $\mu_{\text{sincronia c/ Netflix}} \neq \mu_{\text{sincronia c/ Karamitroglou}}$ ).

### 5.3 Demais Hipóteses Nulas

Repetimos o mesmo procedimento descrito anteriormente para analisar o efeito do fator guia de legendagem na utilidade da legenda para a compreensão do vídeo pelos participantes, assim como para as outras variáveis dependentes como satisfação, motivação e controle durante a apresentação do vídeo legendado (vide Figura 2.c)-f), respectivamente). Em nossa análise, para nenhuma dessas variáveis pudemos constatar que o fator guia de legendagem tenha influenciado de forma estatisticamente significativa os dados reportados pelos participantes ( $p > 0,05$  nível de significância  $\alpha$ ). Em outras palavras, isso significa que não podemos rejeitar nenhuma das hipóteses nulas relacionadas à compreensão, satisfação, motivação e controle ao assistir o vídeo legendado.

Por fim, investigamos a hipótese nula de que não há uma relação estatística entre as variáveis dependentes experiência e qualidade da sincronia das legendas.

$H_0$ : Não há correlação entre as variáveis experiência de assistir o vídeo legendado e qualidade da sincronia atribuída a essas legendas pelos participantes.

Uma maneira de analisar tal relação é através de coeficientes de correlação, que medem o grau de associação ou dependência entre duas variáveis. Aqui, é importante ressaltar que a correlação não implica necessariamente causalidade (ou seja, quanto maior a experiência reportada ao assistir o vídeo legendado, maior ou menor a qualidade atribuída à legenda).

Inicialmente, consideramos o coeficiente de correlação de Pearson ( $\rho$ ), mas logo o descartamos pois este assume, dentre outras condições, que as variáveis sejam medidas no nível contínuo. Sob tais circunstâncias, avaliamos outros coeficientes de correlação para estatísticas não paramétricas. Identificamos alguns testes alternativos, como o coeficiente de correlação de postos de Spearman, o coeficiente de Kendall e o lambda de Goodman e Kruskal. Apesar de não assumir os pressupostos de independência entre observações, normalidade, linearidade e homocedasticidade<sup>14</sup>, esses coeficientes só devem ser usados se existir uma relação monótona<sup>15</sup> entre os dados (condição satisfeita por nosso conjunto de dados).

De acordo com nossos resultados experimentais utilizando o coeficiente de correlação de postos de Spearman, o grau de associação entre as variáveis dependentes, (1) experiência de assistir o vídeo legendado e (2) qualidade da sincronia atribuída a essa legenda ficou em torno de 0,60 ( $p < 0,01$ ). Vale ressaltar que o coeficiente de Spearman registra valores no intervalo [0,1], no qual 0 indica relações completamente independentes, enquanto

<sup>14</sup> Termo para designar variância constante dos erros experimentais para observações distintas.

<sup>15</sup> De um modo geral, em uma relação monótona entre dois conjuntos ordenados, as variáveis devem aumentar nos valores em conjunto (crescente), ou quando uma aumenta e a outra deve diminuir (decrecente).

**Tabela 2. Número de legendas seguindo um determinado guia de legendagem em que, após processo de ajuste da sincronia, o início ou fim da legenda ainda se encontra em conformidade em relação a outro guia de legendagem (entre parêntesis).**

	Início da legenda	Fim da legenda
Amara (relaç. Netflix)	33%	33%
LTV (relaç. Netflix)	32%	35%
Karamitroglou (relaç. Netflix)	0%	0%
Amara (relaç. LTV)	79%	96%
Karamitroglou (relaç. Amara)	0%	61%
Karamitroglou (relaç. LTV)	0%	56%

Nota: A comparação entre o tempo de início e de fim das legendas foi feita da diretiva de sincronia mais ampla em relação a mais estrita (rígida).

que valores próximos de 1 representam alto grau de dependência entre as associações. Portanto, com base em nossa análise, podemos chegar à conclusão de que a correlação entre nossas variáveis dependentes, em torno de 60%, é bastante significativa.

## 6 DISCUSSÃO

Nas seções anteriores, nossa análise dos dados mostrou que a escolha do guia de legendagem teve efeito significativo tanto na experiência de assistir o vídeo legendado, quanto na percepção da qualidade da sincronia das legendas com o vídeo. Devemos também destacar que identificamos uma correlação bastante significativa entre tais variáveis. Por fim, identificamos que dentre os guias selecionados, o que obteve pior avaliação foi o especificado por Karamitroglou (o único a recomendar que a apresentação das legendas inicie após o início da fala).

Após essa análise, decidimos revisitar os arquivos de legendas gerados para o experimento. Neste processo, observamos que certas legendas de um dado arquivo se enquadravam ainda em outros guias de legendagem (vide Tabela 2). Em outras palavras, após o ajuste de sincronia para um dado guia, trechos de legendas também satisfaziam as recomendações de sincronia de outros guias. Chegamos à conclusão que esse comportamento se deve à característica do conteúdo do vídeo utilizado (muitos diálogos próximos uns dos outros), onde o ajuste de um dado bloco de legenda precisou ser revisto para que não impactasse a exibição do bloco de legenda que vinha antes ou depois dele.

Ainda com relação à Tabela 2, podemos notar que o arquivo de legendas criado seguindo o guia do Karamitroglou, como era de se esperar, não possui nenhuma legenda cujo início se enquadre dentro de qualquer outro guia de legendagem que utilizamos. Já os arquivos de legendas para os guias do Amara e do LTV ficaram com uma conformidade do tempo de início da legenda de pouco mais de 30% em relação ao Netflix. Podemos observar ainda que existe uma alta interseção entre o início das legendas

para os guias do Amara e do LTV (79%). Como visto na análise estatística apresentada na seção anterior, tal correspondência é também refletida nas médias calculadas para as diferentes variáveis dependentes ( $\mu_{\text{experiência c/ Amara}} = 4,42$  e  $\mu_{\text{experiência c/ LTV}} = 4,29$ ;  $\mu_{\text{sincronia c/ Amara}} = 4,42$  e  $\mu_{\text{sincronia c/ LTV}} = 4,38$ ).

Já com relação à sincronia de fim de exibição da legenda, observamos que 33% das legendas seguindo o guia do Amara, 35% seguindo o LTV e 0% seguindo o guia do Karamitroglou terminavam exatamente no final da fala do personagem da cena (ou seja, ainda eram compatíveis com o guia do Netflix). Para as legendas seguindo o guia do Amara, identificamos uma equivalência ainda maior em relação ao guia do LTV (96%). Por fim, observamos que as legendas do guia do Karamitroglou tinham uma alta equivalência em relação aos guias do Amara e LTV. Isso quer dizer que, apesar de uma legenda do guia do Karamitroglou poder levar até 2s após o fim da fala para deixar de ser exibida, no nosso estudo, em 61% dos casos, as legendas terminavam com até 500ms depois da fala, ou seja, ainda se encaixavam dentro das restrições mais rígidas de sincronia do guia do Amara para o fim de apresentação da legenda com relação à fala dos personagens em cena. De forma análoga, observamos que 56% das legendas associadas ao guia do Karamitroglou no experimento ainda satisfaziam as diretivas mais rígidas do guia do LTV.

Esses resultados sugerem que a QoE reportada pelos usuários é mais sensível à sincronia de início do que à sincronia de fim de legenda com a fala dos personagens. Essa análise também dá indícios de que o excesso de preciosismo dos legendadores na marcação do início e do fim das legendas pode ser um esforço desnecessário no processo de legendagem, pois desde que satisfeita a condição de que a exibição da legenda se inicie um pouco antes ou até pelo menos o início fala, a QoE do usuário ao assistir o vídeo legendado não é afetada de forma significativa. Além disso, outras diretivas que ajustam as legendas posteriormente de acordo com CPS, linhas, etc. acabam alterando o resultado da sincronia feito de forma manual pelos legendadores. Como eles realizam, geralmente, muitas iterações na etapa de sincronia [3], esse relaxamento do tempo de início de exibição da legenda seria um ponto de melhoria no processo (ou seja, poderiam investir menos tempo nesta fase, reduzindo o tempo total para a legendagem dos vídeos).

Apesar dos resultados indicarem descobertas interessantes, é importante ressaltar que tanto os efeitos da variável independente sobre as variáveis dependentes, quanto à associação significativa (60%) entre as variáveis experiência e qualidade da sincronia não implica que podemos generalizar as conclusões para qualquer guia de legendagem e qualquer população. Isso porque fixamos de antemão quais seriam os guias utilizados no experimento [28]. Também é importante lembrar que aspectos linguísticos, muito importantes para a QoE, não foram o foco deste trabalho.

O estudo também teve algumas limitações. Por um lado, o grupo de participantes pode ser considerado homogêneo em termos de faixa etária, nível de escolaridade e hábitos de consumo de

conteúdo audiovisual. Por outro, utilizamos um computador (e não uma TV ou dispositivo móvel) para apresentação do vídeo legendado [34] e sem focar em outras questões ergonômicas, como iluminação, tamanho e distância da tela, altura e posicionamento do usuário numa cadeira (e não num sofá).

Além disso, não podemos descartar a influência que outros fatores como idade, gênero, grupo étnico, tipo de conteúdo, a língua materna, o idioma apresentado no vídeo, nível de proficiência do usuário na língua estrangeira e duração do vídeo, para citar alguns, podem ter na QoE do usuário. No nosso estudo, nos restringimos a utilizar um vídeo curto com muitos diálogos, ou seja, com grande número de legendas, em Português Brasileiro, aparecendo intercaladamente uma após a outra. Portanto, é preciso ter em mente que nossos resultados não devem ser imediatamente generalizados e alguns desses fatores devem ser devidamente investigados em trabalhos futuros.

## 7 CONCLUSÃO

Este trabalho investigou a influência da sincronia de legendas na QoE dos usuários ao assistir um vídeo legendado. Para isso, realizamos um experimento com um grupo de 24 participantes e respondendo a nossa pergunta de pesquisa, observamos que, para um mesmo conjunto de legendas, as diferentes diretivas de sincronia propostas em guias de legendagem podem influenciar, de forma significativa, a QoE reportada pelos usuários. Isso nos leva a concluir que “legendas em si podem ser boas ou ruins; e em grande parte isso depende do processo de sincronia”<sup>16</sup>.

Os nossos resultados sugerem ainda que a percepção dos usuários em relação à qualidade da sincronia texto-fala é maior no início do que no fim da apresentação das legendas. Ademais, nosso trabalho traz uma contribuição importante para o estado da arte. Apesar da literatura sugerir que legendas não devem aparecer mais de 2s após as palavras terem sido faladas [1], o impacto na QoE do usuário já é perceptível transcorridos os primeiros 250ms. Cabe investigar se o impacto seria perceptível para valores menores (por exemplo, 50 ou 100ms).

Além das possibilidades de trabalhos futuros mencionadas na seção anterior, pretendemos verificar se guias de legendagem são efetivamente seguidos durante o processo de legendagem. Para isso, visamos analisar as características técnicas (ou diretivas) extraídas a partir dos arquivos de legenda. Consequentemente, poderemos investigar a significância da correlação dessas diretivas em relação à qualidade atribuída pelos usuários através de comentários (por exemplo, número de curtidas ou notas que um arquivo de legenda recebe).

Por fim, estamos avaliando a viabilidade de utilizarmos sensores de resposta galvânica da pele (GSR) para medir o nível de *stress*, ou mesmo o nível emocional e de engajamento, dos usuários ao assistir vídeos legendados. Sensores têm atraído o interesse da

<sup>16</sup> Referência ao trecho “Nada em si é bom ou mau; tudo depende daquilo que pensamos”, Hamlet (2.2).



comunidade de IHC [11][26][45], pois têm se tornado mais um instrumento no processo de avaliação, em particular, como forma de complementar abordagens mais subjetivas.

## AGRADECIMENTOS

Os autores agradecem aos experientes legendadores voluntários pela participação em estudos preliminares e também aos usuários do experimento que permitiu a realização e escrita deste artigo. Durante a execução deste trabalho Jessica O. Brito recebeu financiamento da FAPES (#789/2015) e Celso A. S. Santos do CNPq (#312148/2014-3) e FAPES (#67927378/2015).

## REFERÊNCIAS

- [1] BBC. 2017. Subtitle Guidelines. Retrieved February 21, 2018 from <http://bbc.github.io/subtitle-guidelines/>
- [2] Margaret M. Bradley and Peter J. Lang. 1994. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25, 1 (March 1994), 49-59. [http://dx.doi.org/10.1016/0005-7916\(94\)90063-9](http://dx.doi.org/10.1016/0005-7916(94)90063-9)
- [3] Jessica O. Britto. 2018. Análise do Impacto da Sincronia de Legendas na Qualidade de Experiência do Usuário. Master's dissertation. UFES, ES, Brazil. <http://repositorio.ufes.br/handle/10/9858>
- [4] Jessica O. Brito, Rodrigo L. Guimaraes, and Celso A. S. Santos. 2017. Investigating the Collaborative Process of Subtitles Creation and Sharing for Videos on the Web. In *Proc. of the Brazilian Symp. on Multimedia and the Web (WebMedia '17)*, 69-72. <https://doi.org/10.1145/3126858.3131592>
- [5] Dick C. A. Bulterman, Pablo Cesar, and Rodrigo L. Guimarães. 2013. Socially-aware multimedia authoring: Past, present, and future. *ACM Trans. Multimedia Comput. Commun. Appl.* 9, 1s, Article 35 (October 2013), 23 pages. <http://dx.doi.org/10.1145/2491893>
- [6] Jacob Cohen. 1962. The statistical power of abnormal – social psychological research: A review. *Journal of Abnormal and Social Psychology*, 65, 145-153. <http://dx.doi.org/10.1037/h0045186>
- [7] Bruna C. R. Cunha, Olibario J. Machado Neto, and Maria G. C. Pimentel. 2013. A heuristic evaluation of a mobile annotation tool. In *Proc. of the Brazilian symp. on Multimedia and the web (WebMedia '13)*, 89-92. <http://dx.doi.org/10.1145/2526188.2526232>
- [8] Leonardo A. Domingues, Virginia P. Campos, Tiago M. U. Araújo, and Guido L. de S. Filho. 2016. Accessibility in Digital Cinema: A Proposal for Generation and Distribution of Audio Description. In *Proc. of the Brazilian Symp. on Multimedia and the Web (WebMedia '16)*, 119-126. <https://doi.org/10.1145/2976796.2976867>
- [9] Roberto Fagá Jr., Vivian G. Motti, Renan G. Cattelan, Cesar A. C. Teixeira, and Maria G. C. Pimentel. 2010. A social approach to authoring media annotations. In *Proc. of the ACM symp. on Document engineering (DocEng '10)*, 17-26. <http://dx.doi.org/10.1145/1860559.1860566>
- [10] Marcos P. Feitosa. 2009. *Legendagem comercial e legendagem pirata: um estudo comparado*. Doctoral thesis. UFMG, MG, Brazil. <http://hdl.handle.net/1843/ALDR-83ZHSQ>
- [11] Vinícius P. Gonçalves, Vânia P. A. Neris, Jó Ueyama, Gabriel T. Giancristofaro, Geraldo P. R. Filho, and Sibelius Seraphini. 2013. Emoções na interação humano-computador: um estudo considerando sensores. In *Proc. of the Brazilian Symp. on Human Factors in Computing Systems (IHC '13)*, 252-255.
- [12] Rodrigo L. Guimarães, Priscilla Avegliano, and Lucas C. Villa Real. 2016. A Lightweight and Efficient Mechanism for Fixing the Synchronization of Misaligned Subtitle Documents. In *Proc. of the ACM Symposium on Document Engineering (DocEng '16)*, 175-184. <https://doi.org/10.1145/2960811.2960812>
- [13] Rodrigo L. Guimarães, Pablo Cesar, and Dick C. A. Bulterman. 2010. Creating and sharing personalized time-based annotations of videos on the web. In *Proc. of the ACM symposium on Document engineering (DocEng '10)*, 27-36. <https://doi.org/10.1145/1860559.1860567>
- [14] Rodrigo L. Guimarães, Pablo Cesar, and Dick C. A. Bulterman. 2012. "Let me comment on your video": supporting personalized end-user comments within third-party online videos. In *Proc. of the Brazilian symp. on Multimedia and the web (WebMedia '12)*, 253-260. <http://dx.doi.org/10.1145/2382636.2382690>
- [15] Rodrigo L. Guimarães, Carlos S. Soares Neto, and Luiz F. G. Soares. 2008. A visual approach for modeling spatiotemporal relations. In *Proc. of the ACM symposium on Document engineering (DocEng '08)*, 285-288. <https://doi.org/10.1145/1410140.1410202>
- [16] Karen Grace-Martin. 2008. Can Likert Scale Data ever be Continuous? Retrieved February 8, 2018 from [www.articlealley.com/article\\_670606\\_22.htm](http://www.articlealley.com/article_670606_22.htm)
- [17] Blake Hallinan and Ted Striphas. 2014. Recommended for you: The Netflix Prize and the production of algorithmic culture. *New Media & Society*, 18, 1 (June 2014), 117-137. <https://doi.org/10.1177/1461444814538646>
- [18] Elaine C. S. Hayashi, Julián E. Gutiérrez Posada, Vanessa R. M. L. Maiké, and M. Cecilia C. Baranauskas. 2016. Exploring new formats of the Self-Assessment Manikin in the design with children. In *Proc. of the 15th Brazilian Symp. on Human Factors in Computing Systems (IHC '16)*, Article 27, 10 pages. <https://doi.org/10.1145/3033701.3033728>
- [19] Jack Jansen, Pablo Cesar, Rodrigo L. Guimaraes, and Dick C. A. Bulterman. 2012. Just-in-time personalized video presentations. In *Proc. of the ACM symposium on Document engineering (DocEng '12)*, 59-68. <https://doi.org/10.1145/2361354.2361368>
- [20] Fotios Karamitroglou. 1998. A proposed set of subtitling standards for Europe. *Translation Journal*, 2, 2, 1-15.
- [21] Artur H. Kronbauer, Diferson Machado, Omar G. Junior, and Celso A. S. Santos. 2014. Sherlock: uma plataforma para avaliar a experiência dos usuários com dispositivos móveis. In *Proc. of the 13th Brazilian Symp. on Human Factors in Computing Systems (IHC '14)*, 236-245.
- [22] Fons Kuijk, Rodrigo L. Guimarães, Pablo Cesar, and Dick C. A. Bulterman. 2009. Adding dynamic visual manipulations to declarative multimedia documents. In *Proc. of the ACM symposium on Document engineering (DocEng '09)*, 149-152. <https://doi.org/10.1145/1600193.1600227>
- [23] Fons Kuijk, Rodrigo L. Guimarães, P. Cesar, and Dick C. A. Bulterman. 2009. From Photos to Memories: A User-Centric Authoring Tool for Telling Stories with your Photos. *User Centric Media*. Springer Berlin Heidelberg, 13-20. [https://link.springer.com/chapter/10.1007%2F978-3-642-12630-7\\_2](https://link.springer.com/chapter/10.1007%2F978-3-642-12630-7_2)
- [24] Jonathan Lazar, Jinjuan H. Feng, and Harry Hochheiser. 2017. *Research Methods in Human-Computer Interaction*. Elsevier/Morgan Kaufmann Publishers, 2 ed. ISBN: 978-0128053904
- [25] Manuella A. C. B. Lima, Tiago M.U. de Araújo, and Erickson S. de Oliveira. 2015. Incorporation of Syntactic-Semantic Aspects in a LIBRAS Machine Translation Service to Multimedia Platforms. In *Proc. of the Brazilian Symp. on Multimedia and the Web (WebMedia '15)*, 133-140. <http://dx.doi.org/10.1145/2820426.2820434>
- [26] Camila L. B. Maia and Elizabeth S. Furtado. 2016. A study about psychophysiological measures in user experience monitoring and evaluation. In *Proc. of the 15th Brazilian Symp. on Human Factors in Computing Systems (IHC '16)*, Article 7, 9 pages. <https://doi.org/10.1145/3033701.3033708>
- [27] Salvatore S. Mangiafico. 2016. *Summary and Analysis of Extension Program Evaluation in R*. Available at <http://rcompanion.org/handbook/>
- [28] John H. McDonald. 2014. *Handbook of Biological Statistics*. Sparky House Publishing, Baltimore, Maryland, 3rd. edition, 190-208. Available at <http://www.biostathandbook.com/linearregression.html>
- [29] Flabiéli P. F. de Miranda and Soraia S. Prietch. 2016. Emotion-Libras: Instrument for use in Human-Computer Interaction Researches considering People who are Deaf as Potential Users of Technology. In *Proc. of the 15th Brazilian Symp. on Human Factors in Computing Systems (IHC '16)*, Article 44, 4 pages. <https://doi.org/10.1145/3033701.3033746>
- [30] Netflix. 2016. Timed Text Style Guide: General Requirements. Retrieved February 8, 2018 from <https://backlothelp.netflix.com/hc/en-us/articles/215758617-Timed-Text-Style-Guide-General-Requirements>
- [31] Naiara M. Nobre. 2013. A Legendagem no Brasil: Interferências Linguísticas e Culturais nas Escolhas Tradutórias e o Uso de Legendas em Aulas de Língua Estrangeira. *Letras Escreve* 2, 1, 91-108. <https://periodicos.unifap.br/index.php/letras/article/view/489>
- [32] Danilo Peloi. 2012. Legendamento Não Autorizado na Rede: Práticas de Cópia e Resistência. In: *Anais do Congresso Internacional Interdisciplinar Em Sociais e Humanidades*, 1, 2012, Niterói. ISSN 2316-266X
- [33] Mateus Pinheiro, Nayana Carneiro, and Ticianne Darin. 2017. Recommendations for the Design of Urban Mobility Applications Based on the Study of the User Experience. In *Proc. of the XVI Brazilian Symp. on Human Factors in Computing Systems (IHC '17)*, Article 44, 10 pages. <https://doi.org/10.1145/3160504.3160517>

- [34] Jacob M. Rigby, Duncan P. Brumby, Anna L. Cox, and Sandy J. J. Gould. 2016. Watching movies on netflix: investigating the effect of screen size on viewer immersion. In *Proc. of the 18th Int. Conf. on Human-Computer Interaction with Mobile Devices and Services Adjunct (MobileHCI '16)*, 714-721. <https://doi.org/10.1145/2957265.2961843>
- [35] Luis Rivero and Tayana Conte. 2017. A Systematic Mapping Study on Research Contributions on UX Evaluation Technologies. In *Proc. of the XVI Brazilian Symp. on Human Factors in Computing Systems (IHC '17)*, Article 5, 10 pages. <https://doi.org/10.1145/3160504.3160512>
- [36] Vanessa M. M. de Sá. 2015. From Orkut to Facebook: How Brazilian pirate audiences utilize social media to create sharing subcultures. *International Journal of Communication*, 9 (January 2015), 852-869. <http://ijoc.org/index.php/ijoc/article/view/3730>
- [37] Wagner F. de Santana, Claudio S. Pinhanez, and Rogério A. de Paula. 2014. Unsupervised annotation of usability test videos. In *Proc. of the 13th Brazilian Symp. on Human Factors in Computing Systems (IHC '14)*, 207-215.
- [38] Celso A. S. Santos, Alexandre V. S. G. Santos, and Tatiana A. Tavares. 2007. Uma Estratégia para a Construção de Ambientes para a Descrição Semântica de Vídeos. In *Proc. of the Brazilian symp. on Multimedia and the web (WebMedia '07)*.
- [39] Marcelo F. Sousa, Raoni Kulesza, and Carlos A. G. Ferraz. 2016. A Model-driven Approach for MulSeMedia Application Domain. In *Proc. of the Brazilian Symp. on Multimedia and the Web (WebMedia '16)*, 111-118. <https://doi.org/10.1145/2976796.2976872>
- [40] Samira Spolidorio. 2017. *Comunidades online e legendas de fãs: novas formas de produzir e consumir legendas*. Master's dissertation. Unicamp, SP, Brazil. <http://repositorio.unicamp.br/jspui/handle/REPOSIP/322463>
- [41] Robert C. Streijl, Stefan Winkler, and David S. Hands. 2016. Mean opinion score (MOS) revisited: methods and applications, limitations and alternatives. *Multimedia Syst.* 22, 2 (March 2016), 213-227. <http://dx.doi.org/10.1007/s00530-014-0446-1>
- [42] Jörg Tiedemann. 2008. Synchronizing translated movie subtitles. In *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC '08)*, 1902-1906. [http://www.lrec-conf.org/proceedings/lrec2008/pdf/484\\_paper.pdf](http://www.lrec-conf.org/proceedings/lrec2008/pdf/484_paper.pdf)
- [43] Elaine A. Trindade. 2012. *A legendagem da televisão por assinatura do Brasil*. Master's dissertation. USP, SP, Brazil. <https://doi.org/10.11606/D.8.2012.tde-13092012-114245>
- [44] Lucas C. Villa Real, Rodrigo L. Guimarães, and Priscilla Avegliano. 2015. Dynamic Adjustment of Subtitles Using Audio Fingerprints. In *Proc. of the 23rd ACM int. conference on Multimedia (MM '15)*, 975-978. <https://doi.org/10.1145/2733373.2806378>
- [45] Chen Wang, Erik N. Geelhoed, Phil P. Stenton, and Pablo Cesar. 2014. Sensing a live audience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1909-1912. <https://doi.org/10.1145/2556288.2557154>