

Análise da segmentação linguística nas legendas para surdos e ensurdecidos (LSE) do filme *Virada Radical*: estudo baseado em corpus

Analysis of linguistic segmentation in the subtitles for the deaf and the hard-of-hearing (SDH) of the movie *Virada Radical*: a corpus-based study

Vera Lúcia Santiago Araújo*

Daniel de Albuquerque e Arraes**

Resumo: Para inserir-se na sociedade, a comunidade surda depende de programação audiovisual com legendas que atendam suas necessidades, mas quais os critérios para medir a qualidade destas legendas? Segundo apontaram pesquisas feitas pelo grupo *Legendagem e Audiodescrição* (LEAD) (ARAÚJO; NASCIMENTO, 2011; ARAÚJO; VIEIRA; MONTEIRO, 2013), um fator importante é a qualidade da *segmentação* (divisão entre inserções de legenda ou quebra de linha), especialmente a segmentação linguística, feita segundo a sintaxe, tal que o respeito à estrutura dos sintagmas nas quebras de linhas parece facilitar a compreensão dos programas pelos surdos. Este artigo descreve os problemas de segmentação linguística das legendas do filme *Virada*

* Professora do Programa de Pós-graduação em Linguística Aplicada (PosLA) da Universidade Estadual do Ceará (UECE). E-mail: verainnerlight@uol.com.br

** Mestrando do Programa de Pós-graduação em Linguística Aplicada (PosLA) da Universidade Estadual do Ceará (UECE). E-mail: arraes.d@gmail.com

Radical, propondo ressegmentações. Das 1074 legendas coletadas, 5,4% apresentaram problemas, principalmente nos sintagmas verbais e nominais.

Palavras-chave: Tradução audiovisual; Acessibilidade; Legendagem para surdos e ensurdecidos; Estudo baseado em corpus.

Abstract: In order to insert itself in society, the deaf community depends on audiovisual programs with subtitles/captions that can attend to their needs, but what are the criteria to measure its quality? Previous research done by the group *Legendagem e Audiodescrição* (LEAD) (ARAÚJO; NASCIMENTO, 2011; ARAÚJO; VIEIRA; MONTEIRO, 2013), suggest that an important factor is the quality of the *segmentation* (division between subtitles or line break), specially linguistic segmentation, made according to the syntax, so much so that respecting phrase structure at line breaks could facilitate comprehension of programs by the deaf. This paper describes the linguistic segmentation problems of the subtitles/captions of the movie *Virada Radical*, proposing resegmentations. Out of the 1074 collected subtitles, 5.4% presented problems especially within noun phrases and verb phrases.

Keywords: Audiovisual translation; Accessibility; Subtitling for the deaf and the hard-of-hearing; Corpus-based study.

1. Introdução

No século XX, o desenvolvimento de recursos audiovisuais como rádio e televisão possibilitou às massas acesso àqueles bens culturais antes disponíveis apenas à parcela letrada das diferentes sociedades. Apesar disso, pessoas com deficiência visual e auditiva continuam impossibilitadas de usufruir plenamente dos meios de informação e entretenimento disponíveis em diferentes mídias.

Todavia, tem crescido, mesmo que lentamente, a conscientização da sociedade com relação às necessidades dessas minorias. Um dos frutos importantes disso foi a promulgação, pelo Ministério das Comunicações, da Portaria 310 de 27 de junho de 2006, a qual estabeleceu um prazo de onze anos, a partir de sua publicação, para que os meios de radiodifusão brasileiros incorporassem meios de acessibilidade em sua programação. No caso da comunidade surda, a acessibilidade se daria através da incorporação de legendas na programação televisiva.

Entretanto, a exigência legal de meios de acessibilidade não garante a qualidade dos mesmos. Pesquisas de recepção de legendas desenvolvidas pelo grupo Legendagem e Audiodescrição (LEAD) da Universidade Estadual do Ceará (UECE) nos últimos 15 anos, tanto com surdos cearenses (e.g.: FRANCO; ARAÚJO, 2003; ARAÚJO, 2004, 2008), quanto com surdos das diversas regiões do país (e.g.: ARAÚJO; NASCIMENTO, 2011; ARAÚJO, 2012; ARAÚJO; VIEIRA; MONTEIRO, 2013), permitiram, em primeiro lugar, constatar as dificuldades dos surdos em lidar com as legendas de televisão brasileiras - haja vista não estarem de acordo com nenhum dos critérios preconizados pelos pesquisadores de legendagem¹. Ademais, tais pesquisas permitiram também refinar o entendimento sobre as dificuldades supracitadas.

Um primeiro fator a se considerar numa legenda é a sua *velocidade*, que deve respeitar o “tempo que o expectador médio leva para ler e assimilar a informação” (DIAZ-CINTAS; REMAEL, 2007: 23), mas os canais de televisão, ao invés disso, praticam uma transcrição quase completa das falas da programação, sobrecarregando as legendas com informação.

Nas primeiras pesquisas de recepção com surdos cearenses ([estudo 1] FRANCO; ARAÚJO, 2003; [estudo 2] ARAÚJO, 2004), os informantes declararam achar as legendas muito rápidas, levando à conclusão de que a melhor solução seria uma maior condensação e edição das legendas. Uma pesquisa posterior ([estudo 3] ARAÚJO, 2008) testou, também com surdos cearenses, várias configurações diferentes de legendas em encontros mensais ao longo de 8 meses, coletando as sugestões dos informantes até que se chegou a um modelo consistente de legendagem. Finalmente, era preciso testar este modelo de legendagem com uma amostra maior de surdos. Realizou-se uma nova pesquisa de recepção ([estudo 4] ARAÚJO, 2012) com surdos de várias capitais brasileiras.

¹ Tradicionalmente, o processo de legendar um filme tinha duas etapas: a tradução do texto verbal, chamada *legendação*, e a inserção das legendas no filme, chamada *legendagem*. As duas etapas eram laboriosas e demandavam dois profissionais diferentes, respectivamente, legendista e legendador. Porém, com o advento do computador pessoal e de softwares de legendagem, uma única pessoa pode realizar as duas funções. Nesse sentido, entendemos que a diferenciação não se justifique mais, daí adotarmos neste artigo apenas a terminologia legendagem. Para maiores detalhes, vide Franco e Araújo (2011).

Os resultados confirmaram a eficácia do modelo proposto a partir das sugestões dos surdos cearenses. Porém, os resultados da pesquisa (ARAÚJO, 2012) apontaram também para a importância de outro fator além da velocidade, a *segmentação* (divisão das falas na legenda). Os dados sugerem que uma boa segmentação pode permitir aos surdos uma boa recepção de legendas, mesmo que elas estejam rápidas.

Assim, formulou-se uma nova hipótese: legendas rápidas podem ser lidas confortavelmente, desde que estejam bem segmentadas. Porém, a literatura dos Estudos da Legendagem carecia de estudos empíricos descrevendo o que caracterizaria uma segmentação problemática num dado sistema linguístico (no nosso caso, o português brasileiro), dificultando a manipulação desta variável em futuras pesquisas de recepção.

Para suprir essa lacuna teórica, foi idealizado o projeto “A segmentação na legendagem para surdos e ensurdecidos (LSE): um estudo baseado em corpus”, abreviado como CORSEL (Corpus, Segmentação e Legendagem), que coletou legendas de diferentes programas da Rede Globo - abrangendo gêneros como novela, documentário e comédia de situação (*sitcom*) - durante o triênio 2012-2014, com o objetivo de identificar legendas com segmentação problemática, classificar os tipos de problemas e propor soluções para os problemas (ressegmentação), visando construir um banco de dados que possa, entre outras aplicações, auxiliar na formação de futuros legendistas.

O presente artigo apresenta os resultados de uma pesquisa conduzida paralelamente ao projeto CORSEL, oriundos do Trabalho de Conclusão de Curso do coautor e desenvolvido no ano de 2014 como requisito parcial para obtenção do grau de Bacharel em Letras (língua inglesa e respectiva literatura) pela Universidade Estadual do Ceará (ARRAES, 2015). Foram analisadas as legendas do filme *Virada Radical*, de 2006, exibido pela Rede Globo em 04 de junho de 2014 na Sessão da Tarde. A relevância da pesquisa deve-se ao fato de analisar um gênero audiovisual não contemplado pelo projeto CORSEL; filmes exibidos em TV.

Os objetivos do estudo aqui relatado foram: (i) analisar e descrever os problemas de segmentação nas legendas do filme *Virada Radical*, tal como

produzidas pela Rede Globo; (ii) aferir, quanto a dois parâmetros técnicos da legendagem, velocidade e número de linhas, em quais condições estes problemas acontecem; (iii) comparar os resultados da presente pesquisa com aqueles obtidos pelo projeto CORSEL. A coleta dos dados replicou os procedimentos adotados pelo projeto CORSEL para o gênero documentário: gravou-se a tela de uma TV exibindo o filme com as legendas ativadas, as quais foram transcritas manualmente usando o programa *Subtitle Workshop 2.51*; em seguida, houve a anotação do corpus por etiquetagem; finalmente, deu-se a análise eletrônica via software *Wordsmith Tools 5.0*.

Além desta introdução, este artigo é composto por uma seção de bases teóricas, que define a legendagem, caracteriza suas diferentes formas e discorre sobre os parâmetros responsáveis pela boa integração imagem-som-legendagem; uma seção de metodologia, que detalha a tipologia da pesquisa e os passos seguidos até o momento da análise; uma seção de apresentação dos resultados e sua discussão à luz da Tradução Audiovisual; e uma seção de considerações finais.

2. Bases teóricas

2.1. Legendagem: definição e características

A legendagem é uma modalidade de tradução audiovisual surgida na primeira metade do século XX concomitantemente ao advento do cinema falado. A disseminação internacional de grandes produções cinematográficas criou uma grande demanda pela tradução dos diálogos da língua de partida para os idiomas dos diversos países consumidores. Logo:

[...] legendagem consiste em representar, através da escrita, a tradução, para uma língua alvo, da troca original de diálogos enunciados pelos diferentes falantes, bem como de todas as outras informações verbais transmitidas visualmente [...] ou pela audição [...] (DIAZ-CINTAS, 2010: 344).

Enquanto objeto de estudo científico, a legendagem está inserida na disciplina da Tradução Audiovisual (TAV), que é por sua vez parte integrante da área maior denominada Estudos da Tradução.

É difícil estabelecer uma tipologia de legendagem, haja vista que novas formas de produzi-las surgem a cada inovação tecnológica, mas em linhas gerais podem-se categorizar as formas de legendagem quanto a critérios de exibição e linguísticos.

De um ponto de vista da exibição, uma legenda pode ser *aberta*, quando é fixada às imagens e não pode ser desligada, como numa sessão de cinema, ou *fechada*, quando é opcional e pode ser ativada ao bel prazer do expectador, como no caso de DVDs. Já de uma perspectiva linguística, distinguem-se legendas *interlinguísticas*, quando se traduz o discurso falado da língua de partida para o texto escrito na língua de chegada, de legendas *intralinguísticas*, quando o idioma do discurso falado coincide com o do texto escrito na legenda (DIAZ-CINTAS, 2010).

A aplicação mais comum das legendas intralinguísticas é aquela dirigida ao público com algum tipo de deficiência auditiva, denominada *legendagem para surdos ensurdecidos* (LSE) (DIAZ-CINTAS, 2010), embora também possa haver LSE interlinguística, por exemplo, em festivais de filmes estrangeiros com sessões especiais para surdos e ensurdecidos. A principal diferença entre a LSE e a legendagem para ouvintes é o acréscimo de informações que de outra forma seriam acessíveis apenas pelo canal auditivo, como efeitos sonoros e identificação dos falantes.

No Brasil, os canais de televisão utilizam o sistema estadunidense *closed caption* para legendar sua programação, transformando as legendas em códigos eletrônicos inseridos na linha 21 do intervalo vertical em branco da TV (ARAÚJO, 2008). As legendas *closed caption* podem ser apresentadas na tela de duas maneiras: surgindo num bloco único, permanecendo alguns segundos e desaparecendo, semelhante à legenda de cinema (legenda *pop-on*); ou surgindo em blocos de palavras digitados da esquerda para a direita que, ao completarem uma linha, deslocam-se para a linha de cima e permanecem na tela até que a linha inferior seja completada (legenda *roll-up*). O primeiro tipo é mais utilizado para programas pré-gravados, enquanto o segundo é mais

comum em programas ao vivo. Ademais, as legendas *closed caption* podem ser produzidas através de duas tecnologias: a *estenotipia*, que utiliza um teclado fonético chamado estenótipo, comumente usado na transcrição de julgamentos, devido à maior rapidez de digitação; e o *refalamento* (do inglês, *respeaking*), que consiste de uma tecnologia de reconhecimento de voz através da qual um profissional repete o que está sendo dito no programa de TV para um *software*, que, então, transforma a fala do profissional no texto das legendas (ARAÚJO; ASSIS, 2014).

É questionável se as legendas *closed caption* praticadas no Brasil oferecem realmente acessibilidade à comunidade de surdos e ensurdecidos, haja vista os resultados das diversas pesquisas de recepção de legendas descritas anteriormente.

A acessibilidade, de maneira geral, significa garantir que qualquer pessoa possa dispor e utilizar-se, com conforto e segurança, de espaços, serviços e produtos, tanto físicos quanto virtuais, presencialmente ou não, em condição de igualdade aos demais (ALONSO, 2007:17, apud MASCARENHAS, 2012:15).

Para produzir legendas acessíveis, tanto para ouvintes quanto para surdos e ensurdecidos, é necessário primar pela plena integração da legenda ao produto audiovisual, que por sua vez resulta da convergência de um conjunto de fatores, denominados parâmetros técnicos e linguísticos de legendagem.

2.2. Parâmetros técnicos e linguísticos da legendagem

Existem duas grandes classes de parâmetros da legendagem, os técnicos e os linguísticos. Os parâmetros técnicos respondem pelas coerções espaciais e temporais impostas por produtos audiovisuais. Eles são: (i) espaciais, tais como posição da legenda na tela, número de linhas por

inserção de legenda, número de caracteres por linha; e (ii) temporais, como duração mínima e máxima de uma inserção de legenda, intervalo de tempo entre o início da fala e o aparecimento da legenda, intervalo de tempo entre inserções de legenda (KARAMITROGLOU, 1998; DIAZ-CINTAS; REMAEL, 2007).

Já os parâmetros linguísticos consistem de: (i) convenções de pontuação (como o uso de reticências no início ou fim da legenda, o uso de parênteses, colchetes, aspas etc.); e (ii) convenções de edição do texto (segmentação, quebra de linhas, alteração de estruturas sintáticas etc.).

Dentre os parâmetros técnicos, o primeiro que deve ser aqui discutido é o da velocidade. Para alcançar uma boa integração imagem-som-legenda, um dos critérios vitais a ser seguido pelos profissionais de legendagem é a chamada *Regra dos seis segundos*, “que se refere ao tempo que o expectador médio leva para ler e assimilar a informação contida nas [...] linhas de uma legenda” (DIAZ-CINTAS; REMAEL, 2007: 23). Se houver informação em excesso numa legenda (uma legenda rápida), o expectador não terá tempo hábil de lê-la, provocando desconforto e dificuldade de compreensão.

A regra dos seis segundos prevê três velocidades mais propícias para a boa integração de legenda com imagem e som: 145, 160 e 180 palavras por minuto (ppm). A maior eficiência destas três velocidades na recepção de legendas foi confirmada através de pesquisas experimentais desenvolvidas por psicólogos belgas (D'YDEWALLE et al., 1987) via rastreamento ocular.

A velocidade de uma legenda surge da divisão da quantidade de palavras pelo tempo de duração. Visando permitir um controle mais fino dessa relação, Diaz-Cintas e Remael (2007) propõem a conversão da unidade de medida ‘palavras por minuto’ para ‘caracteres por segundo’ (cps), conforme mostra a Tabela 1.

Tabela 1 - Máximo de caracteres para uma velocidade de 145 ppm

Segundos : frames ²	Nº de caracteres	Segundos : frames	Nº de caracteres	Segundos : frames	Nº de caracteres
-----------------------------------	---------------------	----------------------	---------------------	----------------------	---------------------

² A tabela mostra a unidade de conversão em “segundos:frames” considerando o padrão para filmes longa metragem, nos quais 1 segundo de filme tem 24 quadros. Diferentes arquivos de vídeo podem apresentar taxas diferentes de frames por segundo. Portanto, as frações de segundo poderão equivaler a um número diferente de frames em cada caso.

1:00	16	2:00	29	3:00	44
1:04	17	2:04	32	3:04	46
1:08	18	2:08	34	3:08	48
1:12	20	2:12	36	3:12	50
1:16	23	2:16	38	3:16	52
1:20	25	2:20	40	3:20	54

Fonte: adaptado de Diaz-Cintas e Remael (2007)

Com relação ao papel da velocidade na LSE, as três primeiras pesquisas de recepção realizadas na UECE ([estudo 1] FRANCO; ARAÚJO, 2003; [estudo 2] ARAÚJO, 2004; [estudo 3] ARAÚJO, 2008) apontaram para a velocidade como principal fator a interferir na recepção das legendas *closed caption* pelos surdos cearenses. Por isso, o modelo de legendagem desenvolvido a partir destes estudos tinha como hipótese que a velocidade de 145 ppm era a mais apropriada para o público de surdos e ensurdecidos.

Porém, a validação desta proposta demandava testes de recepção com um número maior de informantes surdos. Uma pesquisa de nível nacional intitulada *Em busca de um modelo de legendagem para surdos para o Brasil* (projeto MOLES³) testou o modelo proposto com as sugestões dos surdos cearenses nas três velocidades preconizadas pela regra dos seis segundos, partindo do seguinte pressuposto: quanto mais rápida a legenda, menor será a compreensão do expectador. Participaram da pesquisa 34 surdos de 4 regiões⁴ do Brasil.

Curiosamente, os resultados (ARAÚJO; NASCIMENTO, 2011; ARAÚJO, 2012; ARAÚJO; VIEIRA; MONTEIRO, 2013) mostraram um bom índice de recepção de legendas entre os informantes surdos, mesmo com legendas em velocidade alta (180 ppm), refutando o pressuposto mencionado acima. Em face disso, foi necessário considerar, além da regra dos seis segundo, outro parâmetro de legendagem: a *segmentação*.

³ Sigla de: Modelo de Legendagem para Surdos (vide ARAÚJO, 2012).

⁴ Os dados foram coletados em sete capitais brasileiras: Porto Alegre e Santa Catarina, na região Sul; Rio de Janeiro e Belo Horizonte, no Sudeste; Salvador e Fortaleza, no Nordeste; e Rio Branco, no Norte. Na região centro-oeste, não foram encontrados sujeitos dispostos a participar da pesquisa.

O parâmetro da segmentação determina a integração da legenda a um produto audiovisual respeitando (i) os cortes de cena (segmentação visual), (ii) as pausas respiratórias dos falantes (segmentação retórica) e (iii) a estrutura linguística, tanto sintática quanto semântica, do texto da legenda (segmentação linguística). Dos três tipos de segmentação, o projeto CORSEL ateu-se à linguística, visto que os outros dois tipos não ocorreram em quantidade significativa em outras pesquisas (CHAVES, 2012; CHAVES; ARAÚJO, 2014).

O exame das legendas usadas nas pesquisas de recepção de nível nacional ([estudo 4] ARAÚJO, 2012) revelou que elas respeitavam o critério da segmentação (haja vista terem sido confeccionadas pelos pesquisadores do grupo LEAD), o que poderia ter influenciado na recepção das legendas.

A segmentação linguística pode acontecer ora entre duas inserções de legenda, ora numa mesma inserção (denominada quebra de linha ou *line break*). Segundo Diaz-Cintas e Remael (2007: 72), “uma segmentação cuidadosa da informação pode ajudar a reforçar a coerência e a coesão na legendagem”. Segundo Karamitroglou (1998), isto acontece, pois:

Quando segmentamos uma sentença, forçamos o cérebro a pausar seu processo linguístico por um instante, até os olhos localizarem a próxima informação linguística. Nos casos em que a segmentação é inevitável, devemos tentar levar o cérebro a fazer essa pausa num ponto onde a carga semântica se configure numa informação completa. (KARAMITROGLOU, 1998: 9-10)

Aproveitando-se do ferramental analítico da Linguística Gerativa, que utiliza esquemas arbóreos para analisar hierarquicamente as relações sintáticas entre os constituintes de uma oração, Karamitroglou (1998) postula que o texto de uma legenda deve ser segmentado nos mais altos níveis sintáticos.

Para esta pesquisa, porém, a opção pelo gerativismo dificulta a análise do texto das legendas *closed caption*, cuja transcrição quase literal da fala preserva marcas de oralidade do idioma pátrio, não se permitindo analisar facilmente por descrições formalistas. Por isso, optou-se, no projeto CORSEL, por desenvolver as categorias de problemas de segmentação tomando por

base a premissa de Karamitoglou (1998), porém usando como aporte gramatical uma descrição funcionalista do português brasileiro falado, baseada em dados empíricos do uso real dos falantes: a *Nova Gramática do Português Brasileiro* (CASTILHO, 2012).

Ademais, o estudo do parâmetro da segmentação foi feito através dos métodos da Linguística de Corpus⁵ pelo uso de etiquetas, tal como desenvolvidas inicialmente pelos trabalhos de Chaves (2012), Assis (2013) e Gabriel (2013), e depois aperfeiçoadas pelo projeto CORSEL (ARAÚJO, 2015; ASSIS, 2016; GABRIEL, 2015; VIEIRA; ARAÚJO; CHAVES; MONTEIRO, 2015; FRANCO, CHAVES, 2015). Maiores detalhes a seguir.

3. Metodologia

3.1. Tipo da pesquisa

Trata-se de uma pesquisa descritiva, pautada por análises quali-quantitativas apoiadas teórica e metodologicamente na Tradução Audiovisual e na Linguística de Corpus. Ademais, quanto à abordagem, é uma pesquisa dedutiva (do geral para o específico), haja vista que as categorias de problemas de segmentação foram pensadas expandindo a premissa mais geral de Karamitoglou (1998) com o aporte da *Nova Gramática do Português Brasileiro* (CASTILHO, 2012).

3.2. Corpus

⁵ Ramo da linguística dedicado à análise eletrônica das línguas naturais através da coleta de um corpus ou de vários corpora. Corpora são conjuntos de dados linguísticos textuais, criteriosamente armazenados em forma digital e passíveis de leitura por programas de computador (BERBER SARDINHA, 2004).

É composto pelas legendas *closed caption* do filme *Virada Radical* (2006) (título original, *Stick It*), exibido pela Rede Globo no dia 04 de junho de 2014 no programa Sessão da Tarde. O filme é uma produção teuto-americana de 2006 dirigida por Jessica Bendinger e conta a história de superação de Haley Graham (Missy Peregrym), uma ex-ginasta problemática que, pega em flagrante pela polícia cometendo vandalismo, deve cumprir pena alternativa retornando ao mundo da ginástica olímpica para reparação de danos. Seu treinador é Burt Vickerman (Jeff Bridges), um exigente técnico dono de uma academia de ginástica olímpica, onde Haley deverá cumprir sua pena.

O corpus consiste de 8.636 palavras corridas (tokens)⁶ distribuídas em 1704 inserções de legendas do tipo *pop-on* e é classificado como um corpus pequeno⁷. Apesar do tamanho diminuto, o corpus pode ser considerado representativo, na medida em que não se está descrevendo um fenômeno propriamente linguístico (como é praxe na Linguística de Corpus), mas um fenômeno de legendagem. Ademais, não há um critério objetivo para determinar um tamanho ideal para um corpus representativo, pois este é uma amostra de um fenômeno cuja extensão é desconhecida.

3.3. Procedimentos

Primeiro, filmou-se a tela de uma televisão durante a exibição do filme legendado. Então, manualmente, as legendas foram transcritas e sincronizadas de acordo com o tempo em que foram exibidas utilizando o programa *Subtitle Workshop 2.51*, software livre de legendagem usado para marcar tempo de vídeo (o início e o final de cada legenda), para revisar e para assistir a versão prévia do vídeo legendado. Este programa oferece ao legendista diversas extensões de arquivo para salvar as legendas, mas o

⁶ O total de palavras num texto, mesmo que repetidas.

⁷ Para Berber Sardinha (2002 *apud* BERBER SARDINHA, 2004, p. 26), um corpus é considerado pequeno quando o total de palavras for menor ou igual a 80 mil palavras.

formato aqui adotado foi o mesmo utilizado pelo projeto CORSEL - a extensão *srt* - pela facilidade de ser transformado em arquivo *txt*, o padrão utilizado pela ferramenta de Linguística de Corpus aqui adotada: o *Wordsmith Tools 5.0*.

O arquivo *txt* gerado armazena as inserções de legenda conforme exemplificado no Quadro 1.

Quadro 1 - exemplo de legenda em arquivo *txt*

```
3
00:00:44,036 --> 00:00:45,952
[Poot]
Vai Cracker!
```

Fonte: os autores

Na primeira linha, a legenda é numerada em ordem crescente (no exemplo, a legenda será a 3ª a aparecer); a segunda linha contém os tempos inicial (quando a legenda surge na tela) e final (quando a legenda desaparece); nas duas linhas inferiores está o texto que irá aparecer na legenda (no exemplo, a 3ª linha identifica entre colchetes qual personagem toma a palavra, enquanto a 4ª linha registra sua fala).

Depois de convertidas para *txt*, as legendas sofreram etiquetagem, processo de acrescentar marcações ao corpus para sinalizar informações desejadas. Dentre inúmeras convenções de etiquetagem, adotou-se o padrão *SGML (Standard Generalized Mark-up Language)*, cujos códigos escritos entre parênteses angulares - <etiquetas de abertura> informação </etiqueta de fechamento> - caracterizam e delimitam o conteúdo a ser analisado. São dois os grupos de etiquetas: (i) as de parâmetros técnicos da legendagem e (ii) as de problemas de segmentação linguística. O Quadro 2, lista as etiquetas existentes para os parâmetros técnicos e suas respectivas funções.

Quadro 2 - etiquetas para parâmetros técnicos da legendagem

Número da legenda	<sub1>legenda1</sub1>
Linhas por legenda	<1L>, <2L> e <3L>
Tempo inicial e final de cada legenda	<t>início --> final</t>
Número de caracteres por linha	<cpl>
Velocidade de legenda baixa (145ppm)	<veloc_leg_baixa>

Velocidade de legenda média (160ppm)	<veloc_leg_média>
Velocidade de legenda alta (180ppm)	<veloc_leg_alta>

Fonte: Araújo, 2015

A etiquetagem manual do corpus foi feita no programa *Bloco de Notas*, cujo resultado está exemplificado no Quadro 3.

Quadro 3 - exemplo de legenda etiquetada

```
<sub3> <2L> 3
<t>00:00:44,036 --> 00:00:45,952</t>
<cpl6>[Poot]
<cpl12>Vai Cracker!
<veloc_leg_baixa 18c/1,9s></sub3>
```

Fonte: os autores

O Quadro 3 mostra, na 1ª linha, o acréscimo de uma etiqueta contendo o número de linhas do texto legenda, <2L>, bem como uma etiqueta usada para facilitar o uso da ferramenta de busca do Bloco de Notas, <sub3>. Na 2ª linha, o tempo inicial e final da legenda está cercado pelas etiquetas <t> e </t>. À 3ª e 4ª linhas, são adicionadas etiquetas <cpl>, que registram quantos caracteres cada linha tem. Na última linha, há a etiqueta sinalizadora da velocidade da legenda, indicando a divisão entre o número de caracteres (18 caracteres) e a duração da legenda (1,9 segundos).

Além das etiquetas relativas aos parâmetros técnicos, também foram utilizadas etiquetas concernentes aos problemas de segmentação, desenvolvidas a partir do aporte da *Nova Gramática do Português Brasileiro*, de Castilho (2012), como mostra o Quadro 4.

Quadro 4 - Etiquetas de problemas de segmentação linguística

ETIQUETA INDICATIVA DE PROBLEMA DE SEGMENTAÇÃO		
<PROSEGL>		
ETIQUETAS P/ SINT. VERBAL (SV)	ETIQUETAS P/ SINT. NOMINAL (SN)	ETIQUETAS P/ SINT. PREPOSICIONAL (SP)
<SV_composto> <SV_verbo+SAdv> <SV_(verbo)+oblíquo+SV>	<SN_especif+SN> <SN_subst+SAdj> <SN_núcleo>	<SP_prep+SN> <SP_prep+SV>
ETIQUETAS P/ SINT. ADJETIVAL (SAdj)	ETIQUETAS P/ SINT. ADVERBIAL (SAdv)	ETIQUETAS P/ PERÍODOS COMPOSTOS

<SAdj_especif+SAdj>	<SAdv_adv+SAdv>	<COORD_conj+oração> <SUBORD_conj/pron_rel+oração>
---------------------	-----------------	------------------------------------------------------

Fonte: Araújo, 2015.

Por fim, o corpus já etiquetado foi analisado através de duas ferramentas do programa *Wordsmith Tools 5.0*: o *Wordlist*, que informou as entradas lexicais (*types*) e a ocorrência destas entradas (*tokens*); o *Concord*, que permitiu a análise dos problemas de segmentação em contexto através das linhas de concordância.

4. Resultados

Primeiramente, há incidência baixa de problemas de segmentação (etiqueta <PROSEGL>) em relação ao total de legendas do corpus: das 1704 inserções de legendas, apenas 92 (5,4%) apresentaram problemas de segmentação (vide Tabela 2). Um índice muito baixo se comparados com o dos outros gêneros analisados no projeto CORSEL: 15,5% no gênero novela, 10,6% no gênero seriado comédia de situação e 39,3% no gênero documentário (ARAÚJO, 2015).

Tabela 2 - Ocorrência de problemas de segmentação em *Virada Radical*

Nº de legendas	PROSEGL	
	Valor absoluto	%
1074	92	5,4

Fonte: os autores

A Tabela 3 apresenta, em ordem decrescente da esquerda para a direita, a ocorrência dos problemas de segmentação a partir das sete categorias maiores - os cinco tipos de sintagma do português brasileiro e os períodos compostos por coordenação e subordinação -, tanto em valores absolutos quanto em percentual.

Tabela 3 - Valores absolutos (VA) e valor percentual (%) dos tipos de PROSEGL em *Virada Radical*

	SV	SN	SUBORD	SAdj	SP	SAdv	COORD	TOTAL
VA	65	13	10	2	2	0	0	92
(%)	70,6	14,2	10,8	2,2	2,2	0	0	100

Fonte: os autores

Além disso, embora o presente corpus compartilhe com os corpora do projeto CORSEL uma predominância de problemas de segmentação nos sintagmas verbal e nominal, ocorreu uma diferença significativa entre os resultados aqui relatados e os do CORSEL.

Em dois gêneros analisados pelo projeto CORSEL, novela e seriado, houve maior proximidade entre a quantidade de problemas de segmentação nos sintagmas verbal e nominal: no gênero novela houve 44,3% de problemas de segmentação no sintagma verbal e 26,8% no sintagma nominal (diferença de 17,5%); no gênero seriado, 47,3% de problemas no sintagma verbal e 27,8% no sintagma nominal (diferença de 19,5%) (ARAÚJO, 2015). Na segmentação de *Virada Radical*, a diferença na ocorrência dos dois sintagmas foi bem maior, 56,4%: os problemas de segmentação no sintagma verbal totalizaram 70,6%, enquanto apenas 14,2% dos problemas de segmentação ocorreram no sintagma nominal.

No gênero documentário do projeto CORSEL, diferentemente das novelas e seriados, os sintagmas nominal e preposicional tiveram maior ocorrência de problemas de segmentação: respectivamente, 40,2% e 27,5% (ARAÚJO, 2015). Estes dados parecem apontar para a influência do gênero audiovisual nos tipos de problemas de segmentação mais frequentes, porém, seriam necessárias mais pesquisas para confirmar esta suposição.

Cabe agora pormenorizar a análise qualitativamente dentro de cada uma das sete categorias mencionadas na Tabela 3. As análises estão ordenadas em ordem decrescente de ocorrências.

4.1. Sintagma Verbal

A categoria com a maior incidência de problemas de segmentação foi a dos sintagmas verbais. Dos 65 casos de problemas nesta categoria, a grande maioria (47 casos) foram do tipo <SV_composto>, no qual a quebra de linha acontece entre dois verbos, o auxiliar e o pleno, de um sintagma verbal.

Em segundo lugar de ocorrências (13 casos) estão as do tipo <SV_verbo+SAdv>, cuja quebra de linha interfere na relação predicativa entre um advérbio e um verbo.

Finalmente, sobre o último tipo de problema de segmentação no sintagma verbal, ocorreram apenas 5 casos de <SV_(verbo)+oblíquo+SV>, etiqueta usada nas quebras de linha que separam um verbo de um pronome clítico (oblíquo).

Como proposta de ressegmentação de uma legenda com quebra de linha problemática nesta categoria, utilizaremos como exemplo a legenda 1524:

Quadro 5 - Exemplo de problema de segmentação <SV_composto>

Legenda	Nº de linhas	Duração	Velocidade	Texto
Sub1524	2	4,8s	145ppm (baixa)	Vocês vão me dizer quem vai competir e quem não vai?

Fonte: os autores

No Quadro 5, há um exemplo de problema de segmentação comumente encontrado em legendas para TV: ao transcrever tanto a dublagem de filmes, quanto as falas de novelas e seriados - ambos marcados por forte oralidade - os sintagmas verbais em legendas frequentemente vêm cindidos em dois constituintes (perífrase verbal), tal que na legenda 1524, o conteúdo morfossintático (numero-pessoal e modo-temporal) resta sinalizado pelo verbo auxiliar [vai], enquanto ao verbo principal [competir] cabe apenas indicar o conteúdo lexical do sintagma, permanecendo na forma infinitiva.

O texto da legenda totaliza 51 caracteres, lidos tranquilamente graças à duração da legenda; 4,8 segundos. Logo, o foco da ressegmentação deve ser não ultrapassar o máximo de caracteres por linha, como vemos abaixo.

VOCÊS VÃO ME DIZER QUEM VAI COMPETIR E QUEM NÃO VAI?

Nesta proposta de ressegmentação, optou-se por segmentar a legenda no mais alto nível sintático, realocando os constituintes [quem vai] para a linha inferior, produzindo uma legenda com 18 caracteres na linha superior e 31 na linha inferior.

4.2. Sintagma nominal

A segunda categoria com maior incidência de problemas de segmentação é a dos sintagmas nominais. Dos 13 casos de quebra no sintagma nominal, a maioria (8 casos) foram do tipo <SN_subst+SAdj>, quando a quebra de linha separa o núcleo nominal de um sintagma adjetival posposto.

Prosseguindo a análise, no segundo lugar (5 casos) de problemas de segmentação no sintagma nominal está <SN_especif+SN>, no qual há a separação entre o núcleo nominal e os constituintes a ele antepostos.

Finalmente não ocorreu nenhum caso de <SN_núcleo>, quando a quebra de linha desloca parte do núcleo nominal, como no caso de nomes próprios.

Como proposta de ressegmentação de uma legenda com quebra de linha problemática nesta categoria, utilizaremos como exemplo a legenda 1234:

Quadro 6 - Exemplo de problema de segmentação < SN_subst+SAdj>

Legenda	Nº de linhas	Duração	Velocidade	Texto
Sub 1234	2	2,2	180ppm (alta)	Já deve ter alguma ginasta machucada por aí

Fonte: os autores

O Quadro 6 destaca a legenda 1234, cujo problema de segmentação demanda cuidado na resolução, pois além do adjetivo [machucada], há anteposto ao núcleo nominal [ginasta] o pronome indefinido [alguma], tal que as três palavras deveriam permanecer juntas após a ressegmentação. Todavia, uma legenda de 2,2 segundos deveria ter 34 caracteres, bem abaixo dos 42 caracteres da legenda, demandando a condensação da legenda e impedindo preservar intacto o sintagma [alguma ginasta machucada]. Uma possibilidade de retificar a segmentação seria:

JÁ DEVE TER
GINASTA MACHUCADA POR AÍ.

A supressão do pronome indefinido reduz bastante o texto da legenda, mas ela ainda permanece com um número de caracteres ligeiramente acima do recomendado (36 caracteres), problema que pode ser sanado pelo rearranjo do tempo de duração desta legenda e da anterior, a legenda 1233. A legenda 1233 contém um texto de 16 caracteres - [Ah, o Vickerman.] - exibidos em 1,5 segundos, duração esta que pode ser diminuída em 0,5 segundos sem alterar a velocidade da legenda (145ppm). Este saldo de meio segundo pode então ser repassado para a legenda 1234, a qual ficaria com 2,7 segundos, mais que o necessário para torná-la uma legenda com velocidade de 145ppm.

4.3. Oração subordinada

A categoria de problemas de segmentação em orações subordinadas figura em terceiro lugar, com 10 ocorrências. Nesta categoria existe apenas uma etiqueta, <SUBORD_conj/pron_rel+oração>, para indicar legendas cuja quebra de linha separa uma conjunção ou pronome relativo da oração posposta, como mostra o exemplo do Quadro 7:

Quadro 7 - Exemplo de problema de segmentação
<SUBORD_conj/pron_rel+oração>

Legenda	Nº de linhas	Duração	Velocidade	Texto
Sub 166	2	3s	145ppm (baixa)	Não é à toa que ninguém descobriu você.

Fonte: os autores

A duração de 3 segundos da legenda 166 é suficiente para exibir os 38 caracteres do texto da legenda, restringindo o foco da ressegmentação ao máximo de caracteres e às questões sintáticas, como sugerido abaixo.

NÃO É À TOA
QUE NINGUÉM DESCOBRIU VOCÊ.

A realocação do trecho [que ninguém] para a linha inferior tanto respeitaria a premissa de realizar a quebra de linha no mais alto nível sintático, quanto permitiria respeitar o número máximo de caracteres por linha (11 caracteres na linha superior e 27 na inferior).

4.4. Sintagma adjetival

Foi escasso número de problemas de segmentação no sintagma adjetival (2 casos), categoria com uma única etiqueta, indicadora da cisão entre núcleo adjetival e o advérbio à esquerda, como ilustra o Quadro 8:

Quadro 8 - Exemplo de problema de segmentação <SAdj_especif+SAdj>

Legenda	Nº de linhas	Duração	Velocidade	Texto
Sub 354	2	1,4s	180ppm (alta)	Porque eu já estou meio cansada de adultos loucos.

Fonte: os autores

Segundo Diaz-Cintas e Remael (2007), para que uma legenda de 1,4 segundos permaneça a 145ppm, ela deveria ter aproximadamente 19 caracteres, menos da metade de caracteres da legenda 354 (49 caracteres). Apenas a condensação do texto não será suficiente para uma ressegmentação adequada da legenda. Logo, propõe-se a realocação do tempo da legenda anterior à legenda 354, legenda 353, cuja duração é maior que o necessário para torná-la uma legenda lenta (possui 33 caracteres exibidos em 3,4s).

A nova duração de 2,4 segundos permitiria à legenda 354 conter 36 caracteres, limite este respeitado se o texto condensado retirar a conjunção [Porque] e o advérbio [já], resultando na seguinte proposta de ressegmentação.

ESTOU MEIO CANSADA
DE ADULTOS LOUCOS.

4.5. Sintagma preposicionado

Também o sintagma preposicionado quase não apresentou problemas de segmentação, havendo apenas 2 ocorrências do tipo <SP_prep+SN>, cuja quebra acontece entre o núcleo preposicional e o sintagma nominal posposto. Vejamos o seguinte exemplo:

Quadro 9 - Exemplo de problema de segmentação <SP_prep+SN>

Legenda	Nº de linhas	Duração	Velocidade	Sintagma Preposicional
Sub 1326	2	1,9s	180ppm (alta)	comparada com as das outras competições?

Fonte: os autores

O Quadro 9 apresenta mais um caso de legenda com escassez de tempo: os 1,9 segundos de duração da legenda 1326 só permitiriam um texto de 25 caracteres, espaço curto para condensar os 39 caracteres da legenda. Logo,

seria necessário realocar tempo de alguma legenda vizinha, porém tanto a legenda anterior, quanto a posterior possuem tempo muito reduzido.

O tempo reduzido das legendas neste trecho do filme resultou de um descuido do legendista na edição da legenda: cerca de 50 segundos antes da legenda 1326, a legenda 1309 teve uma duração excessiva de 7,3 segundos para um texto de 67 caracteres, a qual demandaria uma duração de apenas 4,8 segundos. O atraso de 2,5 segundos na legenda 1309 comprometeu a sincronia das legendas num trecho de quase 1 minuto cheio de diálogos longos.

O espaço diminuto deste artigo não permite analisar a ressegmentação de todo este trecho, de modo que será proposta apenas uma solução atenuadora do problema de segmentação.

COMPARADA A OUTRAS COMPETIÇÕES

É possível substituir o trecho [com as das] pela preposição [a], mantendo a forma participial [comparada] na linha superior, resultando numa legenda de 30 caracteres: uma economia de 9 caracteres.

4.6. Sintagma adverbial

Não houve casos de quebra no sintagma adverbial que permitissem uma análise a partir do presente corpus. O exame extenso feito em Araújo (2015), porém, mostra que esse tipo de quebra se dá comumente quando há um advérbio intensificando outro, este último frequentemente agindo como Especificador (segundo a terminologia de Castilho (2012)) de um sintagma adjetival. Por exemplo: se numa legenda aparecesse a oração [João é bem mais alto que Joaquim.], se a quebra acontecesse entre [bem] e [mais], a mesma seria uma quebra do tipo <SAdv_adv+SAdv>.

4.7. Oração coordenada

A ausência de problemas de segmentação nesta categoria não é de todo surpreendente, pois mesmo num corpus bem mais abrangente como o do projeto CORSEL (ARAÚJO, 2015), que tem cerca de 100 mil palavras, naqueles programas com legendas *pop-on*, ocorreram índices baixíssimos de quebra de linha problemática dessa natureza - 3,76% no gênero novela (ARAÚJO, 2015, p. 61); no gênero sitcom, 0% em 'Tapas e Beijos' (ARAÚJO, 2015, p. 67), 3,92% em 'Pé na jaca' (ARAÚJO, 2015, p. 67), e apenas 1,69% em 'A Grande Família' (ARAÚJO, 2015, p. 71). Ocorreram mais casos deste tipo de problema de segmentação no gênero documentário, totalizando 6,93% (ARAÚJO, 2015, p. 77), provavelmente por exibirem legendas do tipo *roll-up*.

Terminadas as análises das sete categorias de problemas de segmentação, resta analisar as condições de número de linhas e velocidade das legendas com segmentação problemática.

4.8. Relação entre problemas de segmentação e parâmetros técnicos

A Tabela 4 mostra, primeiramente, que a grande maioria dos problemas de segmentação linguística (quase três quartos) acontecem em legendas de 2 linhas, fato antevisto se levarmos em conta que apenas 10,8% das 1704 legendas eram de 3 linhas (ARRAES, 2015). Um número pequeno de legendas de 3 linhas implica em um número menor de problemas de segmentação nessa categoria. Desse modo, o número de <PROSEGL> em legendas de 2 linhas é mais que o triplo das de 3 linhas.

Tabela 4 - Número de linhas e velocidade nos problemas de segmentação

	2 linhas	3 linhas	145ppm Vel. Baixa	160ppm Vel. Média	180ppm Vel. Alta
VA	71	21	34	6	52
%	77,2	22,8	36,9	6,6	56,5

Fonte: os autores

A baixa ocorrência de legendas de 3 linhas em *Virada Radical* é intrigante se considerarmos os dados do projeto CORSEL (ARAÚJO, 2015), os quais mostram que, para aqueles gêneros cuja legenda é do tipo *pop-on*, houve uma tendência ao equilíbrio entre os <PROSEGL> de 2 e 3 linhas: no gênero novelas, a diferença é de apenas 3,85% (ARAÚJO, 2015, p. 61); no gênero sitcom, a diferença é de meros 4,44% (ARAÚJO, 2015, p. 76). Quanto ao gênero documentário, como este utiliza legendas *roll-up*, não pode ser devidamente comparado aqui, pois o número de linhas na tela irá depender de quantas linhas são permitidas pelo sistema usado pela emissora.

Com relação ao parâmetro da velocidade das legendas, conforme a Tabela 4, pouco mais da metade dos problemas de segmentação ocorreram em legendas de 180ppm (56,5%), enquanto outra porção significativa dos problemas acontece em legendas de 145ppm (36,9%). Haja vista que, do total de legendas, apenas 4,7% tinham velocidade 160ppm, é natural que haja poucos problemas de segmentação nas legendas com esta velocidade.

Ao contrário do que aconteceu com o número de linhas, os dados acerca da velocidade das legendas se assemelham bastante àqueles obtidos pelo projeto CORSEL (ARAÚJO, 2015): considerando a velocidade de 160ppm como zona de transição entre as outras duas velocidades, seu percentual tanto nas novelas como nas sitcom também girou em torno de 6%; de igual modo, em ambos os gêneros também houve predominância de ocorrências de problemas de segmentação em legendas de 180ppm.

5. Considerações finais

O primeiro objetivo deste estudo foi analisar e descrever os problemas de segmentação nas legendas do filme *Virada Radical*. As análises constataram um número pequeno (5,4%) de problemas de segmentação, se considerarmos o tamanho do corpus, de 1704 legendas, e se este resultado for comparado com o percentual de legendas com problemas de segmentação nos demais gêneros do projeto CORSEL. Isto parece indicar um maior cuidado por parte dos profissionais responsáveis na confecção das legendas, o que também explicaria a baixa quantidade de problemas de segmentação em legendas de três linhas. Das sete categorias de PROSEGL em sintagmas e orações, houve maior quantidade de problemas nos sintagmas verbais, nominais e nas orações subordinadas, porém, houve uma proporção muito maior de quebras problemáticas nos sintagmas verbais, cerca de 77%. Não é possível afirmar se isto é uma característica do gênero ou se foi uma peculiaridade do filme analisado. Em face desses resultados, o primeiro objetivo da pesquisa foi alcançado.

O segundo objetivo da pesquisa foi aferir, quanto aos parâmetros de velocidade e número de linhas, em quais condições estes problemas acontecem. Os resultados indicam uma quantidade baixa de PROSEGL em legendas de três linhas, o que talvez seja o resultado de maior cuidado no momento de confecção de legenda. Entretanto, não é possível dizer se isto acontece na legendagem de outros filmes. Apenas a análise de outros filmes exibidos em TV permitiria tecer generalizações. Quanto ao parâmetro da velocidade, os resultados deste estudo apresentaram certo equilíbrio entre as velocidades baixa e alta - 56,5% de problemas em legendas de 180 ppm e 36,9% em legendas de 145ppm. No projeto CORSEL houve maior proeminência de problemas em velocidade alta (62,7% em sitcom, 56,1,3% nas novelas), mas nenhuma diferença significativa. Em face desses resultados, o segundo objetivo da pesquisa foi alcançado.

O terceiro objetivo da pesquisa foi comparar os resultados da presente pesquisa com aqueles obtidos pelo projeto CORSEL. No início da seção de

análise, apontou-se para o percentual significativamente menor de problemas de segmentação no filme *Virada Radical* em comparação com o percentual obtido pelos gêneros analisados pelo CORSEL. Depois, comparou-se as quantidades de PROSEGL nas sete categorias analisadas e percebeu-se uma quantidade muito maior de problemas no sintagma verbal em *Virada Radical* do que nos demais gêneros do projeto CORSEL. Em face desses resultados, o terceiro objetivo da pesquisa foi alcançado.

Esta pesquisa, junto ao projeto CORSEL, trouxe um passo significativo na descrição dos problemas de segmentação linguística nas LSEs da TV brasileira, mas ainda é incapaz de apresentar uma teoria geral de como os problemas de segmentação interferem no processamento cognitivo das legendas. Sem isso, é impossível estender os achados até aqui conquistados para outras línguas que não o português brasileiro. A elaboração desta teoria ainda está por ser enunciada e acreditamos que só poderá ser formulada através de pesquisas experimentais através de rastreamento ocular (como por exemplo: Vieira, Teixeira e Chaves (2017) e Monteiro e Dantas (2017)), pois sem isso, toda especulação sobre o processamento mental das legendas continuará no plano das conjecturas.

Referências

- ARAÚJO, V. L. S. Closed subtitling in Brazil. In: ORERO, P. (Org.). *Topics in Audiovisual translation*. Amsterdã: John Benjamins, 2004: 199-212.
- _____. Por um modelo de legendagem para Surdos no Brasil. *Tradução e comunicação: Revista Brasileira de Tradutores*, n. 17, 2008: 59-76. Disponível em: <http://sare.anhanguera.com/index.php/rtcom/article/view/148/147>. (24/03/2016).
- _____. *Legendagem para surdos: em busca de um modelo para o Brasil*. Relatório Técnico n. 306948/2008-7. Fortaleza: CNPq. Fev/2012. Não publicado.
- _____. *A segmentação na legendagem para surdos e ensurdecidos (LSE): um estudo baseado em corpus*. Relatório Técnico n. 306441/2011-0. Fortaleza: CNPq. Fev/2015. Não publicado.
- _____; NASCIMENTO, A. K. P. Investigando parâmetros de legendas para Surdos e Ensurdecidos no Brasil. *Tradução em Revista*, n. 11, 2011: 1-

18. Disponível em: <http://www.maxwell.lambda.ele.puc-rio.br/18862/18862.PDF>. (24/03/2016).
- ARAÚJO, V. L. S.; VIEIRA, P. A.; MONTEIRO, S. M. M.. Legendagem para surdos e ensurdecidos (LSE): um estudo de recepção com surdos da região sudeste. *TradTerm*, v. 22, 2013: 283-302. Disponível em: <http://www.revistas.usp.br/tradterm/article/view/69132/71589>. (24/03/2016).
- ARAÚJO, V. L. S.; ASSIS, I. A. P.. A segmentação linguística na legendagem para surdos e ensurdecidos (LSE) de ‘Amor Eterno Amor’: uma análise baseada em corpus. *Letras & Letras*, v. 30, n. 2, jul/dez 2014: 156-184. Disponível em: <http://www.seer.ufu.br/index.php/letraseletras/article/download/28562/15798>. (19/01/2015).
- ARRAES, D. A.. *Legendagem para surdos e ensurdecidos: análise baseada em corpus da segmentação linguística do filme “Virada Radical”*. 2015. 71f. Trabalho de conclusão de curso (Bacharelado em Letras). Curso de Letras, Universidade Estadual do Ceará, Fortaleza, 2015.
- ASSIS, I. A. P.. *A segmentação na LSE de amor eterno amor: uma análise baseada em corpus*. 2013. 72f. Trabalho de conclusão de curso (Bacharelado em Letras). Curso de Letras, Universidade Estadual do Ceará, Fortaleza, 2013.
- _____. *Legendagem para Surdos e Ensurdecidos (LSE): análise baseada em corpus da segmentação linguística em Amor Eterno Amor*. 2015. 107f. Dissertação (Mestrado em Linguística Aplicada). Universidade Estadual do Ceará, Fortaleza, 2016. Disponível em: http://www.uece.br/posla/dmdocuments/Disserta%C3%A7%C3%A3o_%C3%8Dtalo%20Alves.pdf.pdf. (20/10/2016).
- BERBER SARDINHA, T. *Linguística de Corpus*. Barueri, São Paulo: Manole, 2004.
- CASTILHO, A. T. *Nova Gramática do Português Brasileiro*. São Paulo: Contexto, 2012.
- CHAVES, E. G. *Legendagem para Surdos e Ensurdecidos: um estudo baseado em corpus da segmentação nas legendas de filmes brasileiros em DVD*. 126f. Dissertação (Mestrado em Linguística Aplicada). Universidade Estadual do Ceará, Fortaleza- CE, 2012.
- _____; ARAÚJO, V. L. S. Segmentation tags: a proposal for the analysis of subtitles. In: ALUÍSIO, S. M.; TAGNIN, S E. O. (orgs.) *New language, technologies and linguistic research: a two way road*. Newcastle upon Thyne: Cambridge Scholar’s Publishing, 2014: 62-75.
- DIAZ-CINTAS, J. Subtitling. In: GAMBIER, Y.; DOORSLAER, D. (Eds.). *Handbook of Translation Studies*. Amsterdam/Philadelphia: John Benjamins, 2010: 344-349.
- DIAZ-CINTAS, J.; REMAEL, A. *Audiovisual Translation: Subtitling*. Manchester, UK, Kinderhook, NY, UK: St. Jerome Publishing, 2007.
- D’YDEWALLE, G.; PRAET, C.; VERFAILLIE, K.; VAN RENSBERGEN, J. Reading a message when the same message is available auditorily in another language: the case of subtitling. In: O’REGAN, J. K.; LÉVY-SCHOEN, A. (eds.). *Eye Movements: From Physiology to Cognition*. Amsterdam, New York: Elsevier Science Publishers, 1987: 313-321.

- FRANCO, E.; ARAÚJO, V. L. S. Reading Television: Checking deaf people's Reactions to Closed Subtitling in Fortaleza, Brazil. *The Translator: studies in intercultural communication*, v. 9, n. 2, 2003: 249-267. Disponível em: [http://www.tandfonline.com/doi/abs/10.1080/13556509.2003.10799156?src. \(15/03/2015\).](http://www.tandfonline.com/doi/abs/10.1080/13556509.2003.10799156?src. (15/03/2015).)
- _____. Questões terminológico-conceituais no campo da tradução audiovisual (TAV). *Tradução em Revista*, v. 2, 2011: 1-23.
- FRANCO, R. P.; CHAVES, E. G. Legendagem para Surdos e Ensurdidos (LSE): Um Estudo da Segmentação Linguística nas Séries da TV Brasileira. *Mosaico*, v. 14, 2015: 247-265.
- GABRIEL, M. H. C.. *A segmentação linguística na legendagem para surdos e ensurdidos (LSE): uma análise baseada em corpus*. 2013. 59f. Trabalho de conclusão de curso (Especialização em Estudos da Tradução). Universidade Federal do Ceará, Fortaleza, 2013.
- _____. *Problemas de segmentação linguística na legendagem para surdos e ensurdidos (LSE) de "Cheias de Charme": uma análise baseada em corpus*. 109f. Dissertação (Mestrado em Linguística Aplicada). Universidade Estadual do Ceará, Fortaleza, 2015. Disponível em: [http://www.uece.br/posla/dmdocuments/DISSERTA%C3%87%C3%83O%20-%20MARIA%20HELENA%20CLARINDO%20GABRIEL.pdf. \(10/01/2016\).](http://www.uece.br/posla/dmdocuments/DISSERTA%C3%87%C3%83O%20-%20MARIA%20HELENA%20CLARINDO%20GABRIEL.pdf. (10/01/2016).)
- KARAMITROGLOU, F. A Proposed Set of Subtitling Standards in Europe. *Translation Journal*, v. 2, n. 2, 1998: p. 1-15. Disponível em: [http://translationjournal.net/journal/04stndrd.htm. \(10/03/2014\).](http://translationjournal.net/journal/04stndrd.htm. (10/03/2014).)
- MASCARENHAS, R. O. *A audiodescrição da minissérie policial Luna Caliente: uma proposta de tradução à luz da narratologia*. 285f. 2012. Tese (Doutorado). Instituto de Letras, Universidade Federal da Bahia, Salvador, 2012.
- MONTEIRO, S. M. M.; DANTAS, J. F.. Tradução audiovisual acessível (TAVa): a segmentação linguística na legendagem para surdos e ensurdidos (LSE) da campanha política na televisão em Fortaleza. *Trabalhos em Linguística Aplicada*, n. 56, v. 2, 2017: 527-560.
- VIEIRA, P. A.; ARAÚJO, V. L. S.; CHAVES, E. G.; MONTEIRO, S. M. M. A segmentação na legendagem do documentário Globo Repórter: uma análise baseada em corpus. In: PONTES, V. O.; CUNHA, R. B.; CARVALHO, E. P.; TAVARES, M. G. G. (Orgs.). *A tradução e suas interfaces: múltiplas perspectivas*. Curitiba: CRV, 2015: 65-86.
- VIEIRA, P. A.; TEIXEIRA, E. N.; CHAVES, E. G.. Caminhos do olhar: a movimentação ocular de espectadores surdos durante a exibição de vídeos legendados. *Trabalhos em Linguística Aplicada*, v. 56, n. 2, 2017: 493-525.