



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

ÉLIDA GAMA CHAVES

LEGENDAGEM DE FILMES EM LÍNGUA INGLESA E PORTUGUESA: ANÁLISE
DOS PARÂMETROS DE SEGMENTAÇÃO E VELOCIDADE DE LEGENDAS
PARA OUVINTES

FORTALEZA – CEARÁ

2017

ÉLIDA GAMA CHAVES

LEGENDAGEM DE FILMES EM LÍNGUA INGLESA E PORTUGUESA: ANÁLISE DOS
PARÂMETROS DE SEGMENTAÇÃO E VELOCIDADE DE LEGENDAS
PARA OUVINTES

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

Orientação: Profa. Dra. Vera Lúcia Santiago Araújo.

FORTALEZA – CEARÁ

2017

Dados Internacionais de Catalogação na Publicação

Universidade Estadual do Ceará

Sistema de Bibliotecas

Chaves, Élide Gama.

Legendagem de filmes em língua inglesa e portuguesa: análise dos parâmetros de segmentação e velocidade de legendas para ouvintes [recurso eletrônico] / Élide Gama Chaves. - 2017.

1 CD-ROM: il.; 4 ¼ pol.

CD-ROM contendo o arquivo no formato PDF do trabalho acadêmico com 218 folhas, acondicionado em caixa de DVD Slim (19 x 14 cm x 7 mm).

Tese (doutorado) - Universidade Estadual do Ceará, Centro de Humanidades, Programa de Pós-Graduação em Linguística Aplicada, Fortaleza, 2017.

Área de concentração: Linguagem e Interação.

Orientação: Prof.ª Ph.D. Vera Lúcia Santiago Araújo.

1. Tradução Audiovisual. 2. Legendagem. 3. Linguística de Corpus. 4. Segmentação. 5. Velocidade. I. Título.

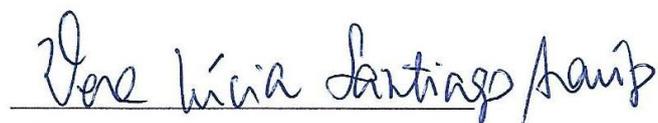
ÉLIDA GAMA CHAVES

LEGENDAGEM DE FILMES EM LÍNGUA INGLESA E PORTUGUESA: ANÁLISE DOS
PARÂMETROS DE SEGMENTAÇÃO E VELOCIDADE DE LEGENDAS
PARA OUVINTES

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

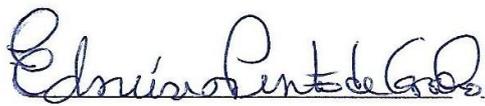
Aprovada em: 19 de setembro de 2017.

BANCA EXAMINADORA



Profa. Dra. Vera Lúcia Santiago Araújo

UECE (Orientadora)



Profa. Dra. Ednúsia Pinto de Carvalho

UFC



Profa. Dra. Patrícia Araújo Vieira

UFC



Prof. Dr. Wilson Júnior de Araújo Carvalho

UECE



Prof. Dr. Pedro Henrique Lima Praxedes

Filho – UECE

À Vera Lúcia Santiago Araújo

DEDICO

AGRADECIMENTOS

Ao meu bom Deus. Minha energia.

À minha família: meus queridos pais Irajá Wagner e Maria Irene, aos meus irmãos Átila e Nélida e aos meus sobrinhos Oscar Guerreiro e Heitor Chaves. Minhas referências, minhas raízes.

Ao meu amor Douglas. Sobrevivemos.

Aos peludos da minha vida: Pink, Caue, Brad, João Billy, Lua e Baruk, o fardo foi muito mais leve com vocês.

Aos meus leitores atentos: Ednúsia Carvalho, Patrícia Vieira, Wilson Carvalho, Pedro Praxedes, Malena Monteiro e Salete Nunes.

Ao meu primo Iago Chaves, pelo desenvolvimento e doação do etiquetador de legendas a essa pesquisa. E à Alexandra Seoane, que atualizou e melhorou a versão do programa.

Ao nutricionista Marcelo Lacerda. A dieta da Tese, além de deliciosa, me deu energia suficiente para produzir nesse momento.

Aos meus queridos amigos do LEAD-LATAV, especialmente minha querida Marisa Aderaldo.

Aos meus queridos amigos: Renatta Franco, Sofia Amoreira e Ednardo Souza. Gratidão. Vocês não imaginam o quanto me ajudaram nesse processo.

Às amigas Patrícia, Malena, Salete, Alexandra e Bruna. Unidas pela tese.

À competente revisora Ana Carla. Sua compaixão diminuiu meu fardo. Você foi meu anjo, não tenho dúvidas.

À Universidade Estadual do Ceará. Toda a minha vida acadêmica e profissional.

RESUMO

Esta tese está inserida no campo da Tradução Audiovisual, mais especificamente nos estudos da Legendagem. Na literatura sobre a pesquisa experimental em legendagem, a velocidade baixa (abaixo de 145ppm), a segmentação do texto audiovisual em uma linha e os problemas de segmentação linguística são considerados condições prejudiciais à recepção e ao processamento de legendas. Desta forma, esta pesquisa avaliou legendas para ouvintes, de filmes em DVD, nos contextos de língua inglesa e portuguesa, analisando a velocidade e os padrões de distribuição do texto audiovisual e de segmentação linguística encontrados e como eles se relacionam entre si e com aqueles da LSE de filmes em DVD e de programas de TV. Com subsídios da Linguística de *Corpus*, a metodologia utilizada foi a baseada em *corpus*, com a qual foram analisadas as legendas (em inglês e português) de cinco filmes estrangeiros em língua inglesa distribuídos no Brasil. Foram feitas as etapas de extração do *corpus*, utilizando o extrator de legendas de DVD *SubRip*; etiquetagem dos parâmetros técnicos e linguísticos da legendagem, com base no conjunto de etiquetas do Projeto CORSEL; e processamento dos dados com auxílio do programa *WordSmith Tools 7.0*. Verificou-se que a LO em inglês apresenta 45% de legendas com velocidade abaixo de 145ppm e 36% de legendas com velocidade de 180ppm e acima de 180ppm, 55% de legendas de uma linha e 4,2% de problemas de segmentação linguística, enquanto a LO em português apresenta 66% de legendas com velocidade abaixo de 145ppm, 62% de legendas de uma linha e 4,7% de problemas de segmentação linguística. Nas duas línguas, há uma predominância de legendas de velocidade abaixo de 145ppm, entretanto, em inglês, há um percentual expressivo de legendas rápidas, o que as torna rápidas em relação às legendas em português. A segmentação ou distribuição do texto audiovisual é feita predominantemente em uma linha de legenda, e todo o *corpus* apresentou tipos e subtipos de problemas de segmentação linguística semelhantes, havendo maior frequência de problemas no sintagma verbal, seguido dos sintagmas nominal e preposicional. Relacionando-se esses parâmetros, verifica-se que em inglês as legendas analisadas são rápidas e mal segmentadas e em português são lentas e mal segmentadas; desse modo, ambas têm sua qualidade comprometida e não condizem com os parâmetros preconizados nas pesquisas exploratórias, descritivas e experimentais da área.

Palavras-chave: Tradução Audiovisual. Legendagem. Linguística de *Corpus*. Segmentação.

ABSTRACT

This thesis is part of the Audiovisual Translation studies, more specifically the Subtitling studies. In the literature about experimental research in subtitling, low subtitle rate (below 145wpm), the audiovisual text segmentation in one line and linguistic segmentation problems are considered harmful conditions to the reception and the reading process of subtitles. Thus, this research evaluated DVD movies subtitles for the hearing, in the contexts of English and Portuguese languages, analyzing the subtitle rate and the patterns of audiovisual text distribution and linguistic segmentations found and how they compare to each other and to SDH subtitles from DVD movies and TV shows. With the assistance of *Corpus Linguistics*, the methodology used was based on *corpus*, with which we analyzed the subtitles (in English and Portuguese) from five English-language foreign movies distributed in Brazil. The *corpus* was obtained using the DVD subtitle extractor *SubRip*; then the technical and linguistic parameters of subtitling were annotated, based on the labels from Project CORSEL; and the data was processed using the software *WordSmith Tools 7.0*. It was observed that the *corpus* in English has 45% of subtitles below 145wpm and 36% at 180wpm and above 180wpm, 55% of one-line subtitles and 4.2% of linguistic segmentation problems, while the *corpus* in Portuguese has 66% of subtitles below 145wpm, 62% of one-line subtitles and 4.7% of linguistic segmentation problems. In both languages, subtitles below 145wpm are predominant; however, in English, there is also a high percentage of fast subtitles, which makes them fast when compared to the subtitles in Portuguese. The segmentation or distribution of the audiovisual text in one line is predominant, and the entire *corpus* revealed similar linguistic segmentation categories and subcategories, with a higher frequency of problems in the verbal syntagma, followed by noun and prepositional syntagmas. Connecting these parameters, it was observed that the subtitles analyzed are fast and ill segmented in English and slow and ill segmented in Portuguese; so, both have their quality compromised and neither follows the parameters preconized by the exploratory, descriptive and experimental researches in the area.

Keywords: Audiovisual Translation. Subtitling. *Corpus Linguistics*. Segmentation.

LISTA DE QUADROS

Quadro 01 –	Formatos de legenda	34
Quadro 02 –	Relação caracteres por segundo na regra europeia	39
Quadro 03 –	Quadro sinótico das etiquetas do Projeto CORSEL	63
Quadro 04 –	Dados quantitativos provenientes da <i>WordList</i>	73
Quadro 05 –	Quantitativo de problemas de segmentação linguística nas legendas em línguas inglesa e portuguesa	79
Quadro 06 –	Características da LO em inglês de filmes em DVD	114
Quadro 07 –	Características da LO em português de filmes em DVD	115

LISTA DE GRÁFICOS

Gráfico 01 –	Categorias de problemas de segmentação linguística: ocorrências em inglês e português	80
Gráfico 02 –	Subcategorias de problemas de segmentação linguística: ocorrências em inglês e português	81
Gráfico 03 –	Velocidade das legendas em inglês e português	102
Gráfico 04 –	Representação da velocidade em linhas	105
Gráfico 05 –	Relação entre PROSEGL e velocidade no <i>corpus</i>	107
Gráfico 06 –	Relação entre legendas sem PROSEGL e velocidade	108
Gráfico 07 –	Percentual de problemas de segmentação dos seis tipos de legenda	110
Gráfico 08 –	Categorias de PROSEGL na LO e na LSE	111
Gráfico 09 –	Interação segmentação e velocidade nas legendas em inglês e português	116

LISTA DE FIGURAS

Figura 01 – Exemplo de intertítulo (<i>intertitle</i>). Letra a: fotograma que antecede o intertítulo. Letra b: intertítulo que diz, em português: “Amanhã os pássaros cantarão”	26
Figura 02 – Legenda sobreposta à imagem	26
Figura 03 – Mapa da legendagem demarcando os tipos de legenda	27
Figura 04 – Esquema de caracterização das legendas	29
Figura 05 – Exemplo de TCR	35
Figura 06 – Exemplo de legenda sem problema (a) e com problema (b) de segmentação linguística	41
Figura 07 – Árvore sintática utilizada para segmentação linguística por Karamitroglou (1998)	44
Figura 08 – Capas dos filmes em DVD	57
Figura 09 – Interface do programa <i>SubRip 1.50</i>	60
Figura 10 – Arquivo de legenda em formato .srt gerado pelo <i>SubRip</i> e aberto no <i>Bloco de Notas</i>	61
Figura 11 – Etiquetador automático de parâmetros técnicos da legendagem	65
Figura 12 – Arquivos .srt não anotado e anotado	68
Figura 13 – Tela da <i>WordList</i> com informações das legendas em inglês e português dos 5 filmes	69
Figura 14 – Tela da <i>Concord</i> com os problemas de segmentação linguística em nóduo	70
Figura 15 – Dados numéricos da <i>WordList</i>	73
Figura 16 – Quebra de sintagma verbal <SV_composto> em inglês (a) e português (b)	83
Figura 17 – Quebra de sintagma verbal <SV_verbo+SAdv> em inglês (a) e português (b)	86
Figura 18 – Quebra de sintagma nominal <SN_especif+SN> em inglês (a) e português (b)	89
Figura 19 – Quebra de sintagma nominal <SN_subst+SAdj> em inglês (a) e português (b)	92

Figura 20 – Quebra de sintagma preposicional <SP_prep+SN> em inglês (a) e português (b)	94
Figura 21 – Quebra de oração subordinada <SUBORD_ conj/pron_rel+oração> em inglês (a) e português (b)	95
Figura 22 – Quebra de oração coordenada <COORD_ conj+oração> em inglês (a) e português (b)	97
Figura 23 – Problema de segmentação pela distribuição de texto nas legendas de duas linhas em inglês	99
Figura 24 – Problema de segmentação pela distribuição de texto nas legendas de duas linhas em português	101

LISTA DE ABREVIATURAS E SIGLAS

AD	Audiodescrição
COORD	Orações coordenadas
CORSEL	A segmentação na legendagem para surdos e ensurdecidos (LSE): um estudo baseado em <i>corpus</i>
Cps	Caracteres por segundo
ETBC	Estudos da Tradução Baseados em <i>Corpora</i>
EXLEG	Estudos Experimentais em Legendagem: análise da velocidade e da segmentação
LATAV	Laboratório de Tradução Audiovisual
LC	Linguística de <i>Corpus</i>
LEAD	Legendagem e Audiodescrição
Libras	Língua brasileira de sinais
LO	Legendagem para Ouvintes
LSE	Legendagem para Surdos e Ensurdecidos
ms	Milissegundos
MOLES	Modelo de Legendagem para Surdos
PosLA	Programa de Pós-Graduação em Linguística Aplicada
ppm	Palavras por minuto
SAdj	Sintagma adjectival
SAdv	Sintagma adverbial
SN	Sintagma nominal
SP	Sintagma preposicional
SUBORD	Orações subordinadas
SV	Sintagma verbal
TAV	Tradução Audiovisual
TAVa	Tradução Audiovisual Acessível
UECE	Universidade Estadual do Ceará
VHS	Video Home System

SUMÁRIO

1	INTRODUÇÃO	15
2	FUNDAMENTAÇÃO TEÓRICA	23
2.1	TRADUÇÃO AUDIOVISUAL LEGENDAGEM	23
2.2	PARÂMETROS DA LEGENDAGEM	30
2.3	ESTUDOS DA LEGENDAGEM BASEADOS EM <i>CORPORA</i> E PESQUISAS SOBRE RECEPÇÃO E PROCESSAMENTO DE LEGENDAS	42
3	METODOLOGIA	56
3.1	CONTEXTO DA PESQUISA	56
3.2	TIPO DE PESQUISA	56
3.3	<i>CORPUS</i>	57
3.4	INSTRUMENTOS	59
3.5	TRATAMENTO E ANÁLISE DO <i>CORPUS</i>	60
3.5.1	Extração das legendas	60
3.5.2	Etiquetagem do <i>corpus</i>	61
3.5.2.1	Etiquetas de problemas de segmentação	66
3.5.2.2	Concepção de sintagmas e orações de Castilho (2012)	67
3.5.3	Processamento do <i>corpus</i> via <i>WordSmith Tools</i>	68
3.6	PROCEDIMENTOS PARA ANÁLISE DOS DADOS	71
4	RESULTADOS E DISCUSSÃO	72
4.1	PANORAMA DO <i>CORPUS</i>	72
4.2	COMPARAÇÃO DOS PROBLEMAS DE SEGMENTAÇÃO E VELOCIDADE NAS LEGENDAS DE DVD PARA OUVINTES: INGLÊS × PORTUGUÊS	76
4.2.1	Problemas de segmentação nas legendas em inglês e em português	77
4.2.1.1	Problemas de segmentação no sintagma verbal	82
4.2.1.2	Problemas de segmentação no sintagma nominal	88
4.2.1.3	Problemas de segmentação no sintagma preposicional	93
4.2.1.4	Problemas de segmentação na oração subordinada	95
4.2.1.5	Problemas de segmentação na oração coordenada	96
4.2.2	Velocidade das legendas em inglês e em português	101
4.2.3	Problemas de segmentação e velocidade: LO em inglês × LO em português	105

4.3	COMPARAÇÃO DOS PROBLEMAS DE SEGMENTAÇÃO E VELOCIDADE NAS LEGENDAS EM PORTUGUÊS: LO × LSE	109
4.3.1	Problemas de segmentação e velocidade: LO × LSE	109
4.4	DESCRIÇÃO DOS PADRÕES DAS LEGENDAS PARA OUVINTES: INGLÊS × PORTUGUÊS	113
4.4.1	Padrão da LO de DVD em inglês e português	114
4.5	LO DE FILMES EM DVD EM INGLÊS E PORTUGUÊS: SÍNTESE DOS RESULTADOS	117
5	CONSIDERAÇÕES FINAIS	120
	REFERÊNCIAS	123
	ANEXOS	128
	ANEXO A – Problemas de segmentação linguística nas legendas em inglês	129
	ANEXO B – Problemas de segmentação linguística nas legendas em português ...	171

1 INTRODUÇÃO

Esta tese se situa no campo da Tradução Audiovisual (TAV). No âmbito institucional, está inserida na linha de pesquisa 2 – Multilinguagem, Cognição e Interação – do Programa de Pós-Graduação em Linguística Aplicada (PosLA) da UECE, e no projeto “A segmentação na legendagem para surdos e ensurdecidos (LSE): um estudo baseado em *corpus*” (CORSEL – 2012-2015). O projeto descreveu e analisou, por meio de uma metodologia baseada em *corpus*, o parâmetro da segmentação, ou seja, a distribuição das falas de uma produção audiovisual em uma ou mais legendas, em alguns programas de TV brasileiros (novelas, seriados e documentários). Também se desenvolve a partir dos pressupostos teórico-metodológicos de pesquisas descritivas, experimentais e exploratórias em legendagem (D’YDEWALLE, 1987; KARAMITROGLOU, 1998; DE LINDE; KAY, 1999; DIAZ CINTAS; REMAEL, 2007; PEREGO, 2008; ARAÚJO; NASCIMENTO, 2011; CHAVES, 2012; ARAÚJO, 2015; VIEIRA, 2016; MONTEIRO, 2016).

A legendagem é uma modalidade de TAV encontrada em produtos audiovisuais diversos, como filmes em DVD, cinema, televisão, vídeos disponíveis na internet, eventos ao vivo e teatro. Consiste na tradução das falas de uma produção audiovisual exibida em forma de texto escrito. Com base na taxonomia de Jakobson (1959) para os aspectos linguísticos de uma tradução, a legendagem pode ser interlinguística (entre duas línguas diferentes), intralinguística (dentro da mesma língua) e intersemiótica (entre meios semióticos diferentes), como, por exemplo, a tradução dos elementos acústicos na legendagem para surdos e ensurdecidos (LSE) – os efeitos sonoros – em forma de texto escrito – as legendas. Por exemplo, para que o surdo saiba que a campanha está tocando, cola-se entre colchetes a legenda [Som da campanha].

No que diz respeito à classificação das legendas nesta tese analisadas, são do tipo: legendagem para ouvintes, de filmes em DVD. Quanto ao aspecto técnico, são fechadas; quanto ao aspecto linguístico das legendas em inglês, são intralinguísticas, e das legendas em português, são interlinguísticas. São produzidas por programas editores de legendas e sincronizadas ao produto.

Assim como as demais modalidades de tradução, a legendagem é, em sua essência, acessível, pois facilita o acesso a uma fonte de informação e entretenimento anteriormente hermética (DÍAZ CINTAS, 2005). Para isso, é preciso que no momento da tradução, atenda a parâmetros técnicos e linguísticos, como velocidade da legenda, número de linhas, caracteres por linha, formato, posição das legendas, convenções linguísticas, redução,

explicitação, segmentação etc. no momento da tradução. Tais parâmetros, a serem detalhados na seção teórica, presentes em legendas tanto para ouvintes (LO) quanto para surdos e ensurdecidos (LSE), têm sido investigados por pesquisadores da tradução audiovisual em vários países, especialmente os da Europa e o Brasil.

Algumas dessas pesquisas (D'YDEWALLE *et al.*, 1987; KARAMITROGLOU, 1998; DE LINDE; KAY, 1999; DÍAZ CINTAS; REMAEL, 2007; PEREGO, 2008; ARAÚJO; NASCIMENTO, 2011; CHAVES, 2012; ARAÚJO, 2015; VIEIRA, 2016; MONTEIRO, 2016), realizadas na busca por parâmetros que orientem melhores práticas e contribuam para melhores produtos audiovisuais, sugerem que problemas técnicos e/ou linguísticos podem acarretar problemas tradutórios, bem como problemas de recepção. D'Ydewalle *et al.* (1987) testaram o parâmetro da velocidade a partir da regra dos seis segundos, que era uma medida empírica utilizada pela maioria dos canais de TV na Europa. Para testá-la, d'Ydewalle *et al.* (1987) montaram um experimento procurando saber qual combinação de legenda (1 ou 2 linhas de 32 caracteres vs tempo de exibição de 2, 4 ou 6 segundos) os participantes julgavam apropriada. Ao final do experimento, os pesquisadores verificaram que a melhor legenda foi a de 2 linhas em 6 segundos.

A partir dessa regra, Díaz Cintas e Remael (2007) elaboraram uma tabela de velocidade como forma de dar suporte à prática da legendagem. Eles partiram da regra dos seis segundos, mas entendiam que o máximo de caracteres por linha poderia ser 35 ou 37, em vez de 32. O cálculo foi feito com base nas seguintes medidas: 37 caracteres por linha significam 74 caracteres em duas linhas, que devem ser exibidas em seis segundos. A divisão dos caracteres por segundo resultou na média de 12cps, que, escalonados pela fórmula do programa de legendagem inglês *WinCAPS*, resultaram na tabela de 145 palavras por minuto (ppm). Os autores consideram que essa tabela de 145ppm pode ser adequada para televisão, mas não para cinema e DVD; dessa forma, Díaz Cintas e Remael (2007) propuseram mais duas tabelas (160ppm e 180ppm) pensando em cinema e DVD, que aceitam mais caracteres por segundo, pelo tamanho do telão e pela facilidade de parar e iniciar, respectivamente, e também para atender públicos com velocidades de leituras maiores (160 e 180ppm). No contexto dos programas editores de legendagem, as medidas utilizadas na tradução para legendas são dadas em caracteres por segundo, e não palavras por minuto. Essas tabelas, mesmo sendo medidas aproximadas, ajudam a fazer uma marcação de legendas mais eficiente e mais sistemática.

Ainda sobre os parâmetros, De Linde e Kay (1999), em pesquisa experimental, encontraram que legendas de velocidade baixa, em média 74ppm, são mal recebidas, pois

causam aumento na deflexão e releituras em excesso. Sobre as legendas de uma ou duas linhas, d'Ydewalle e De Bruycker (2003) fizeram testes com crianças e adultos recebendo legendas de uma e de duas linhas em várias condições experimentais e perceberam que legendas de 2 linhas são melhor recebidas em todas as condições e que legendas de uma linha são problemáticas, pois não favorecem leitura sistemática e acabam causando maior custo de processamento.

Na investigação do parâmetro da segmentação, a pesquisa de Perego (2008) descreveu os tipos de problema de segmentação linguística encontrados na quebra de linhas de legendas de filmes para cinema e DVD, em línguas italiana, inglesa e húngara, com base apenas na análise sintática, mas sem fazer relação com o contexto técnico da legendagem. Além disso, não fez análise quantitativa dos dados e nem contrastou a segmentação nos idiomas analisados (italiano, inglês e húngaro); apenas ilustrou as quebras sintáticas com alguns problemas de segmentação. Contudo, trouxe grande contribuição ao estudo da segmentação na legendagem porque descreveu categorias linguísticas em que ocorriam os problemas de quebra de linha. Já pelo viés experimental, Perego *et al.* (2010) manipularam problemas de segmentação linguística no sintagma nominal de legendas em italiano de um filme húngaro e encontraram que não há diferença significativa no processamento de legendas com e sem problemas de segmentação, concluindo que as preocupações psicolinguísticas sobre a segmentação podem ser exageradas. Os resultados de Perego *et al.* (2010) divergem dos resultados de Monteiro (2016) e Vieira (2016), que comprovam que problemas de segmentação dificultam o processamento de legendas por surdos e ouvintes.

Ciclos de pesquisas desenvolvidos no Laboratório de Tradução Audiovisual (LATAV) da Universidade Estadual do Ceará (UECE) pelos pesquisadores do grupo Legendagem e Audiodescrição (LEAD) têm contribuído para a consolidação dos parâmetros da legendagem no Brasil e, assim como as demais pesquisas sobre a recepção e o processamento de legendas, são ponto de partida para a presente tese de doutorado. O projeto MOLES¹, uma das fases do ciclo de pesquisas exploratórias, explorou a recepção de LSE de curtas-metragens por surdos brasileiros, tendo como foco o parâmetro técnico da velocidade. Foram testadas legendas de 145, 160 e 180 palavras por minuto, e os resultados mostraram que os surdos receberam bem os filmes nas três velocidades, especialmente na velocidade mais alta, de 180ppm, conseguindo harmonizar imagem e legenda e fornecer detalhes sobre o conteúdo do filme (ARAÚJO; NASCIMENTO, 2011; ARAÚJO, 2012). Esses resultados

¹ Título do projeto: LEGENDAGEM PARA SURDOS: EM BUSCA DE UM MODELO PARA O BRASIL.

foram contrários aos encontrados em Araújo (2008), em que surdos cearenses indicaram a velocidade de 145ppm como a melhor recebida. Tal divergência suscitou a investigação do parâmetro da segmentação, uma vez que este não havia sido explorado nas pesquisas de recepção.

A segmentação linguística, parâmetro relacionado à divisão das legendas em blocos de texto escrito, foi analisada em programas de TV, no âmbito do projeto CORSEL (ARAÚJO, 2015), por um viés descritivo e baseado em *corpus*. A partir da compilação de um *corpus* de legendas para surdos e ensurdecidos em DVD e em programas da TV brasileira, os resultados mostraram que legendas de telenovelas, séries e documentários apresentam problemas na quebra de linha ao separar os constituintes dos sintagmas verbal, nominal, preposicional, adverbial e adjetival e das orações coordenadas e subordinadas. Os resultados também mostraram que esses problemas são mais frequentes em legendas de velocidade igual ou maior que 180ppm (CHAVES, 2012; ASSIS, 2013, 2015; GABRIEL, 2013, 2015; ARRAES, 2015; FRANCO, 2016; LEONARDO, 2016; ARAÚJO, 2015). Essas pesquisas não apresentaram resultados sobre a influência dos problemas de segmentação e de velocidade na recepção de legendas, uma vez que não tinham tal propósito, por isso abriram espaço para uma investigação experimental-exploratória via rastreamento ocular.

A partir desses achados, os problemas de segmentação na LSE passaram a ser testados experimentalmente no âmbito do projeto EXLEG², ainda em andamento. O EXLEG tem como objetivo verificar a influência dos problemas de segmentação em legendas de velocidade lenta (145ppm) e de velocidade alta (180ppm) sobre a recepção de LSE por surdos e ouvintes brasileiros, a partir de rastreamento ocular. Os resultados encontrados em Vieira (2016) para legendas de documentários e Monteiro (2016) para legendas de campanha política apontaram para a possibilidade de os problemas de segmentação afetarem a recepção de legendas rápidas e lentas, tanto para surdos quanto para ouvintes. Os resultados experimentais corroboram os exploratórios de Araújo (2012), os quais sugerem que surdos desempenham melhor processamento de legendas quando estas são rápidas e bem segmentadas. Além disso, sugerem também que os problemas de segmentação dificultam o processamento leitor de surdos e ouvintes em ambas as velocidades.

Desse modo, a pesquisa exploratória em LSE no Brasil (ARAÚJO; NASCIMENTO, 2011; ARAÚJO, 2012; VIEIRA, 2016; MONTEIRO, 2016) revela um novo dado sobre a velocidade com a pressuposição de que a LSE de velocidade alta (180ppm) é

² Título do projeto: ESTUDOS EXPERIMENTAIS EM LEGENDAGEM PARA SURDOS E ENSURDECIDOS (LSE): ANÁLISE DA VELOCIDADE E DA SEGMENTAÇÃO.

melhor recebida por surdos brasileiros, por proporcionar mais conforto e percepção de detalhes, ao contrário daquela de velocidade baixa (145ppm), que atrasa o processamento por proporcionar releituras, no caso dos surdos. Além disso, sugere que a segmentação adequada das legendas facilita a recepção de produtos audiovisuais. Já a pesquisa descritiva baseada em *corpus* mostra que há problemas de segmentação quando os constituintes de um sintagma ou de uma oração são separados. A pesquisa descritiva teve importante papel ao categorizar problemas nos sintagmas verbal, nominal, preposicional, adverbial e adjetival e nas orações coordenadas e subordinadas e ao revelar que os problemas de segmentação ocorrem com maior frequência em legendas de velocidade alta (180ppm).

A pesquisa experimental-exploratória via rastreamento ocular, por sua vez, partindo da premissa que uma boa segmentação favorece a recepção de legendas, concluiu que legendas bem segmentadas diminuem o custo de processamento por surdos e ouvintes, proporcionando maior conforto. Concluiu também que, no caso dos surdos, a LSE de velocidade baixa (145ppm) dificulta o processamento, enquanto a de velocidade alta (180ppm), quando bem segmentada, favorece o processamento de vídeos legendados. No caso dos ouvintes, tanto a legendagem rápida quanto a lenta são bem recebidas, desde que bem segmentadas. Em outras palavras, o que as pesquisas exploratórias, descritivas e experimentais sobre a LSE no Brasil têm mostrado é que a segmentação é um parâmetro importante na recepção de uma legenda. Como veremos no decorrer deste trabalho, ela também influencia os demais parâmetros da legendagem.

A partir dos resultados das pesquisas em LSE no Brasil e sua importância para a discussão teórica da LSE no campo da Tradução Audiovisual Acessível (TAVa), percebo serem necessários mais estudos sobre a LO brasileira, já que o foco na UECE tem sido a LSE. Já que, segundo Vieira (2016), a influência da segmentação acontece tanto na recepção de surdos quanto de ouvintes, esta tese realiza o estudo desse parâmetro na legendagem para ouvintes nos contextos das línguas inglesa e portuguesa, uma vez que estes não foram ainda bastante explorados no Brasil.

No contexto da LSE brasileira, as pesquisas exploratórias, descritivas e experimentais buscaram consolidar parâmetros, contribuindo com o entendimento da LSE em estudos na área da tradução audiovisual. Tal feito elucidou a LSE de telenovelas, séries, documentários e programas políticos da TV brasileira (ARAÚJO, 2008, 2009; CHAVES, 2009; ASSIS, 2013, 2015; GABRIEL, 2013, 2015; ARAÚJO, 2015; FRANCO, 2016; LEONARDO, 2016; VIEIRA, 2016; MONTEIRO, 2016), mas, como já foi dito

anteriormente, há lacunas no que diz respeito à legendagem de filmes, especialmente a legendagem para ouvintes, que começaram a ser supridas pela pesquisa de Perego (2008).

No âmbito do LATAV, o projeto MOLES (ARAÚJO, 2012) trabalhou com legendas de filmes ao explorar diferentes modelos de LSE em curtas-metragens para surdos brasileiros, mas não teve como foco a segmentação; já as pesquisas descritivas e baseadas em *corpus* tiveram como foco a segmentação (CHAVES, 2012; DINIZ, 2012; ARRAES, 2015). A pesquisa de Chaves (2012) investigou a LSE de um filme brasileiro em DVD, enquanto a pesquisa de Diniz (2012) fez uma comparação da segmentação na LSE e na legendagem para ouvintes em português de um filme brasileiro em DVD, e a de Arraes (2015) analisou a segmentação na LSE em português de um filme estrangeiro exibido na TV.

Tais pesquisas contemplaram a legendagem de filmes, enfatizando principalmente a segmentação linguística na LSE em programas televisivos. Nenhuma delas contemplou a segmentação na legendagem para ouvintes. Portanto, esta pesquisa preenche uma lacuna, abordando não somente a segmentação linguística, mas também a interação da segmentação com os outros parâmetros de legendagem, tais como a velocidade, o número de linhas, o número de caracteres e as edições dessas legendas por meio de condensação (redução ou omissão) do conteúdo. Propõe uma análise da segmentação das legendas em filmes estadunidenses e europeus distribuídos em DVD no Brasil. O estudo faz uma comparação dessas duas realidades, buscando avaliar a segmentação na legendagem nos dois contextos (língua portuguesa e inglesa). Com este estudo, pretendo descrever as características técnicas e linguísticas dos problemas de segmentação encontrados nas legendas em DVD em português e em inglês, no intuito de conhecer os problemas de ambas as realidades e propor soluções motivadas pela constatação de que problemas de segmentação causam problemas tradutórios e de recepção.

Esta pesquisa se justifica pela ausência de investigação dos parâmetros de segmentação e velocidade na legendagem para ouvintes, em português e em inglês. Academicamente, a pesquisa contribui para o desenvolvimento dos estudos da legendagem do grupo LEAD, além de fomentar discussão teórica sobre os parâmetros da legendagem, lançando luz para os contextos de línguas inglesa e portuguesa, e dar visibilidade à legendagem enquanto campo de estudo profícuo. Socialmente, há uma demanda por filmes em DVD, um tipo de mídia que disponibiliza legendas em vários idiomas, feitas com todo o rigor exigido pela indústria do audiovisual. Por acreditar na qualidade desse produto, esta tese defende, em princípio, que as legendas de filmes em DVD atendem aos padrões de uma

legendagem adequada, mostrando harmonia dos parâmetros técnicos e linguísticos, tanto em inglês quanto em português.

Dado o escopo da pesquisa, as perguntas que norteiam a investigação recaem sobre a segmentação e a velocidade em diferentes tipos de legenda, buscando responder:

1. Como se caracterizam os problemas de segmentação e a velocidade nas legendas para ouvintes, em filmes em DVD, nos contextos de línguas inglesa e portuguesa?
2. Quais as semelhanças e diferenças entre as legendas para ouvintes em filmes em DVD em português e as legendas para surdos de filmes em DVD e programas de TV em língua portuguesa?
3. Qual o padrão das legendas para ouvintes em DVD em inglês e português com base no padrão resultante das pesquisas sobre a recepção e o processamento de legendas?

A partir das questões formuladas, esta tese tem como objetivo geral: avaliar em que medida as legendas para ouvintes de filmes em DVD, nos contextos de língua inglesa e de língua portuguesa, dialogam com os padrões de segmentação e velocidade encontrados nas pesquisas descritivas e experimentais da legendagem.

E como objetivos específicos:

1. Comparar os problemas de segmentação e a velocidade das legendas para ouvintes, de filmes em DVD, nos contextos de língua inglesa (britânica e americana) e de língua portuguesa (brasileira) – LO inglês/português.
2. Comparar os problemas de segmentação e a velocidade das legendas em português, nos contextos de legendas para ouvintes (DVD) e de legendas para surdos no Brasil (DVD e TV), descritas na pesquisa de Chaves e Amoreira (2015) e no projeto CORSEL (ARAÚJO, 2015) – português LO/LSE.
3. Descrever os padrões de segmentação e velocidade das legendas para ouvintes de filmes em DVD, nos contextos de línguas inglesa e portuguesa, e comparar ao padrão resultante das pesquisas sobre a recepção e o processamento de legendas.

Esta tese está dividida em cinco seções: esta introdução, que prepara o leitor para adentrar os temas e as discussões da pesquisa; a fundamentação teórica, que aborda as principais discussões teóricas da Tradução Audiovisual Legendagem, os parâmetros da

legendagem e o aporte teórico-metodológico das principais pesquisas baseadas em *corpus* que dialogam com esta; a metodologia, que expõe como se deu o desenvolvimento desta pesquisa em todas as etapas, bem como os instrumentos de coleta e processamento do *corpus* e os procedimentos de análise dos dados; a discussão dos resultados, que mostra os achados desta pesquisa em etapas e as implicações teóricas, práticas e sociais; por fim, as considerações finais, que retomam as principais discussões em culminância com os objetivos e apresentam lacunas que podem ser preenchidas por futuras pesquisas.

2 FUNDAMENTAÇÃO TEÓRICA

Esta seção mostra os fundamentos teóricos que embasaram esta pesquisa. Estudos sobre a Tradução Audiovisual Legendagem, sobre os parâmetros da legendagem e sobre os aportes teóricos-metodológicos das pesquisas descritivas, exploratórias e experimentais.

2.1 TRADUÇÃO AUDIOVISUAL LEGENDAGEM

A tradução audiovisual (TAV) é uma subárea dos Estudos da Tradução que passou a ganhar espaço no final dos anos 1980 com a popularização dos vídeos domésticos em VHS (*Video Home System*). Díaz Cintas (2009, p. 1) observou que nos 20 anos antes do início da pesquisa acadêmica em TAV a indústria audiovisual gerou um campo fértil de dados para esse tipo de pesquisa. E que, além de crescer enquanto atividade profissional, graças, principalmente, à revolução digital, a TAV tornou-se uma área estabelecida e de destaque na pesquisa acadêmica, muito embora o reconhecimento do nome tenha levado mais tempo para acontecer. A indefinição do nome da área foi tema de Franco e Araújo (2011) no artigo sobre as questões terminológicas da TAV, no qual explicam que a abrangência do termo *audiovisual* era confusa na taxonomia proposta por Yves Gambier em 2003 no volume especial do periódico *The Translator*, intitulado *Screen Translation*. Isso porque havia inconsistência entre os nomes e os tipos que representavam, o que demonstrava a indefinição terminológica do autor.

No entanto, a abrangência do termo *audiovisual* em Díaz Cintas (2005) deixa claro que o meio audiovisual inclui todos os espaços onde há um sinal acústico e um sinal visual e expõe que inicialmente a TAV foi usada para encapsular práticas de tradução diferentes usadas nas mídias audiovisuais cinema, televisão e VHS, nas quais há transferência de uma língua fonte para uma língua meta, como dublagem, legendagem, *voice-over*, dublagem parcial, narração e interpretação, além da *surtitling* (supra-legendagem) para teatro e ópera, que veio depois, assim como a legendagem para surdos e ensurdecidos (LSE) e a audiodescrição (AD), levando em consideração o novo cenário de acessibilidade na taxonomia de TAV, mesmo não sendo feitas entre duas línguas diferentes. O autor argumenta que, assim como a LSE e a AD, legendar para ouvintes, dublar e traduzir em *voice-over* compartilham a mesma ideia de acessibilidade. Em consonância com o proposto por Díaz Cintas (2005), Franco e Araújo (2011) consideram como modalidades de TAV a legendagem

para ouvintes, tema desta pesquisa, a LSE, a legendagem eletrônica ou *surtitling*, a dublagem, o *voice-over* e a audiodescrição.

Sobre a legendagem, tema desta pesquisa, Franco e Araújo (2011) afirmam que, no Brasil, a primeira tentativa de propor uma terminologia para a prática partiu de Alvarenga (1998), que não diferenciava a legendagem para ouvintes da legendagem para surdos, acreditando que “legendagem” referia-se apenas à legenda interlinguística. A partir de uma concepção mais abrangente da legendagem, as autoras propõem terminologias diferentes, tendo em vista o crescimento das áreas de estudo e a necessidade de demarcá-las, e diferenciam a legendagem para ouvintes (LO) da legendagem para surdos e ensurdecidos (LSE). Em geral, essa diferença reside no público-alvo e na tradução de elementos dependentes do canal auditivo, como efeitos sonoros e identificação de falantes, no caso das legendas para surdos.

No sentido amplo e sobre os aspectos linguísticos de uma tradução, Jakobson (1959) a classifica em três tipos: tradução interlinguística – ocorre entre línguas diferentes; tradução intralinguística – ocorre dentro da mesma língua; e tradução intersemiótica – ocorre entre meios semióticos diferentes (como do meio acústico para o verbal escrito). Essa taxonomia proposta por Jakobson, antes mesmo do reconhecimento dos estudos da tradução enquanto área de pesquisa (HOLMES, 1972), permanece bastante atual e contribui sobremaneira para a inserção de diferentes práticas de tradução nos estudos da tradução e para o reconhecimento daquelas que ocorrem dentro da mesma língua, a exemplo da LSE em português de filmes brasileiros, e daquelas que ocorrem entre meios semióticos diferentes, como a audiodescrição.

Franco e Araújo (2011) explicam que a legendagem para surdos e ensurdecidos (LSE) é a tradução do inglês SDH (*Subtitling for the deaf and the hard-of-hearing*). Esse termo foi sugerido por Selvatici (2010) e por membros da sua banca de mestrado. As autoras esclarecem que legenda para surdos não é o mesmo que *closed caption*, como muitos pensam. *Closed caption* é um sistema de *legendagem fechada ou oculta* (o espectador precisa acessá-la no seu controle remoto do aparelho de TV ou DVD) em oposição à *legendagem aberta* (a legenda aparece sem precisar acionar). É um modelo norte-americano de legendagem usado em muitos países, inclusive no Brasil. Já a legendagem eletrônica ou *surtitling* (BARTOLL, 2011) é constituída de legendas projetadas em teatros e mostras de cinema. Também é usada em óperas, por isso o nome *opera titling* (FREDDI; LURAGGI, 2011).

Na definição de dublagem, as autoras asseveram que é uma das modalidades de TAV mais conhecidas pelos brasileiros porque, apesar de ter sua origem no rádio, foi o modo

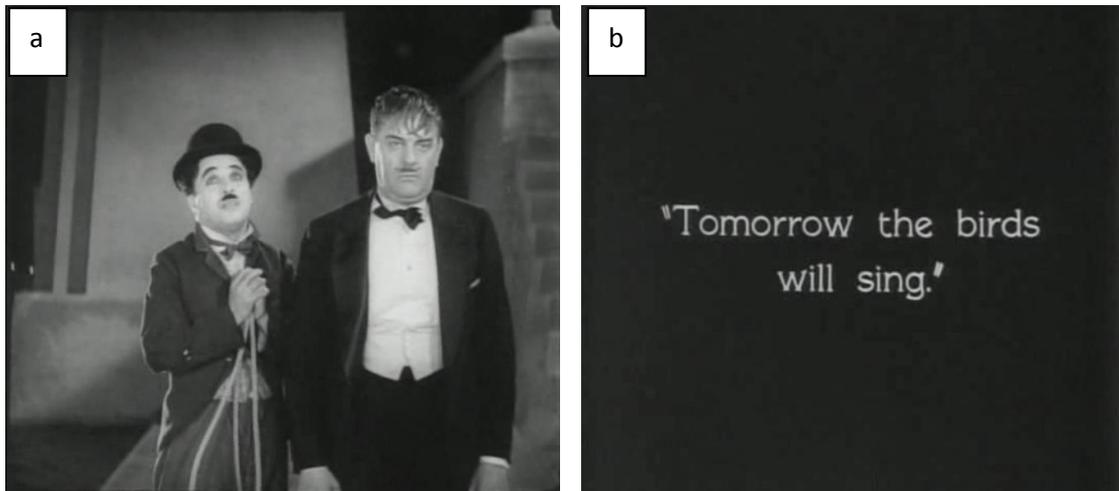
escolhido e amplamente divulgado pela televisão brasileira para divulgar a importação de filmes estrangeiros, especialmente os americanos, outros programas de ficção, como os *sitcoms*, e, mais recentemente, até programas que não se encaixam no rótulo ficcional (FRANCO; ARAÚJO, 2011). A dublagem pode ser definida por três características básicas: 1) uma tradução interlingual de um discurso oral para outro discurso oral, das falas dos personagens de um filme ou programa de ficção pré-gravado; por esse motivo, é também chamada de “revocalização”; 2) elimina a presença do discurso oral estrangeiro; 3) é regida pelo sincronismo labial, fundamental para fazer com que o público-alvo acredite na ilusão de que o personagem esteja falando na sua língua. Por causa disso, a dublagem é às vezes chamada, em países de língua inglesa, de *lip-sync translation* ou simplesmente de *lip-sync* (tradução de sincronismo labial).

Sobre o *voice-over*, as autoras explicam que, assim como a dublagem, é pré-gravado e representa a revocalização de um discurso oral em língua estrangeira para um discurso oral na língua da tradução, mas sob uma perspectiva diferente. As autoras esclarecem que o *voice-over* se caracteriza: 1) por ser uma tradução audiovisual interlingual usada para produtos audiovisuais estritamente pertencentes ao gênero de não ficção, tais como documentários, noticiários com entrevistas pré-gravadas, programas de entrevistas pré-gravados, debates políticos etc.; 2) pela ausência de sincronismo labial, embora seja de extrema importância a sincronia com a duração do discurso em língua estrangeira e com a imagem; 3) pela copresença do discurso em língua estrangeira, em volume baixo, ao qual se sobrepõe o discurso em língua de chegada (FRANCO; ARAÚJO, 2011).

Sobre a modalidade de TAV acessível (TAVa) audiodescrição, Franco e Araújo (2011) compreendem que AD é a tradução em palavras das impressões visuais de um objeto, seja ele um filme, uma obra de arte, uma peça de teatro, um espetáculo de dança ou um evento esportivo. O recurso tem o objetivo de tornar esses produtos acessíveis a pessoas com deficiência visual e pode ser pré-gravado ou ao vivo.

Tais discussões ajudaram a contextualizar o tema e o objeto de estudo desta tese: a legendagem para ouvintes de filmes em DVD, que é intralinguística na medida em que ocorre da língua oral inglesa para a língua escrita inglesa e é interlinguística na medida em que ocorre da língua oral inglesa para a língua escrita portuguesa. Dos tipos de TAV, a legendagem é a mais antiga, concebida desde o início do cinema, no início do século XX. As legendas eram chamadas de intertítulos (*intertitles*) e apareciam entre as sequências dos filmes (IVARSSON; CARROLL, 1998). Os intertítulos eram geralmente desenhados em cartazes ou placas, filmados e inseridos na montagem do filme, conforme figura 01.

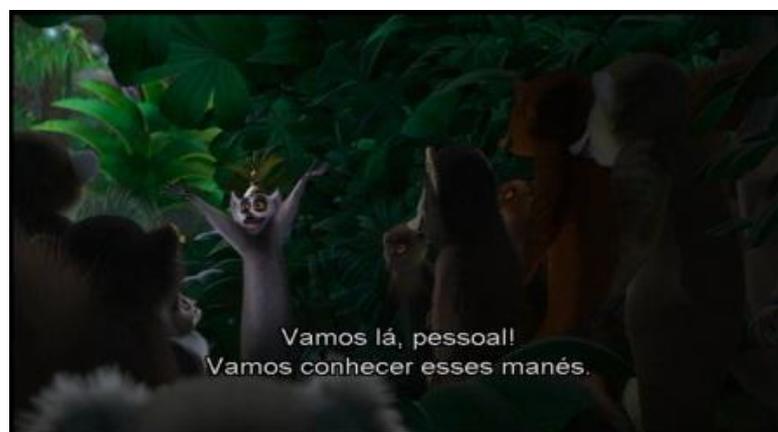
Figura 01 – Exemplo de intertítulo (*intertitle*). Letra a: fotograma que antecede o intertítulo. Letra b: intertítulo que diz, em português: “Amanhã os pássaros cantarão”



Fonte: Fotograma do filme *City Lights* (CHAPLIN, 1931).

Como mostra a figura 01, os intertítulos eram inseridos para explicar e/ou comentar as cenas, ou mesmo para funcionar como alguma fala importante do filme, e não eram sobrepostos às imagens. Nessa época, os filmes eram bastante acessíveis para surdos e ouvintes, pois toda a narrativa era pensada a partir do visual, e não do áudio. Os intertítulos eram facilmente traduzidos e gravados em outras línguas, mas, com o advento do cinema falado, as dificuldades começaram a aparecer para surdos e ouvintes. De um lado, os filmes demandavam a tradução dos diálogos, e, de outro, a mudança também foi positiva porque, nesse momento, houve expansão do cinema. Logo, a legendagem passou a ser sobreposta às imagens, como ainda acontece (vide figura 02).

Figura 02 – Legenda sobreposta à imagem



Fonte: Elaborada pela autora. Fotograma do filme *Madagascar* (2005).

A legendagem para ouvintes, exclusividade do cinema, tornou-se popular entre o público geral com a indústria de *home video*, no final dos anos 1980. A partir de então, novas mídias e produtos audiovisuais surgiram, e, com eles, novos tipos de legenda. A exemplo da legendagem de fãs, que, segundo Díaz Cintas e Sánchez (2006), nasceu em meados da década de 1990, a partir da carência dos fãs estadunidenses para entender animes japoneses que não eram comercializados no país. Outro exemplo é a legendagem para surdos, que, no Brasil, nasceu a partir das normas gerais e critérios básicos para promoção da acessibilidade dispostos na Lei nº 10.098, conhecida como Lei da Acessibilidade (2000). Com o surgimento de novos produtos audiovisuais, vários tipos de legenda foram tomando forma e ampliando o escopo da tradução audiovisual enquanto área de pesquisa. Dessa forma, apresento uma proposta inicial de mapa da legendagem como forma de elucidar os campos de atuação específicos da legendagem, ampliar as discussões na área e caracterizar o tipo de legenda.

Figura 03 – Mapa da legendagem demarcando os tipos de legenda

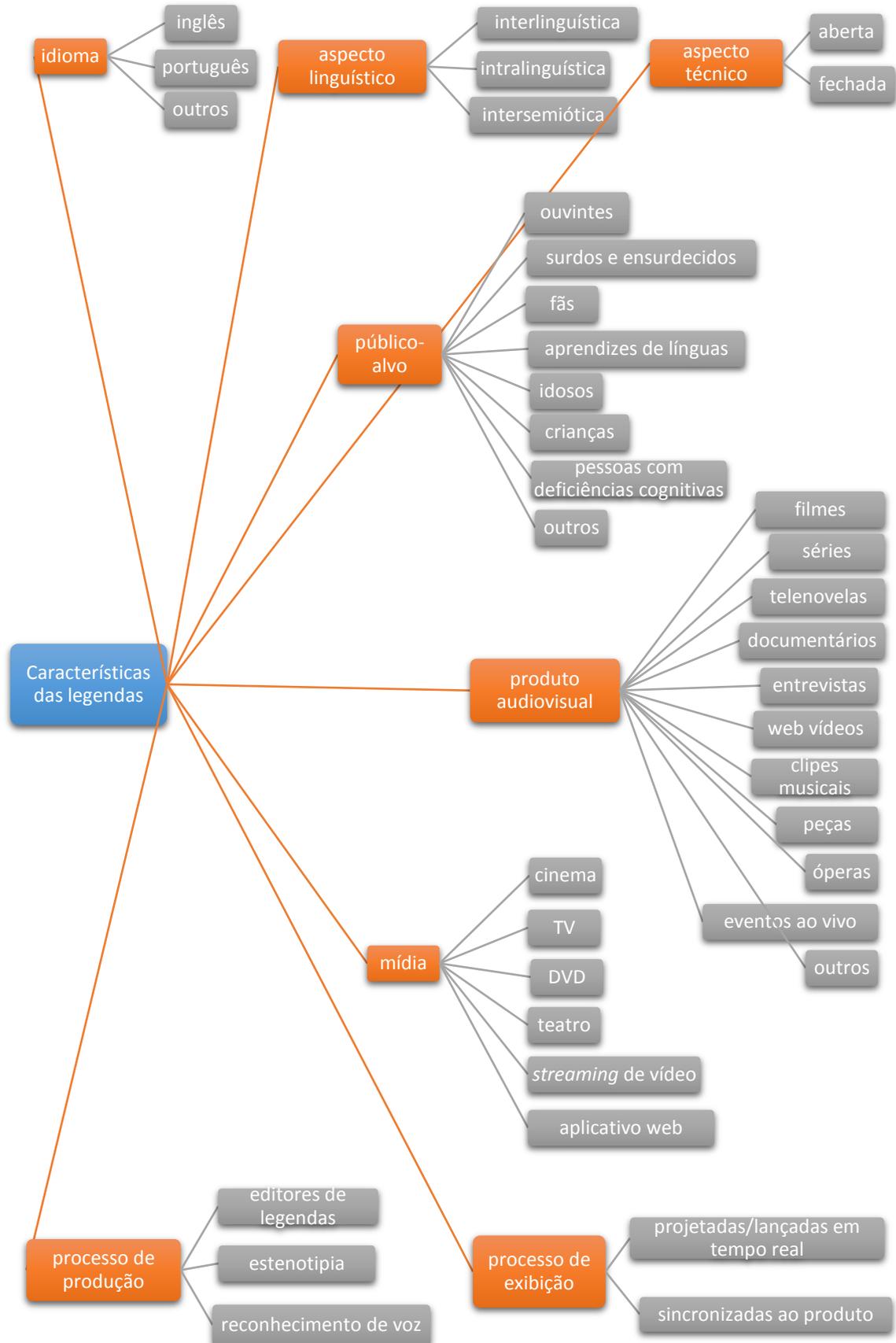


Fonte: Elaborada pela autora.

Os tipos de legenda apresentados são: legendagem para ouvintes, legendagem para surdos e ensurdidos, legendagem amadora (*fansubs*), legendagem eletrônica (*surtitling*), legendagem refalada (*respeaking*) e audiolegendagem. Cada tipo de legendagem envolve um processo específico que pode ser detalhado pelas características a seguir.

Dentro dessa proposta, cada tipo de legenda pode apresentar características próprias, semelhantes e/ou diferentes umas das outras. Quanto ao público-alvo, as legendas podem ser para ouvintes, surdos, fãs, aprendizes de línguas, idosos, crianças, pessoas com deficiência cognitiva etc. Quanto ao idioma: inglês, português, bilíngue, multilíngue, entre outros. Quanto ao aspecto linguístico, podem ser interlinguísticas, intralinguísticas ou intersemióticas. Quanto ao aspecto técnico visual, podem ser abertas ou fechadas (*roll up ou pop on*). Quanto ao produto audiovisual, podem ser de filmes, séries, telenovelas, documentários, web vídeos, clipes musicais, peças, óperas, entre outros. Quanto à mídia, podem ser para: cinema, TV, teatro, DVD, *streaming*, aplicativo *web* (*movie reading, whatscine*), entre outros. Quanto ao processo de confecção, podem ser feitas por programas editores de legendas, estenotipia e reconhecimento de voz. Quanto à exibição, podem ser projetadas ou lançadas em tempo real e podem ser sincronizadas ao produto audiovisual. A figura 04 resume as características das legendas.

Figura 04 – Esquema de caracterização das legendas



Fonte: Elaborada pela autora.

Dadas as características das legendas, as analisadas nesta pesquisa, em inglês e em português, são do tipo: legendagem para ouvintes, de filmes em DVD. Quanto ao aspecto técnico, são fechadas; quanto ao aspecto linguístico das legendas em inglês, são intralinguísticas, e das legendas em português, são interlinguísticas. São produzidas por programas editores de legendas, sincronizadas ao produto.

2.2 PARÂMETROS DA LEGENDAGEM

A legendagem é uma modalidade de tradução audiovisual (TAV) que consiste na tradução de falas e de informações visuais e sonoras de uma produção audiovisual em texto escrito. O produto da legendagem consiste em unidades de legendas que são inserções de texto de uma ou duas linhas. Quanto à finalidade, as legendas podem servir de tradução, de ferramenta para o ensino e aprendizagem de línguas estrangeiras, de acesso a produtos audiovisuais sem áudio ou com volume baixo, de acompanhamento de letras de música em karaokê, entre outros fins. No que diz respeito ao público-alvo, as legendas podem ser direcionadas, mas não restritas, ao público de pessoas ouvintes e de pessoas surdas e ensurdecidas.

Sobre os parâmetros da legendagem, Díaz Cintas e Remael (2007) descrevem parâmetros técnicos e linguísticos, além de questões tradutórias, tomando como base a legendagem para ouvintes na Europa. Os autores entendem que os parâmetros técnicos da legendagem são elementos de dimensões espacial e temporal, como o número de linhas da legenda, o número de caracteres por linha, a posição das legendas no produto audiovisual, a velocidade, a duração, a sincronização, entre outros, e entendem que os parâmetros linguísticos da legendagem são: redução textual, coesão e coerência e segmentação (entre legendas e entre linhas de uma legenda). Os autores (DÍAZ CINTAS; REMAEL, 2007) também discutem sobre questões tradutórias da legendagem, como a variação linguística, o discurso marcado, os termos culturais, a tradução de músicas e a tradução de humor, e sobre questões ideológicas.

Mesmo tomando como base a legendagem para ouvintes na Europa, os parâmetros descritos por Díaz Cintas e Remael (2007) ajudaram a fomentar pesquisas no mundo inteiro e serviram para explicar fenômenos tanto na legendagem para ouvintes quanto na legendagem para surdos e ensurdecidos. O que os autores fizeram foi propor uma sistematização dos parâmetros da legendagem a partir de estudos prévios (D'YDEWALLE, 1987; GOTLIEB, 1994; IVARSSON; CARROLL, 1992, 1998; KARAMITROGLOU, 1998; DE LINDY; KAY,

1999) e orientações sobre as práticas no contexto europeu. Essa sistematização dos parâmetros da legendagem ajudou, sobretudo, a diferenciá-la das demais modalidades de tradução, contribuindo para ampliar seu status científico.

As discussões sobre legendagem eram feitas a partir de regras formuladas por estudiosos e profissionais do mercado da legendagem, e o resultado era geralmente prescritivo, apresentado como soluções práticas para os problemas e/ou orientações subjetivas com foco no produto. A falta de um aporte teórico levou estudiosos a buscarem reflexões mais consistentes que pudessem ampliar as discussões teórico-metodológicas da área, visando alcançar diversos contextos de produção, recepção e exibição de legendas. Nesse sentido, estudos sobre a legendagem passaram a analisar fenômenos a partir dos pontos de vista linguístico, gramatical ou dos estudos da tradução, fazendo interface com áreas como Linguística Sistêmico-Funcional, Multimodalidade e Linguística de *Corpus*, aderindo a concepções gramaticais funcionalistas e gerativistas e a conceitos dos estudos descritivos da tradução e da Linguística de *Corpus*. A interação dos estudos da legendagem com áreas do conhecimento mais consolidadas contribuiu para o seu crescimento científico, mas ainda era preciso relacionar essas áreas aos parâmetros da legendagem, que respondem a questões específicas e dão conta de uma série de restrições próprias dessa modalidade de tradução, justificando eventos como: a redução das falas ou explicitação de informações não contidas nelas, legendas de uma ou duas linhas, números mínimo e máximo de caracteres por linha, segmentações retórica, visual e linguística, entre outras características e fenômenos dessa modalidade de tradução audiovisual que a diferencia das demais.

Nesse sentido, os pesquisadores do Laboratório de Tradução Audiovisual – LATAV da UECE, do grupo LEAD, têm contribuído para a consolidação dos parâmetros da legendagem no Brasil, uma vez que costumam fazer uma interface multidisciplinar entre os parâmetros da legendagem e as áreas com as quais se relacionam, a exemplo das pesquisas exploratórias (Projeto MOLES), descritivas (Projeto CORSEL) e experimentais (Projeto EXLEG) realizadas desde 2002. Os pesquisadores do grupo LEAD entendem que a operacionalização dos parâmetros resulta no aspecto mais inerente à legendagem, o aspecto tradutório, contribuindo para o delineamento teórico-metodológico da legendagem com base em reflexões sobre os parâmetros técnicos e linguísticos em interface multidisciplinar com outras áreas de estudos.

No Brasil, as primeiras pesquisas que analisaram os parâmetros da legendagem se deram com a recepção de legendas para surdos e ensurdecidos, para verificar se as legendas fechadas (*closed caption*) exibidas pela TV brasileira atendiam às necessidades dos surdos

brasileiros. A primeira pesquisa, com surdos do Instituto Cearense de Educação de Surdos (ICES) da cidade de Fortaleza e ouvintes que assistiram com volume abaixado (FRANCO; ARAÚJO, 2003; ARAÚJO, 2004a; ARAÚJO, 2005; ARAÚJO, 2007), demonstrou que tanto os surdos quanto os ouvintes não conseguiram acompanhar as legendas porque as mesmas eram rápidas e sem sincronismo com fala e imagem, sugerindo que a condensação e a edição das legendas são elementos fundamentais para torná-las mais eficazes. Em um segundo momento da referida pesquisa, os mesmos trechos receberam nova legendagem, sendo condensadas e sincronizadas com as falas e as imagens. Os resultados mostraram que os surdos tiveram mais sucesso ao assistir aos programas; contudo, novas pesquisas precisariam ser realizadas para investigar o nível de edição e o formato que melhor atenderiam às necessidades de surdos do Brasil.

A segunda pesquisa (ARAÚJO, 2008), com surdos do Centro de Atendimento às Pessoas com Surdez (CAS) de Fortaleza, investigou os níveis de condensação e formato para definir um modelo de legendagem que desse aos surdos brasileiros acesso ao lazer e à informação proporcionados pela televisão. Ao longo de oito meses, tendo um encontro mensal, os surdos assistiram a programas legendados com o modelo *closed caption* da Rede Globo, que transcreve todas as falas sem editar o texto, e com o modelo semelhante ao de ouvintes em três níveis de edição diferentes: 16, 15 e 14 caracteres por segundo. Além dos quatro diferentes níveis de edição, outros parâmetros foram testados, como cor, tamanho, fonte e informações sonoras. No último encontro, os surdos foram expostos também ao formato de legenda português (europeu) para que pudessem ter contato com outros formatos antes de definir o modelo que melhor atendesse às necessidades da comunidade surda.

Os resultados chegaram a um modelo de legendagem definido pelos surdos cearenses, posteriormente testado com surdos de outras regiões do Brasil: o nível de condensação preferido foi o de 14 caracteres por segundo, cor amarela sem tarja, identificação de falante sempre que houvesse mudança de turno, tradução de efeitos sonoros sempre feita, mesmo que a ação fosse visível na tela, uso de norma culta do português, e ritmos musicais não precisariam ser discriminados, a não ser que seja música tema. A essa época, o parâmetro de velocidade parecia ser o que mais influenciava na recepção das legendas, e, de acordo com os surdos de Fortaleza, as legendas mais condensadas, as que continham apenas 14 caracteres por segundo, eram as que proporcionavam melhor recepção e compreensão de conteúdo.

O mesmo não foi observado no projeto MOLES, que testou tal formato definido pelos surdos cearenses em legendas com velocidades de 180, 160 e 145 palavras por minuto (ppm), ou 18, 16 e 14 caracteres por segundo (cps), respectivamente, com 34 surdos

de 4 regiões do Brasil. Os resultados mostraram que os surdos brasileiros conseguem ler e compreender o conteúdo dos filmes com legendas nas três velocidades. Contrariando o consenso e as pesquisas anteriores, a velocidade mais alta, de 180ppm, foi a que proporcionou melhor recepção, possibilitando aos surdos apreciarem legendas e imagens, além de captar detalhes sobre o conteúdo. A principal explicação de tal resultado, além do cumprimento dos parâmetros técnicos e linguísticos da legendagem, deve-se à segmentação, haja visto que todos os parâmetros foram controlados nesta pesquisa, mas nas anteriores a segmentação não havia sido analisada com afinco. Os resultados dessas primeiras pesquisas abriram lacunas no que diz respeito aos parâmetros da legendagem, o que levou a uma série de pesquisas com desenhos metodológicos diferenciados para a investigação das características em diferentes tipos de legenda.

As pesquisas ora citadas têm demonstrado que a harmonia dos parâmetros é característica de uma boa legenda, o que justifica a importância de conhecer os parâmetros da legendagem e compreender as relações estabelecidas entre eles. O entendimento dos parâmetros é imprescindível para acompanhar a análise dos dados desta pesquisa. Os parâmetros da legendagem levados em conta nesta pesquisa são os técnicos: formato, marcação, número de linhas, número de caracteres por linha, velocidade da legenda, e o linguístico: segmentação. Esses parâmetros costumam ser discutidos na pesquisa em legendagem e também estão presentes no Guia de LSE para produções audiovisuais brasileiras (ARAÚJO; CHAVES, 2016), doravante Guia, no qual é descrita uma legenda (para surdos ou ouvintes) que seria aquela preconizada por pesquisadores da área tanto nacionais quanto internacionais.

O formato é um parâmetro técnico da legenda e pode aparecer de três formas: na primeira, são exibidos quase o mesmo número de caracteres para as duas linhas, tendo um formato semelhante a um retângulo. No segundo e no terceiro, aparecem mais caracteres na linha de cima ou na de baixo, respectivamente, lembrando uma pirâmide (ARAÚJO; ASSIS, 2014) (Vide quadro 01).

Quadro 01 – Formatos de legenda

Opções de Formato	Legendas
<p>Em forma de retângulo</p>	 <p>[Leo] Queria ver se você tivesse que resolver equação em Braille.</p>
<p>Em forma de pirâmide invertida com a linha de cima maior</p>	 <p>[Gi] Você prefere pisar descalço num formigueiro</p>
<p>Em forma de pirâmide com a linha de cima menor</p>	 <p>[Gi] Devia aceitar. Em duas semanas tem prova.</p>

Fonte: Elaborado pela autora. Fotogramas do filme *Eu não quero voltar sozinho* (2010).

O formato é um parâmetro levado em consideração especialmente na hora de segmentar uma legenda. Segundo Ivarsson e Carroll (1998), a linha de cima deve ser preferencialmente menor para que haja mais área de imagem do filme preservada. O formato deve ser definido em prol da segmentação linguística, que tem como fim manter a integridade dos sintagmas, e, para Díaz Cintas e Remael (2007), a segmentação linguística deve ser priorizada, e não o aspecto estético da legenda.

A marcação de uma legendagem é um parâmetro técnico e consiste na determinação dos tempos de entrada e saída de cada legenda. Segundo orientações do Guia (2016), a marcação de um diálogo deve tentar seguir os ritmos de fala do filme e a performance dos atores, tendo em mente as pausas, as interrupções e outros elementos prosódicos que caracterizam as falas de um filme. Períodos longos, na maioria das vezes, poderão ser divididos em várias legendas, enquanto que períodos curtos podem ser agrupados para evitar o estilo telegráfico (DÍAZ CINTAS; REMAEL, 2007).

Uma boa marcação de legenda acontece quando se consegue obter o exato sincronismo entre as falas do filme e as legendas. Essa sincronização pode ser realizada com o auxílio de um cronômetro, conhecido na área como TCR (*Time Code Reader*), que localiza as falas do filme pelas horas, minutos, segundos e pelos *frames* ou quadros, conforme figura 05.

Figura 05 – Exemplo de TCR



Fonte: Elaborada pela autora. Fotograma do filme *Madagascar* (2005).

Pelo TCR, é possível saber que o filme está com oito minutos e vinte e cinco segundos de duração, e no fotograma número dez. Os fotogramas marcam a velocidade do filme, que no caso do cinema é, normalmente, de vinte e quatro quadros por segundo, enquanto na TV e no DVD giram em torno de 29 quadros por segundo.

Segundo o Guia (ARAÚJO; CHAVES, 2016), as empresas de legendagem do mundo inteiro adotam a mesma estratégia no que diz respeito ao número de linhas, ou seja, são utilizadas, no máximo, duas linhas, as quais devem ter, no máximo, 37 caracteres cada uma. Essa é a medida usada na Europa na chamada regra dos seis segundos. D'Ydewalle *et al.* (1987) testaram essa regra experimentalmente, e seus resultados comprovaram essa pressuposição. Embora o sistema americano de *closed caption* utilizado para legendas da TV no Brasil possibilite a exibição de uma legenda em mais de duas linhas para LSE, as autoras corroboram a posição de d'Ydewalle *et al.* (1987), porque com mais de duas linhas a

legendagem pode prejudicar o movimento de deflexão, no qual o espectador lê as legendas e olha as imagens para poder harmonizá-las e, assim, assistir confortavelmente a uma produção audiovisual. Qualquer esforço maior nessa leitura pode prejudicar essa harmonização.

D'Ydewalle *et al.* (1987) explicam que as legendas de uma linha geraram mais relatos de problemas que as legendas de duas linhas. Os autores acreditam que um dos motivos para esses problemas é que os canais de televisão apresentaram legendas de uma linha maiores (acima de 32c) por um tempo maior que o adequado para a regra dos seis segundos (acima de 6s). Assim, embora as legendas de uma linha apresentem mais perturbações na sua recepção, parece que as perturbações foram motivadas não apenas pelo formato de uma linha, mas pelo excesso de palavras e pela longa permanência em tela, fatores que tornam a legenda lenta. Isso, a meu ver, reforça a influência que os parâmetros exercem sobre a qualidade das legendas e da recepção e reforça o argumento desta tese de que a harmonia dos parâmetros contribui para a qualidade das legendas e da recepção, tanto de uma quanto de duas linhas.

Na literatura sobre o processamento de legendas, legendas de duas linhas, embora levem mais tempo para serem processadas, são processadas fluentemente, por terem maior conteúdo informativo (D'YDEWALLE; DE BRUYCKER, 2003). As legendas de uma linha, por sua vez, são processadas em menos tempo, mas deixam de ser fixadas com frequência, porque apresentam menor conteúdo informativo. Por um lado, pesquisadores acreditam que legendas de duas linhas são processadas fluentemente por causa do maior conteúdo informativo presente em duas linhas; por outro lado, a densidade das informações contidas em uma linha de legenda, no caso das campanhas políticas, parece causar desconforto na recepção de surdos e ouvintes (MONTEIRO, 2016).

Especialmente no caso das legendas de uma linha das campanhas políticas no Brasil, Monteiro (2016) observou que as legendas de uma linha foram recebidas com desconforto pela maioria dos participantes surdos e ouvintes. A autora considera que as legendas de uma linha na velocidade rápida exijam mais atenção dos espectadores quanto à completude do conteúdo.

No experimento feito por De Linde e Kay (1999) com participantes surdos e ouvintes, a velocidade da legenda teve efeito sobre o tempo de leitura, sobre o número de palavras por fixações e sobre a releitura. Quanto à velocidade, esta foi obtida por uma média das legendas dos vídeos. Aquelas consideradas rápidas (139ppm) foram processadas em menos tempo, enquanto as legendas mais lentas (74ppm) foram processadas em mais tempo (300ms). O maior efeito da releitura foi expresso em duas legendas de uma única palavra

[applause], cujo tempo de exposição foi de 0,6s e o tempo de leitura foi de 0,43s. O tempo de processamento de uma legenda de única palavra foi maior que o tempo de processamento de legendas de 139ppm e 74ppm. Dessa forma, é correto afirmar que os autores encontraram muito mais releitura em legendas com longa duração, como, por exemplo, em uma legenda de 8,5s. Em resumo, há dois fatores que dificultam o processamento e que levam à releitura de legendas. São as legendas lentas, de 74ppm, e legendas com muito tempo de exposição. Entendo que o tempo de exposição em excesso gera uma legenda de velocidade muito lenta, e esse acaba sendo o único causador da releitura.

Diferentemente dos outros autores que analisaram o processamento de legendas, Vieira (2016) e Monteiro (2016) controlaram as velocidades testadas de 145 e 180ppm nos experimentos sobre o processamento de legendas com e sem problemas de segmentação, para surdos e ouvintes. Os resultados também demonstraram releitura das legendas na velocidade de 145ppm.

O Guia (CHAVES; ARAÚJO, 2016) afirma que, para a elaboração adequada de uma legenda, é preciso que a velocidade de leitura de uma legenda seja compatível com a velocidade da fala que ela traduz (DÍAZ CINTAS; REMAEL, 2007). Os padrões de velocidade utilizados (145, 160 e 180ppm) advêm da regra dos 6 segundos, que é empírica, sem base científica, sendo a regra geral adotada por emissoras de TV em muitos países, que usavam a mesma regra geral para a marcação de legendas. D'Ydewalle *et al.* (1987) explicam que, pela regra, duas linhas juntas de até 32 caracteres cada poderiam ser usadas uma por vez, ou seja, virariam duas legendas de 1 linha. E se fossem duas linhas juntas que somadas tivessem 32 caracteres, essa legenda de duas linhas seria exibida por 6 segundos. Depois disso, as legendas passaram a ser marcadas proporcionalmente por essa regra. Os autores ainda enfatizaram que ninguém parece saber de onde veio essa regra.

D'Ydewalle *et al.* (1987) montaram um experimento procurando saber qual combinação de legenda (1 ou 2 linhas vs tempo de exibição de 2, 4 ou 6 segundos) os participantes julgavam apropriada e explicaram que a velocidade da legenda era algo empírico e a regra dos 6 segundos era a referência utilizada pela maioria das emissoras de TV, em que legendas de 1 linha continham sentenças com até 32 caracteres e legendas de duas linhas continham sentenças de 33 a 64 caracteres. Ao final do experimento, os pesquisadores verificaram que a melhor legenda foi a de 2 linhas em 6 segundos.

Díaz Cintas e Remael (2007), por sua vez, elaboraram uma tabela com auxílio de uma fórmula usada pelo programa de legendagem utilizado no Reino Unido, WinCAPS com medidas proporcionais para cada segundo/*frame*, como forma de dar suporte à prática da

legendagem. Eles partiram da regra dos seis segundos, mas entendiam que o máximo de caracteres por linha poderia ser 35 ou 37, em vez de 32, que também era uma medida empírica utilizada na prática da legendagem adotada por algumas empresas.

Díaz Cintas e Remael (2007) basearam seus cálculos em 37 caracteres por linha. Assim, uma legenda cheia, de duas linhas, exibida em seis segundos poderá conter até 74 caracteres. Ao dividir o total de caracteres pelo total de segundos, é possível obter a média de 12 caracteres por segundo. Essa média foi escalonada pela fórmula do WinCAPS, que resultou na tabela de 145 palavras por minuto. Os autores consideram que essa tabela de 145ppm pode ser adequada para televisão, mas não para cinema e DVD, e, por isso, propuseram mais duas tabelas, uma de 160ppm e outra de 180ppm, pensando em cinema e DVD, respectivamente, com a justificativa de que essas mídias aceitam mais caracteres por segundo, pelo tamanho do telão do cinema e pela facilidade de parar e iniciar o DVD a qualquer momento, e também para atender públicos com velocidades de leitura maiores que 145ppm. Para que essas velocidades conversassem com a linguagem dos programas editores de legendas, mesmo se tratando de medidas aproximadas, as legendas eram pensadas em caracteres por segundo, o que ajudava a fazer uma marcação de legendas mais eficiente e mais sistemática (DÍAZ CINTAS; REMAEL, 2007).

Quadro 02 – Relação caracteres por segundo na regra europeia

145 palavras por minuto	Segundos: <i>frames</i>	Caracteres	Segundos: <i>frames</i>	Caracteres	
	01:00	16	02:00	29	
	01:04	17	02:04	32	
	01:08	18	02:08	34	
	01:12	20	02:12	36	
	01:16	23	02:16	38	
	01:20	25	02:20	40	
Segundos: <i>frames</i>	Caracteres	Segundos: <i>frames</i>	Caracteres	Segundos: <i>frames</i>	Caracteres
03:00	44	04:00	58	05:00	71
03:04	46	04:04	60	05:04	71
03:08	48	04:08	62	05:08	73
03:12	50	04:12	64	05:12	73
03:16	52	04:16	65	05:16	74
03:20	54	04:20	67	05:20	74
160 palavras por minuto	Segundos: <i>frames</i>	Caracteres	Segundos: <i>frames</i>	Caracteres	
	01:00	17	02:00	31	
	01:04	18	02:04	34	
	01:08	20	02:08	37	
	01:12	23	02:12	40	
	01:16	26	02:16	42	
	01:20	28	02:20	44	
Segundos: <i>frames</i>	Caracteres	Segundos: <i>frames</i>	Caracteres	Segundos: <i>frames</i>	Caracteres
03:00	48	04:00	63	05:00	75
03:04	50	04:04	65	05:04	75
03:08	53	04:08	67	05:08	76
03:12	56	04:12	69	05:12	76
03:16	58	04:16	71	05:16	77
03:20	60	04:20	73	05:20	77
				06:00	78
180 palavras por minuto	Segundos: <i>frames</i>	Caracteres	Segundos: <i>frames</i>	Caracteres	
	01:00	17	02:00	35	
	01:04	20	02:04	37	
	01:08	23	02:08	39	
	01:12	26	02:12	43	
	01:16	28	02:16	45	
	01:20	30	02:20	49	
Segundos: <i>frames</i>	Caracteres	Segundos: <i>frames</i>	Caracteres	Segundos: <i>frames</i>	Caracteres
03:00	53	04:00	70	05:00	78
03:04	55	04:04	73	05:04	78
03:08	57	04:08	76	05:08	78
03:12	62	04:12	76	05:12	78
03:16	65	04:16	77	05:16	78
03:20	68	04:20	77	05:20	78
				06:00	78

Fonte: Adaptado de Díaz Cintas e Remael (2007, p. 97-99).

O quadro 02 mostra a relação de caracteres nas velocidades de 145, 160 e 180ppm. Por exemplo, para uma fala que tem um segundo e aparece no fotograma 20, ela

pode ter até 25, 28 ou 30 caracteres, se a velocidade desejada for de 145, 160 ou 180ppm, respectivamente. Vieira (2016) observou que legendas de duas linhas bem segmentadas são melhor recebidas na velocidade de 180ppm e que legendas de 145ppm foram mal recebidas por surdos, causando releitura da legenda.

A segmentação diz respeito à divisão das falas em blocos semânticos e sintáticos e deve ser feita de modo que cada legenda seja facilmente compreendida no curto espaço de tempo em que é exibida (DÍAZ CINTAS; REMAEL, 2007, p. 172). As falas são distribuídas em legendas, e a segmentação dessas falas se dá entre as linhas de uma mesma legenda, ou seja, na quebra da linha, e entre legendas diferentes. As orientações para segmentar entre linhas e entre legendas são as mesmas, e, segundo Reid (1990), essa segmentação é realizada pelo visual (com base nos cortes das cenas), pela retórica (com base no fluxo da fala) e pela linguística (com base nas unidades semânticas e sintáticas). Não há uma ordem para realizar a segmentação; os três tipos podem ser priorizados de acordo com as necessidades do tradutor e do vídeo. Embora estejam mostrados os três tipos de segmentação, para fins desta análise, apenas a segmentação linguística foi levada em consideração.

A segmentação visual relaciona-se à distribuição do texto com base nos cortes das cenas e, por isso, está mais relacionada à marcação do que a questões linguísticas. Esse tipo de segmentação deve levar em consideração os cortes da câmera, porque, quando há mudança de cena, nossos olhos reconhecem essa mudança e esperam que haja mudança também de legenda. Mas se, ao fazer a deflexão da imagem para a legenda, esta for a mesma da cena anterior, levaremos mais tempo para recalcular que aquela não é uma nova legenda e certamente faremos releitura. Para evitar que maior esforço seja gasto na releitura de legendas, é importante segmentá-las no momento do corte da câmera, sempre que a mudança da fala coincidir com a mudança da cena. Do contrário, se houver mudança de cena, mas não de fala, é preferível priorizar a segmentação retórica em detrimento da visual, e a legenda deve permanecer na cena seguinte.

A segmentação retórica diz respeito à distribuição do texto a partir do fluxo das falas, ou seja, ela deve ocorrer por blocos, e cada bloco de fala deve corresponder a uma nova legenda. Esse tipo de segmentação também está relacionado à marcação, especificamente ao sincronismo entre falas e legendas, que toma como base as pausas dos falantes. Para Díaz Cintas e Remael (2007), o modo como as legendas estão segmentadas deve refletir a dinâmica dos diálogos e ajudar a reforçar surpresa, suspense, ironia, hesitação, entre outras características da língua falada. Quando um desses marcadores da língua falada é antecipado ou atrasado na legenda, pode haver uma quebra da expectativa de quem está assistindo. Se o

espectador for surdo ou ensurdecido, o atraso ou a antecipação de uma surpresa ou suspense pode destoar das imagens e prejudicar a experiência audiovisual, uma vez que ele não terá as informações no exato momento em que elas foram dadas. É importante, na segmentação retórica, que cada legenda só apresente nova informação no momento em que esta for dada pelo falante ou pela trilha sonora, principalmente se for alguma informação relevante para o filme, como é o caso de segredos e revelações e das músicas de suspense.

Já a segmentação linguística, principal parâmetro da legendagem, consiste na distribuição das falas em legendas e deve ser feita por blocos, com base nas unidades semânticas e sintáticas, reforçando coesão e coerência. A segmentação ocorre na quebra da linha de uma mesma legenda e entre legendas diferentes. Para Karamitroglou (1998), o processo deve se dar no mais alto nível sintático possível e, se um bloco de informação não couber em uma única linha de legenda, o mesmo deve ser segmentado em duas linhas de modo a manter o maior agrupamento de carga semântica possível em cada linha. A figura 06 ilustra legendas com e sem problemas de segmentação.

Figura 06 – Exemplo de legenda sem problema (a) e com problema (b) de segmentação linguística



Fonte: Elaborada pela autora. Fotogramas do filme *Êxodo* (2015).

A segmentação no fotograma *a* preserva a integridade dos sintagmas, mantendo a carga semântica. A fala expressa na legenda tem duas orações, com a segunda intercalada, separadas por vírgulas: 1- *Seu primeiro decreto*, 2- *quando for a hora, será aposentá-la*. A estratégia de manter a oração menor na linha de cima e a maior na linha de baixo promove uma harmonia sintática, semântica e estética; já no fotograma *b*, a segunda oração está quebrada, porque separou *quando for a hora*, que representa uma oração subordinada de tempo.

Para finalizar esta subseção sobre os parâmetros da legendagem, retomo a importância de harmonizar os parâmetros com vista a proporcionar melhor recepção de legendas pelo público espectador.

2.3 ESTUDOS DA LEGENDAGEM BASEADOS EM *CORPORA* E PESQUISAS SOBRE RECEPÇÃO E PROCESSAMENTO DE LEGENDAS

O sucesso da interface entre Estudos da Tradução e Linguística de *Corpus* (LC), como previsto pela precursora dos Estudos da Tradução Baseados em *Corpora* (ETBC), Mona Baker (1993, 1995, 1996), está expresso nos inúmeros e profícuos trabalhos desenvolvidos sob as duas perspectivas, a exemplo das pesquisas do LEAD-LATAV, em Tradução Audiovisual Acessível (TAVa), que utilizaram a Linguística de *Corpus* como abordagem metodológica principal para a coleta e análise de dados. As pesquisas de Chaves (2009, 2012, 2015), o projeto CORSEL (ARAÚJO, 2015) e todos os seus desdobramentos (ASSIS, 2013; GABRIEL, 2013; NASCIMENTO, 2013; AGUIAR, 2013; CHAVES; ARAÚJO, 2014; ARAÚJO; ASSIS, 2014; VIEIRA *et al.*, 2015; ARRAES, 2015; FRANCO, 2016; LEONARDO, 2016; ASSIS, 2016) promoveram interface entre legendagem para surdos e ensurdecidos e LC.

O nascimento dos ETBC foi de extrema importância para a pesquisa em tradução porque ofereceu um novo modelo metodológico, resultando em uma nova forma de pensar a tradução. Isso ajudou na transformação da área, que ainda se encontrava presa à visão de tradução como atividade de segunda classe e, por isso, não oferecia nada além da investigação da equivalência e da questão de escolha entre duas línguas diferentes. Nesse período, o escopo da tradução ainda era restrito à comparação de duas línguas diferentes (tradução interlinguística), mesmo tendo reconhecidas as naturezas intralinguística e intersemiótica da tradução classificadas por Jakobson (1959). Os resultados de tais análises geravam resultados contestáveis, não conclusivos, além de deixar de lado o processo tradutório e a formação de tradutores.

Baker (1996) inaugurou seus pressupostos teóricos realizando uma análise baseada em *corpus* dos aspectos recorrentes em textos traduzidos. Os resultados mostraram quatro aspectos regulares em todo o *corpus* e, por essa razão, candidatos a *universais da tradução*, que, segundo a autora, incluem: simplificação (a ideia inconsciente que o tradutor tem de simplificar tanto a linguagem quanto a mensagem), explicitação (tendência de explicar ou adicionar informações detalhadas ao texto traduzido), normalização (tendência de adaptar

ou domesticar o texto fonte a padrões e práticas típicas da língua alvo, mesmo a ponto de parecer exagerado) e estabilização (consiste na tendência que os textos traduzidos têm de manter sempre o mesmo padrão textual, padrão este não encontrado em textos não traduzidos).

Na presente pesquisa, a legendagem para ouvintes é entendida como tradução, estando devidamente amparada pelos estudos da tradução. É texto autêntico, produzido por pessoas em contextos de situações reais, repleto de peculiaridades que precisam ser exploradas e descritas, o que justifica a análise baseada em *corpus*. Assim como Baker (1996) empenhou esforços na busca pelos universais da tradução, este trabalho busca encontrar padrões e características com base na análise de semelhanças e diferenças entre tipos de legendas.

As pesquisas baseadas em *corpus* do LATAV são pioneiras no desenvolvimento de metodologia baseada em *corpora* eletrônicos para análise de legendagem com a utilização de etiquetas para categorização dos dados e o uso do programa de análise linguística *WordSmith Tools*. A primeira pesquisa (CHAVES, 2009) comprovou a viabilidade do uso de *corpora* na análise de LSE a partir da análise baseada em *corpus* de seis trechos de programas da Rede Globo de TV com auxílio da versão 5.0 do programa *WordSmith Tools*. Os parâmetros analisados na legendagem desses programas foram: densidade lexical, nível de condensação, segmentação e explicitação. Os resultados sugeriram que a análise baseada em *corpus* é viável porque agiliza a análise de legendas em grande quantidade e revela características do *corpus*, contribuindo na definição de parâmetros de análise, que não seriam viáveis por meio de análise manual.

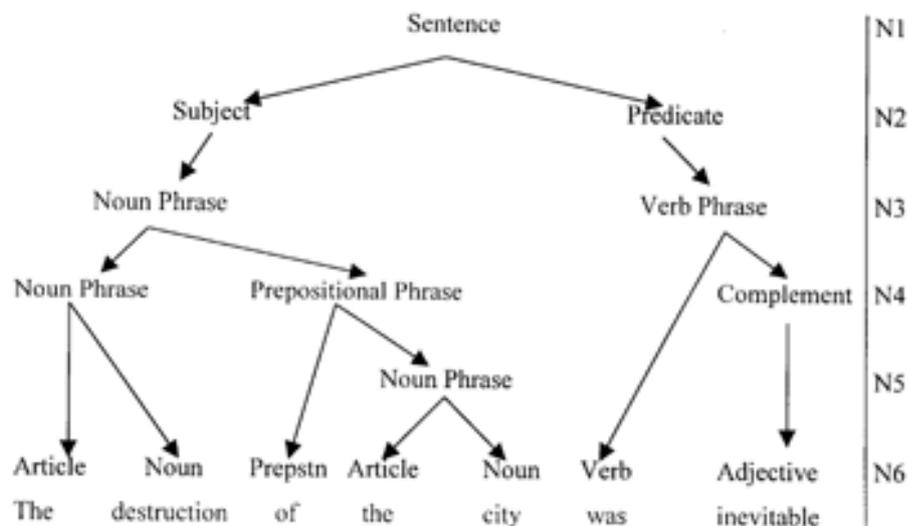
Certa da viabilidade do uso de *corpora* na análise de legendas, Chaves (2012) analisou problemas de segmentação na LSE inaugurando nova metodologia, mais robusta, que proporcionava a análise de legendas em grande escala. O projeto CORSEL compilou um *corpus* de legendas de séries, novelas e documentários da televisão brasileira e definiu um conjunto de etiquetas de parâmetros técnicos e de problemas de segmentação linguística para analisar legendas desses programas. Tal metodologia foi replicada nas demais pesquisas do LATAV e pela sua aplicabilidade contribuíram efetivamente para o delineamento teórico-metodológico da TAV.

A nova abordagem emergiu da necessidade de analisar o parâmetro da segmentação em legendas para surdos, visto que, na pesquisa exploratória do LATAV, o projeto MOLES (ARAÚJO, 2012), este parâmetro demonstrou ter influência sobre a recepção de legendas por surdos brasileiros (ARAÚJO; NASCIMENTO, 2012). Inicialmente, os

estudos sobre a segmentação das legendas foram realizados no contexto europeu e inspiraram a nova abordagem para análise de legendas, pois também demonstraram que os problemas de segmentação linguística em legendas poderiam prejudicar o entendimento de produtos audiovisuais (KARAMITROGLOU, 1998; PEREGO, 2008; KALANTZI, 2008, PEREGO *et al.*, 2010). Esses trabalhos têm em comum a discussão da segmentação na legendagem, mas divergem nas metodologias adotadas.

Karamitroglou (1998) se empenhou em descrever parâmetros de legendagem baseado na observação do contexto de legendagem do Reino Unido. Dentre os parâmetros descritos, o autor fez suas considerações sobre a segmentação com base na gramática gerativista, entendendo que a segmentação deve ser pautada nos mais altos níveis sintáticos possíveis. Ele ressalta que, idealmente, uma oração deve ocupar toda uma linha de legenda, mas nem sempre a oração completa cabe em uma linha só, e, quando isso acontece, o texto da oração deve ser distribuído em duas linhas. O autor demonstra como se dá a segmentação de uma oração utilizando nódulos sintáticos de base gerativista (vide figura 07).

Figura 07 – Árvore sintática utilizada para segmentação linguística por Karamitroglou (1998)



Fonte: Karamitroglou (1998, p. 9).

Na figura 07, Karamitroglou demonstra como pode ser feita a segmentação linguística da oração *The destruction of the city was inevitable* (em português: A destruição da cidade foi inevitável). A oração tem 44 caracteres e, por isso, excede a quantidade de caracteres a ser exibida em apenas 1 linha de legenda; assim, a oração deve ser distribuída em

duas linhas de legenda. Desta forma, a quebra da linha pode ocorrer a partir de diferentes nódulos sintáticos, sempre prezando pelo mais alto nível sintático.

A árvore sintática demonstra que a oração *The destruction of the city was inevitable* está dividida em 6 nódulos sintáticos, do mais alto N1 (oração) ao mais baixo N6 (léxico), como detalhado a seguir: O nódulo 1 (N1) é constituído da oração em si (*The destruction of the city was inevitable*); o nódulo 2 (N2) é constituído de sujeito (*The destruction of the city*) e predicado (*was inevitable*); o nódulo 3 (N3) é constituído de sintagma nominal (*The destruction*), sintagma preposicional (*of the city*) e sintagma verbal (*was inevitable*); o nódulo 4 (N4) é constituído de sintagma nominal (*The destruction*), sintagma preposicional (*of the city*), sintagma nominal (*the city*), sintagma verbal (*was inevitable*); o nódulo 5 (N5) é constituído de sintagma nominal (*The destruction*), sintagma preposicional (*of the city*), sintagma nominal (*the city*), sintagma verbal (*was inevitable*) e complemento (*inevitable*); e o nódulo 6 (N6) é constituído do léxico e sua função dentro da estrutura gramatical: artigo (*The*), substantivo (*destruction*), preposição (*of*), artigo (*the*), substantivo (*city*), verbo (*was*) e adjetivo (*inevitable*).

O mais alto nível sintático nessa representação de Karamitroglou mantém toda a oração de 44 caracteres junta, em um único nódulo sintático (N1). Entretanto, seguindo os parâmetros técnicos da legendagem, 44 caracteres devem ser distribuídos em duas linhas de legenda:

The destruction of the city
was inevitable

Nesse caso, o mais alto nível sintático é o N2, em que cada linha apresenta um nódulo, sendo a linha de cima o sujeito *The destruction of the city*, e a linha de baixo, o predicado *was inevitable*. Assim, “quanto mais alto é o nível sintático, maior é o agrupamento de carga semântica e mais completa é a informação a ser apresentada ao cérebro” (KARAMITROGLOU, 1998, p. 9, tradução nossa)³, facilitando, assim, a compreensão da legenda como um todo.

A partir dessa recomendação sobre segmentação de legendas, Perego (2008) começou a investigar a natureza dos problemas de segmentação na legendagem procurando conhecer os tipos existentes, por entender que problemas de segmentação podem ser

³ No original: “This occurs because the higher the node, the greater the grouping of the semantic load and the more complete the piece of information presented to the brain”.

prejudiciais no processamento de legendas, e corrobora a afirmação de Ivarsson e Carroll (1998) de que qualquer inconsistência experienciada pelos espectadores atrasa o processo de compreensão e faz dispersar do filme.

Levando em consideração que uma segmentação criteriosa, respeitando o mais alto nível sintático, pode aumentar a capacidade de processamento da legendagem, Perego (2008) analisou como ocorre a segmentação linguística entre linhas de legendas em inglês, italiano e húngaro de 15 filmes de gêneros variados e, como resultado dessa investigação, a autora verificou problemas de segmentação nos sintagmas nominal, preposicional e verbal e nas orações coordenada e subordinada. Perego (2008) demonstrou problemas de segmentação com base apenas na análise sintática, sem fazer relação ao contexto técnico da legendagem nem à tradução audiovisual. Além disso, não fez análise quantitativa dos dados levando em consideração a frequência desses problemas no *corpus*, nem contrastou a segmentação nos idiomas analisados (italiano, inglês e húngaro) nem a relação deles com os demais parâmetros da legendagem. Em decorrência dos achados sobre as categorias de problemas de segmentação, Perego *et al.* (2010) observam se o processamento cognitivo de legendas (texto e imagem) é eficaz, verificando se há muita dispersão no percurso entre legenda e imagem, e observam também se legendas de duas linhas com problemas de segmentação têm efeito perturbador na leitura e no reconhecimento de palavras; assim, testam a segmentação manipulando problemas no sintagma nominal de legendas em italiano de um filme húngaro. Perego *et al.* (2010) encontraram que não há diferença significativa no processamento de legendas com e sem problemas de segmentação, concluindo que as preocupações psicolinguísticas sobre a segmentação podem ser exageradas. Os resultados de Perego *et al.* (2010) divergem dos resultados de Monteiro (2016) e Vieira (2016), que comprovam que problemas de segmentação dificultam o processamento de legendas por surdos e ouvintes. Segundo Vieira (2016), os resultados encontrados em Perego *et al.* (2010) podem não ter indicado influência dos problemas de segmentação sobre a recepção de legendas por falha no desenho metodológico, e considero também que os autores discutiram os resultados sob a influência dos modelos de leitura preconizados pela psicolinguística para a leitura de texto impresso.

Em outra pesquisa sobre segmentação, Kalantzi (2008) utilizou metodologia baseada em *corpus* ao compilar um *corpus* com 360 legendas de dois documentários e um programa de atualidades da BBC e teve como foco analisar e descrever dois parâmetros linguísticos da LSE (segmentação e redução), além de descrever os parâmetros técnicos (estratégia de tradução, tempos inicial e final, tempo de exibição, caracteres por legenda,

velocidade da legenda, restrição de tempo, intervalo entre legendas, duração dos intervalos entre legendas, número de palavras por legenda, natureza das exclusões, número de linhas por legenda e erros de segmentação). Para a análise do parâmetro linguístico de segmentação, a pesquisadora procedeu com anotação automática do *corpus* se valendo do etiquetador *Claws5* e do conjunto de etiquetas morfossintáticas desenvolvidas pelo Centro de Pesquisa da Universidade de Lancaster UCREL⁴. Para análise do parâmetro de redução, fez comparação por alinhamento entre as legendas e as respectivas transcrições das falas dos programas utilizando o comparador de documentos do *Word*. Para os demais parâmetros técnicos, procedeu com anotação manual do *corpus*, criando as próprias etiquetas. Além de objetivar descrever parâmetros da LSE, Kalantzi (2008) teve como meta prescrever regras linguísticas para a segmentação gramatical de legendas.

Tais pesquisas impulsionaram a pesquisa de Chaves (2012), já motivada pelos resultados do projeto MOLES (ARAÚJO, 2012) que sugeriram que a boa segmentação facilitou a recepção de legendas por surdos brasileiros, em todas as velocidades testadas (145, 160 e 180ppm), contribuindo inclusive para a percepção de detalhes sobre os filmes. Chaves (2012) investigou o parâmetro de segmentação por um viés descritivo e baseado em *corpus* e teve como objetivo propor parâmetros para analisar os problemas de quebra de linha na LSE do filme *Nosso Lar* (2010). Foram utilizadas as recomendações de Karamitroglou (1998) sobre a segmentação de legendas no mais alto nível sintático. Para identificar os problemas de quebra de linha no *corpus* e prosseguir com a categorização dos problemas de segmentação, foram utilizadas definições de oração e dos sintagmas na sintaxe do português brasileiro propostas a partir da abordagem descritiva de orientação formalista de Perini (2010).

A análise dos problemas de quebra de linha das legendas para surdos do filme *Nosso Lar* possibilitou a criação de um conjunto de etiquetas para sistematizar a segmentação linguística na LSE. Chaves (2012) formulou 3 etiquetas indicativas de problemas de segmentação (linguística, retórica e visual) e 8 para análise de parâmetros técnicos da LSE (número de linhas, número de caracteres por linha, velocidade, tempo inicial e final da legenda). A partir dos conceitos, nomenclaturas e observações de Perini (2010) sobre a gramática do português brasileiro, a autora elencou um número de 19 subetiquetas referentes a categorias de problemas de segmentação linguística. Os resultados apontaram para uma maior quantidade de problemas de segmentação linguística nos níveis dos sintagmas verbal e

⁴ Sigla da *University Centre for Computer Corpus Research on Language*.

nominal (41% e 26%, respectivamente), em sua maioria na quebra entre dois ou mais verbos e em legendas com velocidades de 145ppm.

Outra pesquisa (DINIZ, 2012), inserida no âmbito do Programa de Cooperação Acadêmica (PROCAD) entre UECE e UFMG, também realizou estudo baseado em *corpus* para investigar a segmentação na LSE e na legendagem para ouvintes do filme *Irmãos de Fé* (2004). Para identificar os problemas de quebra de linha no *corpus*, Diniz utilizou definições dos sintagmas do português brasileiro, em forma de regras sintagmáticas, propostas a partir da abordagem descritiva de orientação gerativista de Lemle (1989) e verificou que os problemas de segmentação, chamados pela autora de desvios linguísticos, incidiram, na grande maioria, também no sintagma verbal, seguido pelo nominal.

As pesquisas de Assis (2013) e Araújo e Assis (2014), no escopo do projeto CORSEL (ARAÚJO, 2015) e impulsionadas pela pesquisa de Chaves (2012), passaram a embasar a análise da segmentação linguística por um viés funcionalista. Para tal, a “Nova Gramática do Português Brasileiro” (CASTILHO, 2012) – pela grande riqueza de detalhes desta no que se refere à descrição dos sintagmas e das sentenças complexas da língua portuguesa brasileira, sendo menos sintética que a descrição de Perini (2010) – foi adotada (ARAÚJO; ASSIS, 2014, p. 164). A análise das legendas da telenovela *Amor Eterno Amor* (2012) resultou em uma reformulação do conjunto de etiquetas proposto por Chaves (2012) e revelou maior ocorrência de problemas nos sintagmas verbal e nominal, também corroborando os resultados de Chaves (2012). Como uma proposta mais definitiva, esse conjunto de etiquetas foi adotado pelo projeto CORSEL e veio a ser utilizado pelas demais pesquisas aqui discutidas. Dentro dessa perspectiva, o Projeto CORSEL analisou a segmentação linguística de legendas de gêneros diversos da TV brasileira (séries, novelas e documentários) no intuito de identificar padrões nas legendas de TV e caracterizar os problemas de segmentação.

Os trabalhos de Gabriel (2013, 2015), impulsionados pelas pesquisas baseadas em *corpus* de Chaves (2012), Diniz (2012) e Assis (2013), investigaram a quebra de linha nas legendas para surdos e ensurdecidos da telenovela *Cheias de Charme*. A autora utilizou o conjunto de etiquetas de Assis (2013) e também as descrições linguísticas de Castilho (2012). Os resultados alcançados pela autora mostraram, assim como em Diniz (2012), Chaves (2012), Assis (2013), Araújo e Assis (2014) e Arraes (2015), um maior número de ocorrências de problemas de segmentação linguística no nível do sintagma verbal (35%). Em segundo lugar, bem próximo ao sintagma verbal, ficou o sintagma nominal (33%), o que reforça a relevância desses dois sintagmas no que diz respeito aos problemas de segmentação

linguística ocorridos nos gêneros de ficção telenovela e filme. Assim como em Assis (2013) e Araújo e Assis (2014), a maior parte desses problemas de segmentação linguística ocorreu em legendas de 3 linhas (53%) e com velocidade de 180ppm, considerada alta (54,5%).

Outra pesquisa realizada no âmbito do projeto CORSEL, Arraes (2015), analisou a segmentação linguística na LSE do filme *Virada Radical* (2006), transmitido pela Rede Globo em junho de 2014, utilizando metodologia quanti-qualitativa, baseada em *corpus*, também nos moldes do projeto CORSEL. Na análise das legendas do filme *Virada Radical*, os resultados mostraram um número baixo de problemas de segmentação; cerca de apenas 5,41% das legendas apresentaram algum tipo de problema, predominantemente nos sintagmas verbal (70,6%) e nominal (14,2%) e nas orações subordinadas (10,8%). A maioria desses problemas ocorreu em legendas de 180ppm (56,5%).

A pesquisa de Franco (2016), no escopo do CORSEL, utilizou a mesma metodologia das demais pesquisas, modificando apenas o gênero, que são as séries da TV aberta brasileira. As séries analisadas foram *A Grande Família*, *Pé na Cova* e *Tapas e Beijos*, exibidas pela Rede Globo. Os resultados foram semelhantes aos obtidos nas pesquisas com telenovelas e filmes, uma vez que o sintagma verbal foi também o que apresentou a maior quantidade de problemas (46,50%), seguido pelo sintagma nominal (27,70%), e a maior ocorrência dos problemas se deu em legendas de duas linhas (54,44%) e de velocidade de 180ppm (62,72%).

Já a pesquisa de Leonardo (2016) replicou a metodologia do CORSEL, analisando o documentário exibido pela Rede Globo *Profissão Repórter*. Essa pesquisa utilizou os mesmos recursos e metodologia das pesquisas anteriores e, analisando um *corpus* de 370 legendas, a pesquisadora chegou à conclusão de que quase metade das legendas do programa *Profissão Repórter* apresentou problemas de segmentação linguística (46,49%) e que estes ocorreram mais em legendas com velocidade alta (58,72%) do que em legendas de velocidade baixa (37,7%). Os problemas de segmentação linguística ocorreram em todas as categorias, havendo uma concentração maior nos sintagmas nominal (40,38%) e preposicional (33,65%), diferentemente dos problemas encontrados em filmes, séries e telenovelas, que ocorreram mais nos sintagmas verbal e nominal.

Na investigação de documentários, também no escopo do CORSEL, Vieira *et al.* (2015) e Vieira (2016) analisaram a segmentação linguística do documentário *Globo Repórter* exibido pela Rede Globo. Os resultados mostraram que, na amostra de 1097 legendas, há 462 problemas de segmentação linguística, que equivale a 42,1%. Assim como Leonardo (2016), as autoras também verificaram que em relação a esse gênero, e diferentemente dos gêneros

telenovelas, filmes e séries, sobressaíram-se o sintagma nominal, com 41,7%, e o sintagma preposicional, com 25,9%. Assim como nos outros gêneros da TV, os problemas ocorreram muito mais nas legendas com velocidade 180ppm (59,3%) do que nas legendas de velocidade 145ppm (39,1%).

Além das pesquisas que analisaram a segmentação das legendas de programas de TV, há também a pesquisa de Chaves e Amoreira (2015), no âmbito do projeto CORSEL, mas com proposta de análise de LSE de filmes brasileiros em DVD. A metodologia foi a mesma do projeto CORSEL, bem como o conjunto de etiquetas. Os resultados apontam semelhança com os demais gêneros de ficção do projeto CORSEL e concentram os problemas de segmentação no sintagma verbal (43,4%) e no sintagma nominal (22,7%), seguidos da oração subordinada (15,6%), dos sintagmas preposicional (11,3%) e adverbial (3,9%), da oração coordenada (1,7%) e do sintagma adjetival (1,1%). A maioria dos problemas ocorreu em legendas de velocidade baixa.

As pesquisas descritivas e baseadas em *corpus* que subsidiaram a investigação da segmentação na LSE de filmes e programas de TV brasileiros demonstraram, mesmo a partir de fundamentos linguísticos distintos, como o estruturalismo formalista de Perini (2010), o gerativo de Lemle (1989) e o estruturalismo funcionalista de Castilho (2012), que os problemas de quebra de linha se mostraram mais frequentes em sintagmas verbais e nominais, com exceção dos documentários, que tiveram maioria de problemas nos sintagmas nominal e preposicional. Outra consideração é a de que todos os resultados referentes à LSE de programas de TV tiveram problemas de segmentação em legendas mais rápidas, de 180ppm, enquanto a LSE dos filmes em DVD teve mais problemas de segmentação em legendas mais lentas, de 145ppm.

Apesar de os resultados do CORSEL terem mostrado que legendas de filmes em DVD e legendas de programas de TV apresentam percentuais significativos de problemas, sobretudo os documentários, e que esses problemas ocorrem em sua maioria nas legendas mais rápidas, de 180ppm, com exceção da LSE de filmes em DVD, que ocorre em velocidade mais baixa, de 145ppm, e apesar de os resultados do projeto MOLES sugerirem influência da segmentação sobre a recepção de legendas rápidas e lentas, não foram suficientes para concluir sobre a influência dos problemas de segmentação e da velocidade na recepção de legendas. Apesar de não ter sido objetivo dessas pesquisas, os resultados foram inconclusivos nesse sentido, abrindo espaço para uma investigação experimental-exploratória via rastreamento ocular. A partir desses achados, os problemas de segmentação na LSE passaram a ser testados experimentalmente no âmbito do projeto EXLEG.

O projeto EXLEG é um dos estudos com legendagem no âmbito do LEAD-LATAV e está em andamento. O projeto visa analisar experimentalmente a influência da segmentação e da velocidade na legendagem para surdos e ensurdecidos (LSE) na recepção de produções televisivas brasileiras por espectadores surdos. Baseado nas pesquisas exploratórias e descritivas anteriores, que demonstraram a deficiência da LSE brasileira em relação às suas características técnicas, principalmente velocidade e segmentação, o projeto EXLEG investiga a influência da segmentação das legendas na recepção de LSE pelos participantes.

Inicialmente, na pesquisa de d'Ydewalle *et al.* (1987) sobre a recepção e o processamento de diferentes velocidades, foram feitos 3 experimentos: o primeiro testou a recepção de três tempos de exposição de legendas (regra dos 4, 6 e 8 segundos), sendo a regra dos 6 segundos a mais utilizada por emissoras de TV. No segundo experimento, analisou-se o processamento ocular de legendas nas regras de 2, 4 e 6 segundos em três fitas diferentes do mesmo filme, cujos sujeitos dominassem ou não a língua falada. No terceiro teste, o grupo não recebeu a trilha sonora. Os autores encontraram que, na regra dos seis segundos, conhecer ou não a língua e escutar ou não a língua não influencia no tempo de leitura de legendas, devido à experiência que os sujeitos já têm com essa regra. O número de deflexões de linhas é explicado pela confusão com o número de linhas. Levando em consideração que o tempo gasto na imagem e nas legendas é quase sempre o mesmo, mais tempo é gasto com legendas de duas linhas. Em geral, o processamento de legendas parece ser uma atividade automática se estiver dentro do padrão da regra dos seis segundos, na qual são permitidos 32 caracteres por linha, exibidos em seis segundos.

Com o suporte teórico-metodológico da TAV e dos estudos experimentais em legendagem, foram realizados dois estudos exploratórios com procedimentos experimentais, envolvendo dois *corpora* diferentes: LSE de documentário e de programas eleitorais. A pesquisa contou com 16 participantes, sendo 8 surdos e 8 ouvintes, assistindo a quatro trechos diferentes com a LSE manipulada em quatro condições experimentais: lenta e bem segmentada (LBS), lenta e mal segmentada (LMS), rápida e bem segmentada (RBS) e rápida e mal segmentada (RMS). O primeiro estudo está concluído e mostrou estatisticamente que a legenda mais eficiente na recepção tanto de surdos quanto de ouvintes seria a RBS.

Ainda são poucos os estudos sobre a influência da segmentação linguística na recepção de espectadores surdos e ouvintes com rastreamento ocular. Até o momento, as principais pesquisas que se encaixam nesse contexto são Perego *et al.* (2010), Vieira (2016) e Monteiro (2016), estando as duas últimas no âmbito do projeto EXLEG-LEAD-LATAV.

Essas pesquisas apresentaram os primeiros resultados sobre a influência da segmentação linguística durante a recepção de legendas.

Perego *et al.* (2010) verificaram se legendas com problemas de segmentação no sintagma nominal comprometeriam o conforto da recepção de 15 minutos de um drama húngaro com legendas em italiano para espectadores ouvintes. A análise compreendeu as 28 legendas com problemas de segmentação. Em relação aos dados obtidos pelo rastreamento ocular, ocorreram três vezes mais fixações na área da legenda do que na área da imagem, porém a média de fixações foi maior na imagem do que na legenda (422×221) e 67% do tempo total foi gasto nas legendas. As legendas mal segmentadas tiveram duração média de 228ms, e as bem segmentadas, de 216ms. Os autores revelam que os dados estatísticos utilizando a ANOVA não revelou efeito significativo para a qualidade das legendas. Assim, os autores encerram a pesquisa afirmando, como já dito anteriormente, que as preocupações psicolinguísticas sobre a segmentação linguística da linha da legenda são provavelmente exageradas.

Vieira (2016) investigou o comportamento ocular de oito ouvintes e oito surdos assistindo a curtos vídeos de um documentário legendado. Todos os participantes eram graduandos ou graduados e tinham idade entre 18 e 41 anos. Todos os participantes surdos afirmaram que têm a Libras (Língua Brasileira de Sinais) como sua primeira língua. Foram quatro vídeos com legendas na velocidade de 145 e 180ppm, sendo que esses vídeos tiveram as legendas de duas linhas manipuladas de modo que as quebras entre as linhas respeitassem a segmentação linguística ou não. Não foi testado um sintagma específico, a autora testou os diversos problemas nos sintagmas e orações elaborados pelo grupo LEAD nos vídeos testados. Os grupos assistiram a vídeos com legendas bem segmentadas e mal segmentadas. Cada participante assistiu a quatro vídeos, sendo dois na velocidade lenta (bem e mal segmentados) e dois na velocidade rápida (bem e mal segmentados). O estudo envolveu uma análise do comportamento ocular nas 360 legendas de duas linhas com e sem quebras indevidas e em duas velocidades (145ppm e 180ppm) assistidas pelos participantes ouvintes e 360 legendas assistidas pelos participantes surdos. Mostraremos agora os achados sobre a movimentação ocular em legendas com boa e má segmentação linguística na pesquisa de Vieira (2016).

Os resultados encontrados em Vieira (2016) para a recepção de LSE de documentários mostraram que, para surdos, a legenda mais confortável, ou seja, com menos tempo de processamento, menos releitura e mais deflexão (movimento entre imagem e legenda) é a de tipo rápida e bem segmentada. Já as legendas lentas, bem ou mal segmentadas,

assim como as legendas rápidas mal segmentadas, propiciam releitura das legendas, afetando o processamento. Isso significa que, para os surdos, tanto os problemas de segmentação quanto a velocidade baixa (145ppm) atrasam o processamento de legendas. Para os ouvintes, as legendas mais confortáveis são as bem segmentadas, sejam rápidas ou lentas, resultando em menos processamento, menos releitura e mais movimento de deflexão. Já as legendas mal segmentadas, sejam rápidas ou lentas, atrapalham o processamento de legendas e provocam sensação de desconforto. Isso significa que, para os ouvintes, os problemas de segmentação atrapalham a recepção, mas a velocidade, não.

Outro dado de Vieira (2016) revela que legendas mal segmentadas estimulam mais movimentos regressivos, aumentando o tempo gasto no processamento de legendas, tanto para surdos quanto para ouvintes. Os movimentos regressivos ocorreram muito mais nos sintagmas do que na mesma palavra, na quebra de linha ou nas palavras diferentes. A autora complementa que sintagmas são estruturas complexas que encapsulam uma ideia e acredita que o movimento regressivo, ou seja, o rompimento dos sintagmas pode provocar interrupções na leitura.

Já o estudo de Monteiro (2016) buscou descobrir qual a influência da segmentação linguística entre legendas durante a recepção de propagandas políticas. Sua hipótese era a de que os problemas de segmentação entre legendas dificultariam a leitura por surdos e ouvintes, tornando a recepção das legendas desconfortável, mesmo que elas seguissem as configurações técnicas adequadas. A autora compartilhou da mesma metodologia de Vieira (2016). A diferença entre as duas pesquisas relaciona-se ao gênero e ao número de linhas, uma vez que, na pesquisa de Monteiro (2016), os vídeos apresentam legendas de apenas uma linha. Em relação às medidas analisadas, Monteiro observou somente o número de fixações e a duração média das fixações.

No referido estudo, mais uma vez, a hipótese de que as legendas, independentemente da velocidade, quando mal segmentadas, teriam a recepção prejudicada foi confirmada, tanto pelos dados gerados pelo rastreamento ocular como pela parte exploratória. A má segmentação nas legendas pode ter influenciado a recepção de surdos e ouvintes, causando um número de fixações maior nas legendas mal segmentadas e, em geral, durações de fixação maiores nas legendas do que nas imagens, fazendo com que os participantes não conseguissem relatar, de forma satisfatória, o conteúdo e os detalhes dos vídeos. Da mesma forma, os resultados de Monteiro (2016) também sugerem que as legendas bem segmentadas podem ter facilitado a recepção dos participantes. No entanto, a autora

relata que muitas vezes os dados numéricos não foram claros, por falta de estudos estatísticos mais detalhados, o que a levou a observar os dados qualitativos com maior minúcia.

Os resultados encontrados em Vieira (2016) para legendas de documentários e Monteiro (2016) para legendas de campanha política apontaram para a possibilidade de os problemas de segmentação afetarem a recepção de legendas rápidas e lentas, tanto para surdos quanto para ouvintes. Os resultados experimentais corroboram os exploratórios de Araújo (2012) e mostram que surdos desempenham melhor processamento de legendas quando estas são rápidas e bem segmentadas. Mostram também que os problemas de segmentação dificultam o processamento leitor de surdos e ouvintes em ambas as velocidades.

Com base na literatura sobre o processamento ocular, no que diz respeito ao processamento de legendas de língua nativa (flamenga), no contexto belga, com áudio em língua estrangeira, a média de duração da fixação de uma palavra por um adulto é de 178ms e por uma criança é de 248ms (D'YDEWALLE; DE BRUYCKER, 2003). Os autores expõem que legendas em língua nativa de uma ou duas linhas são melhor recebidas que legendas reversas (em língua estrangeira). Os dados do processamento de legendas por adultos e crianças mostraram que adultos leem mais rápido que crianças porque a duração média de fixação é menor em adultos, entretanto em outras situações, crianças e adultos têm leituras semelhantes. E legendas de uma e de duas linhas são recebidas diferentemente, sendo a de duas linhas melhor recebida que a de uma linha em todas as condições. Outro fato que os autores observaram é que as legendas de duas linhas nativas são melhor lidas que as de uma linha, reforçando que legendas de duas linhas são lidas “*word by word*” (palavra por palavra), enquanto nas legendas estrangeiras o olhar pula bastante. Em língua estrangeira, as legendas de uma linha são lidas com dificuldade, e no geral, as legendas estrangeiras não apresentam uma leitura sistemática, como a *word by word*, pois há muitos pulos nos quais o espectador escolhe algumas palavras para compor o texto.

Em resumo, d'Ydewalle e De Bruycker (2003) concluem que as legendas de uma linha foram recebidas com restrições na língua estrangeira porque não foram lidas de forma sistemática (*word by word*) e, nas legendas nativas, as de uma linha não são mal recebidas, mas as de duas linhas são melhores. Em todas as condições, legendas de uma linha apresentam leitura pouco sistemática, em que o sujeito pula muitas palavras e escolhe apenas algumas palavras para compor a legenda, ao contrário das legendas de duas linhas, que têm uma leitura mais palavra por palavra.

No processamento de legendas para surdos e ensurdecidos em português, de documentário brasileiro, com o volume baixo, portanto, sem o estímulo sonoro, a duração

média da fixação de palavras por surdos foi de 290ms/palavra, e por ouvintes foi de 270ms/palavra (VIEIRA, 2016). Isso significa que, no contexto europeu da legendagem de filme, com estímulos de som em língua estrangeira, imagem e legenda, uma palavra seria processada em 178ms por um adulto e em 248ms por uma criança, e no contexto brasileiro da LSE de documentários, com estímulo de imagem e legendas e sem estímulo sonoro, uma palavra seria processada em 290ms por surdos e em 270ms por ouvintes. Em nenhuma das pesquisas experimentais (VIEIRA, 2016; MONTEIRO, 2016; D'YDEWALLE, 1987, 2003, 2007; DE LINDY; KAY, 1999) a velocidade rápida acima de 180ppm foi controlada, abrindo uma lacuna para investigar se as legendas acima de 180ppm com todos os outros parâmetros controlados prejudica o processamento.

A partir dos resultados das pesquisas em LSE no Brasil e sua importância para a discussão teórica da legendagem para surdos e ensurdecidos no campo da Tradução Audiovisual e Acessibilidade, percebemos o avanço da pesquisa em LSE sobre a pesquisa em LO. As pesquisas exploratórias, descritivas e experimentais do LEAD-LATAV (ARAÚJO, 2008, 2009; CHAVES, 2009; ASSIS, 2013, 2015; GABRIEL, 2013, 2015; ARAÚJO, 2015; FRANCO, 2016; LEONARDO, 2016; VIEIRA, 2016; MONTEIRO, 2016) buscaram consolidar parâmetros contribuindo com o entendimento da LSE para os estudos da tradução audiovisual. Tal feito elucidou a LSE de filmes em DVD, telenovelas, séries, documentários e programas políticos da TV brasileira.

Sobre a legendagem enquanto área de estudos, inicialmente esta se baseava na prática profissional, oferecendo à área recomendações e regras sobre a legendagem para ouvintes com pouca reflexão teórica. Com o passar do tempo, as reflexões sobre a prática passaram a sistematizar parâmetros da legendagem, sendo possível propor interfaces promissoras com diferentes áreas de estudo, como a tradução e a Linguística de *Corpus*. Nesse sentido, a legendagem para surdos e ensurdecidos está à frente nas pesquisas e contribui para a ampliação da área da legendagem dentro do campo da tradução audiovisual.

Esta fundamentação teórica sobre a tradução audiovisual legendagem, os parâmetros da legendagem e o aporte teórico-metodológico das pesquisas do LEAD-LATAV possibilita a discussão sobre os parâmetros das legendas para ouvintes de filmes em DVD, em inglês e português, a partir do ponto de vista da pesquisa descritiva e baseada em *corpus*, com fins de dialogar com os outros tipos de legendas descritos em pesquisas prévias.

Antes de apresentar os resultados da pesquisa, será discutida a metodologia na próxima seção.

3 METODOLOGIA

Esta seção apresenta a metodologia utilizada para desenvolver esta pesquisa. São descritos: contexto e tipo de pesquisa, *corpus*, instrumentos de coleta e análise, tratamento e análise do *corpus* e procedimentos para análise dos dados.

3.1 CONTEXTO DA PESQUISA

Esta pesquisa consiste em uma tese de doutorado vinculada à linha de pesquisa Multilinguagem, Cognição e Interação do Programa de Pós-Graduação em Linguística Aplicada (PosLA) da UECE. A tese é desenvolvida no escopo da linha de pesquisa em Tradução Audiovisual Acessível: Legendagem e Audiodescrição (LEAD) do grupo de pesquisa Tradução e Semiótica, realizada no Laboratório de Tradução Audiovisual (LATAV), no Centro de Humanidades da UECE, sob a coordenação da Profa. Marisa Ferreira Aderaldo.

O estudo em questão teve origem nas pesquisas descritivas e baseadas em *corpus* sobre LSE desenvolvidas no âmbito do Projeto CORSEL (2015) coordenado pela profa. Vera Lúcia Santiago Araújo. Também é uma continuação das pesquisas em legendagem desenvolvidas com a contribuição da presente pesquisadora desde 2007, ainda na iniciação científica.

A pesquisa não envolve seres humanos, por isso não foi submetida ao comitê de ética da UECE, e não teve apoio financeiro de agências de fomento. A pesquisa recebeu doação de um aplicativo *web* para viabilizar a etiquetagem do *corpus*, o qual foi desenvolvido especialmente para fins da pesquisa⁵. Ao longo da pesquisa, o trabalho foi feito com a colaboração de três bolsistas de iniciação científica, sendo dois remunerados pelo programa de bolsas IC/UECE e um voluntário⁶.

3.2 TIPO DE PESQUISA

Com relação aos objetivos, trata-se de uma pesquisa descritivo-comparativa, por partir da observação da segmentação e da velocidade na legendagem para ouvintes, com fins

⁵ Programa desenhado pelo desenvolvedor independente mestre em Ciências da Computação Iago Castro Chaves e atualizado pela Doutora em Linguística Aplicada Alexandra Frazão Seoane.

⁶Bolsistas IC/UECE: Sofia Nicolau Amoreira (Bacharelado em Letras Inglês – CH/UECE) no período de agosto de 2014 a julho de 2015 e Raimundo Ednardo Soares de Souza Júnior (Licenciatura em Letras Inglês – FECLESC/UECE) no período de agosto de 2016 a julho de 2017; e bolsista voluntária: Renatta Pires Franco, no período de agosto de 2014 a julho de 2015. Aos três agradeço imensamente.

de confrontar tais características àquelas presentes nas legendas para ouvintes em inglês e português e nas legendas para surdos sem manipulação do *corpus*.

Quanto à coleta e ao processamento do *corpus*, a metodologia utilizada foi a abordagem baseada em *corpus*, advinda da Linguística de *Corpus*, por envolver procedimentos característicos da área, como etiquetagem do *corpus*, e uso de ferramentas de análise linguística, como o gerador de lista de palavras *WordList* e o concordanciador *Concord*, ambas ferramentas do programa *WordSmith Tools 7.0*. Além disso, propõe-se a analisar, por meio desses procedimentos metodológicos, *corpora* de textos autênticos, representativos da linguagem oral, em grande quantidade.

A análise dos dados foi de natureza quanti-qualitativa. A quantitativa gerou dados numéricos classificados quanto à frequência de ocorrência e às regularidades e padrões observados. A qualitativa foi feita a partir da descrição e comparação dos dados e proporcionou avaliar semelhanças e diferenças entre vários tipos de legendagem.

3.3 CORPUS

O *corpus* é composto das legendas de cinco filmes estrangeiros em língua inglesa distribuídos no Brasil, sendo do tipo paralelo, pois é constituído de textos em inglês e suas respectivas traduções em português. A escolha dos filmes levou em consideração datas de lançamento de até 15 anos, a partir de 2004, e foi feita de modo a abranger gêneros fílmicos variados de diferentes distribuidoras. São eles: *Flightplan – Plano de voo* (2004), suspense (Disney/Buena Vista); *Madagascar – Madagascar* (2005), animação (DreamWorks); *Click – Click* (2006), comédia (Columbia Pictures do Brasil); *Exodus: Gods and kings – Êxodo: Deuses e reis* (2015), ação (Twentieth Century Fox); *The social network – A rede social* (2010), drama (Sony Pictures), vide capas das versões em português na figura 08.

Figura 08 – Capas dos filmes em DVD



Fonte: Elaborada pela autora.

*Plano de voo*⁷ é um suspense com classificação indicativa de 14 anos. Kyle Pratt (Jodie Foster) é uma mulher devastada emocionalmente, devido à recente morte súbita de seu marido. Em meio a uma viagem de Berlim a Nova York, estando a mais de 40 mil pés de altitude e a bordo de um moderno avião, Kyle entra em pânico após perceber o desaparecimento de sua filha de 6 anos, Julia (Marlene Lawston). Desesperada, Kyle precisa provar à tripulação e aos passageiros sua sanidade, já que não há pista alguma sobre o paradeiro de Julia. Ao mesmo tempo, precisa convencer a si mesma que não está enlouquecendo.

*Madagascar*⁸ é uma animação com classificação livre. O leão Alex (Ben Stiller) é a grande atração do zoológico do Central Park, em Nova York. Ele e seus melhores amigos, a zebra Marty (Chris Rock), a girafa Melman (David Schwimmer) e a hipopótamo Gloria (Jada Pinkett Smith), sempre passaram a vida em cativeiro e desconhecem o que é morar na selva. Curioso em saber o que há por trás dos muros do zoo, Marty decide fugir e explorar o mundo. Ao perceberem, Alex, Gloria e Melman decidem partir a sua procura. O trio encontra Marty na estação Grand Central do metrô, mas, antes que consigam voltar para casa, são atingidos por dardos tranquilizantes e capturados. Embarcados em um navio rumo à África, eles acabam na ilha de Madagascar, onde precisam encontrar meios de sobrevivência em uma selva verdadeira.

*Click*⁹ é uma comédia de classificação livre. Michael Newman (Adam Sandler) é casado com Donna (Kate Beckinsale) e tem dois filhos. Com dificuldades em ver as crianças, já que tem feito serão no escritório de arquitetura em que trabalha, no intuito de chamar a atenção de seu chefe (David Hasselhoff), ele sente que nunca terá tempo de conciliar as duas coisas. Mas, após entrar em uma loja misteriosa com intuito de comprar um novo controle remoto para sua casa, ele parece ter encontrado uma solução. Isso porque, ao chegar no local, conhece o excêntrico funcionário Morty (Christopher Walken) e acaba comprando um controle remoto experimental, com a promessa de resolver vários problemas. Porém, Michael logo descobre que o controle tem outras funções, como abafar o som dos latidos de seu cachorro e também adiantar os fatos de sua própria vida, o que mudará as coisas para sempre, não necessariamente para melhor.

⁷ PLANO de voo. Título original: Flightplan. Direção: Robert Schwentke. EUA, 2005. Distribuição: Disney/Buena Vista. 1 DVD (98min), região 4, NTSC, color.

⁸ MADAGASCAR. Título original: Madagascar. Direção: Eric Darnell e Tom McGrath. EUA, 2005. Distribuição: Paramount Pictures. 1 DVD (86min), NTSC, color.

⁹ CLICK. Título original: Click. Direção: Frank Coraci. EUA, 2006. Distribuição: Columbia Pictures do Brasil. 1 DVD (107min), NTSC, color.

*A rede social*¹⁰ é um drama de classificação indicativa de 14 anos. Em uma noite de outono em 2003, Mark Zuckerberg (Jesse Eisenberg), analista de sistemas graduado em Harvard, senta-se em seu computador e começa a trabalhar em uma nova ideia. Apenas seis anos e 500 milhões de amigos mais tarde, Zuckerberg se torna o mais jovem bilionário da história com o sucesso da rede social Facebook. O sucesso, no entanto, leva a complicações em sua vida social e profissional.

*Êxodo: Deuses e reis*¹¹ é uma ação épica de classificação indicativa de 14 anos. Uma adaptação da história bíblica de Êxodo, segundo livro do Antigo Testamento. O profeta Moisés (Christian Bale) nasce entre os hebreus na época em que o faraó ordenava que todos os homens hebreus fossem afogados. Moisés é resgatado pela irmã do faraó e criado na família real. Quando se torna adulto, Moisés recebe ordens de Deus para ir ao Egito, na intenção de liberar os hebreus da opressão. No caminho, ele deve enfrentar a travessia do deserto e passar pelo Mar Vermelho.

3.4 INSTRUMENTOS

Foram instrumentos desta pesquisa, o extrator de legendas *SubRip 1.50*¹² para o procedimento de extração dos arquivos de legendas do DVD; o etiquetador de parâmetros técnicos da legendagem para a etiquetagem do corpus, a partir do conjunto de etiquetas do projeto CORSEL, que também é um dos instrumentos de categorização do corpus. Também são instrumentos duas das três ferramentas do *WordSmith Tools 7.0*¹³: o gerador de lista de palavras *WordList* para gerar listas de palavras com dados estatísticos sobre as legendas, e o concordanciador *Concord* para gerar linhas de concordância, importantes para a verificação dos padrões e das frequências de ocorrências dos parâmetros técnicos e linguísticos da legendagem.

¹⁰ A REDE social. Título original: The social network. Direção: David Fincher. EUA, 2010. Distribuição: Sony Pictures. 1 DVD (120min), NTSC, color.

¹¹ ÊXODO: Deuses e reis. Título original: Exodus: Gods and kings. Direção: Ridley Scott. EUA, Reino Unido, Espanha. 2014. Distribuição: Fox Film do Brasil. 1 DVD (151min), NTSC, color.)

¹² SUBRIP. Versão 1.50 Beta 4. 2015. Disponível em: <www.divx-digest.com/software/subrip.html>. Acesso em: 09 ago. 2017.

¹³ SCOTT, M. **WordSmith Tools**. Versão 7.0. 2017. Disponível em: <<http://www.lexically.net/wordsmith/version7/>>. Acesso em: 09 ago. 2017.

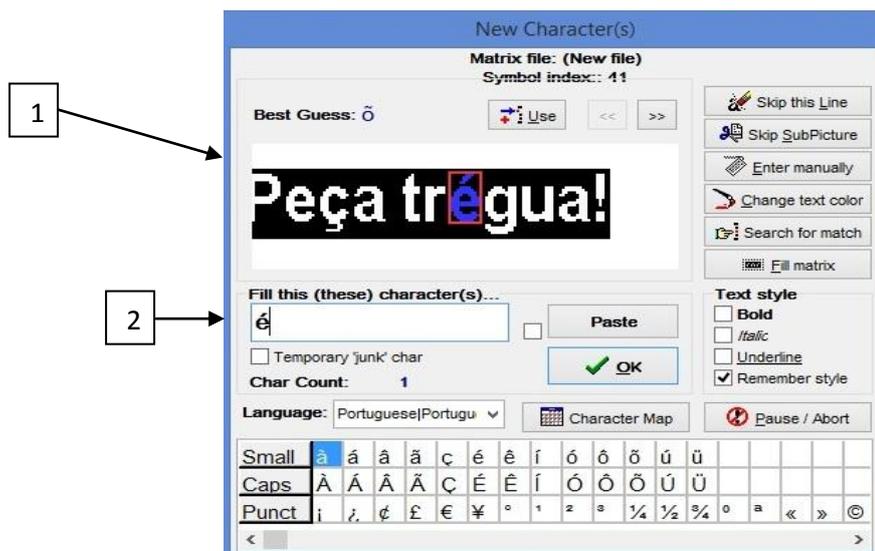
3.5 TRATAMENTO E ANÁLISE DO *CORPUS*

O tratamento do *corpus* é a preparação dos dados para análise. Essa etapa tem início no procedimento de extração das legendas de DVD, depois ocorre a etiquetagem e, por último, o processamento dos dados a partir do *Wordsmith Tools*.

3.5.1 Extração das legendas

A extração das legendas foi feita com auxílio do programa de computador *SubRip 1.50* (vide figura 09), que consiste num extrator de legendas via reconhecimento de caracteres – OCR (*Optical Character Recognition*). A legenda é extraída em formato srt., sendo facilmente visualizada quando aberta com o *Bloco de Notas* do *Windows*.

Figura 09 – Interface do programa *SubRip 1.50*

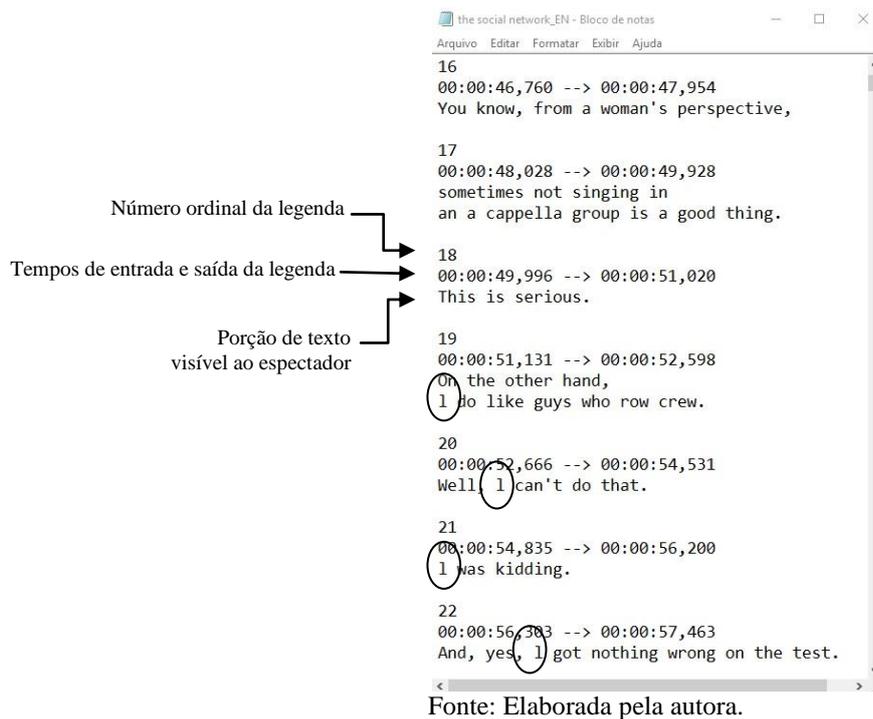


Fonte: Elaborada pela autora.

A figura 09 mostra a interface do programa *SubRip* após selecionado o arquivo de legenda (.vob) que será extraído. A extração dos arquivos .vob pode ser feita diretamente tanto de um diretório contendo a mídia DVD como de arquivos .vob sem a mídia. O campo indicado pela seta 1 é o espaço de visualização dos caracteres à espera de serem reconhecidos. A letra “é” delimitada em azul corresponde ao caractere que o programa pede reconhecimento. O campo indicado pela seta 2 é a caixa de texto na qual deve ser digitada a letra correspondente. Após digitar o caractere atentando para o formato, negrito, itálico ou sublinhado, é preciso validar o reconhecimento teclando *enter*.

Por ser um programa de reconhecimento de caracteres, a leitura é passível de erros, como, por exemplo, a distinção entre as letras “ri” e a letra “n” e entre as letras “I” maiúscula e “L” minúscula, conforme exemplos circutados na figura 10. É preciso, portanto, revisar as legendas extraídas a fim de evitar inconsistências nas etapas de análise.

Figura 10 – Arquivo de legenda em formato .srt gerado pelo *SubRip* e aberto no *Bloco de Notas*



O arquivo .srt é um arquivo de texto com formato simples e limpo que reúne as informações básicas de uma legenda como mostram as setas indicativas na figura 10: número ordinal informando a ordem numérica da referida legenda, tempos de entrada e saída da legenda e a legenda propriamente dita, que é a porção de texto que é visualizada pelo espectador. A partir dos arquivos de legendas (.srt) gerados pelo programa (vide figura 10), é possível fazer a etiquetagem do *corpus*, procedimento explicado na próxima seção.

3.5.2 Etiquetagem do *corpus*

A etiquetagem é um processo de anotação do *corpus* considerado por O'Donnell (1999) a análise do *corpus*. A análise ou categorização do *corpus* é um procedimento típico da abordagem metodológica baseada em *corpus*. Ainda sobre anotação, Leech (2005) esclarece

que é a prática de adicionar informações linguísticas interpretativas ao *corpus*¹⁴. De acordo com McEnery *et al.* (2006), essas informações adicionais podem ocorrer num nível metalinguístico, como, por exemplo, informações sobre autor, público leitor ou data de publicação para compor um cabeçalho, como também podem codificar uma análise de aspectos nos níveis discursivo, semântico, sintático, lexical, morfológico ou fonético¹⁵.

Nesta pesquisa, utilizei o conjunto de etiquetas do Projeto CORSEL, que apresenta etiquetas personalizadas para a análise de parâmetros técnicos da legendagem e de problemas de segmentação linguística. A etiquetagem dos parâmetros técnicos foi feita com auxílio de um etiquetador automático, e a dos problemas de segmentação foi feita manualmente a partir de anotação das legendas no bloco de notas. A anotação manual deve ser feita com auxílio de editores de texto, e o arquivo deve ser salvo em formato .txt (texto sem formatação) para poder ser decodificado pelos programas de análise linguística, como o escolhido para esta tese, *Wordsmith Tools 7.0*.

O conjunto de etiquetas formulado no Projeto CORSEL (ARAÚJO, 2015) compreende duas etiquetas de identificação da legenda no *corpus*, três para análise de parâmetros técnicos da legendagem, uma para identificação de problemas de segmentação e quatorze subetiquetas para análise dos problemas de segmentação linguística de acordo com o que é apresentado no quadro 03.

¹⁴ Minha tradução de: *the practice of adding interpretative linguistic information to a corpus*.

¹⁵ Minha tradução de: *Annotation can also occur at the meta-linguistic level – for example, by adding information such as author, level of readership or date of publication to a text’s header*.

Quadro 03 – Quadro sinótico das etiquetas do Projeto CORSEL

ETIQUETA DE IDENTIFICAÇÃO DA LEGENDA NO <i>CORPUS</i>	
Número da Legenda <sub1>legenda 1</sub1>	<sub1>legenda1</sub1>
Tempos inicial e final	<t>início --> final</t>
ETIQUETAS DE ANÁLISE DE PARÂMETROS TÉCNICOS DA LEGENDAGEM	
Linhas por legenda	<1L>, <2L> e <3L>
Número de caracteres por linha	<cpl>
Velocidade da legenda (145ppm) 14 e 15cps Velocidade de legenda (160ppm) 16cps Velocidade de legenda (180ppm) 17 e 18cps	<veloc_leg_>
ETIQUETA INDICATIVA DE PROBLEMA DE SEGMENTAÇÃO LINGUÍSTICA	
<PROSEGL>	
ETIQUETAS DE ANÁLISE DE SINTAGMA VERBAL (SV)	
<SV_composto> <SV_verbo+SAdv> <SV_(verbo)+oblíquo+SV> <SV_verbo+SP>	
ETIQUETAS DE ANÁLISE DE SINTAGMA NOMINAL (SN)	
<SN_especif+SN> <SN_subst+SAdj> <SN_subst+SP> <SN_núcleo>	
ETIQUETAS DE ANÁLISE DE SINTAGMA PREPOSICIONAL (SP)	
<SP_prep+SN> <SP_prep+SV>	
ETIQUETAS DE ANÁLISE DE SINTAGMA ADJETIVAL (SAdj)	
<SAdj_especif+SAdj>	
ETIQUETAS DE ANÁLISE DE SINTAGMA ADVERBIAL (SAdv)	
<SAdv_adv+adv>	
ETIQUETAS DE ANÁLISE DE ORAÇÃO SUBORDINADA (SUBORD)	
<SUBORD_conj/pron_rel+oração>	
ETIQUETAS DE ANÁLISE DE ORAÇÃO COORDENADA (COORD)	
<COORD_conj+oração>	

Fonte: Adaptado do relatório do Projeto CORSEL (ARAÚJO, 2015).

Conforme mostra o quadro 03, o conjunto de etiquetas do CORSEL oferece seis etiquetas e quatorze subetiquetas para análise da segmentação na legendagem de produtos audiovisuais brasileiros, sendo duas etiquetas de identificação da legenda no *corpus*, três etiquetas para análise dos parâmetros técnicos da legendagem, uma etiqueta indicativa de problema de segmentação linguística, quatro subetiquetas para análise do sintagma verbal, quatro subetiquetas para a análise dos problemas de segmentação no sintagma nominal, duas subetiquetas para análise do sintagma preposicional, uma subetiqueta para análise do sintagma adjetival, uma subetiqueta para análise do sintagma adverbial, uma subetiqueta para análise das orações subordinadas e uma subetiqueta para análise das orações coordenadas.

Na Linguística de *Corpus*, os parênteses angulares, com ou sem barra, caracterizam as etiquetas e servem para delimitar o conteúdo analisado. Os parênteses angulares vazios <> indicam onde se inicia uma categoria, e os parênteses angulares com uma barra </> indicam onde termina aquela categoria. O sinal de mais (+) representa a quebra indevida da estrutura linguística.

As etiquetas de identificação atribuem codificações que ajudam na busca no contexto e a posição em que se localiza a legenda. No entanto, elas não foram levadas em consideração durante a análise dos dados, pois servem apenas para orientar o pesquisador na localização das legendas enfocadas. As etiquetas de número da legenda <sub1>legenda1</sub1> e as de tempos inicial e final <t>início --> final</t> ajudam a identificar mais rapidamente a legenda em seu contexto e sua posição exata no vídeo.

As etiquetas de parâmetros técnicos codificam as legendas conforme os parâmetros da legendagem e serão levadas em consideração na análise dos dados. As etiquetas de linhas por legenda informam se a legenda é de uma, duas ou mais linhas <1L> ou <2L>, dado que influencia nos resultados, já que influencia na recepção (D'YDEWALLE; DE BRUYCKER, 2003; MONTEIRO, 2016; VIEIRA, 2016). A etiqueta de número de caracteres por linha <cpl> informa a quantidade de caracteres de cada linha de legenda, dado importante para harmonizar os parâmetros. Na literatura da área, baseada na prática de legendagem europeia, como foi visto no capítulo 2, Diaz Cintas e Remael (2007) sugerem que o número médio de caracteres é 37 caracteres por linha, para que a linha não se estenda pelas laterais da tela a ponto de ser cortada.

As etiquetas de velocidade da legenda correspondem às três velocidades de leitura investigadas por d'Ydewalle (1987) na leitura de legendas. A etiqueta de velocidade <veloc_leg_> vem acrescida do número de caracteres dividido pelo tempo de exibição da legenda, resultando em uma média de caracteres por segundo. Esse dado é muito importante,

porque durante muito tempo as velocidades eram classificadas como baixa (145ppm) média (160ppm) e alta (180ppm), mesmo não havendo evidências de que essa classificação desse conta dos padrões de velocidade e de leitura para determinados contextos. Nas pesquisas de Monteiro (2016) e Vieira (2016), as pesquisadoras perceberam que a velocidade em interação com a segmentação resulta em diferentes padrões de recepção de legendas que influenciam sobremaneira nos resultados da pesquisa. Desse modo, a etiqueta informa a média de caracteres por segundo, e não mais as classificações de baixa, média ou alta, e a velocidade da legenda: 145ppm (14 e 15cps), 160ppm (16cps) ou 180ppm (17 e 18cps).

A etiquetagem dos parâmetros técnicos da legendagem é automatizada, feita com um etiquetador desenvolvido por programadores independentes para fins de pesquisa em legendagem. O programa foi desenvolvido em *JavaScript* e consiste em um aplicativo web que usa um navegador de internet para ser executado (vide figura 11).

Figura 11 – Etiquetador automático de parâmetros técnicos da legendagem



Fonte: Elaborada pela autora.

De interface simples, o programa é dividido em duas caixas de texto: a de cima, própria para carregar o arquivo de legenda (.srt) não etiquetada e a de baixo apresenta as legendas já etiquetadas, com identificações das legendas e os parâmetros técnicos, após clicar o botão etiquetar indicado por uma seta na figura 11.

3.5.2.1 Etiquetas de problemas de segmentação

As etiquetas de problemas de segmentação codificam os problemas nas legendas contribuindo para encontrar padrões e regularidades no *corpus*. A etiqueta indicativa de problema de segmentação <PROSEGL> é aplicada logo após o problema de segmentação, na quebra de linha. E as quatorze etiquetas de problemas nos sintagmas e nas orações são aplicadas à medida que os problemas vão aparecendo. A primeira versão das etiquetas foi criada na minha pesquisa de mestrado (CHAVES, 2012). À época, o número de etiquetas era maior e o conceito de sintagmas e orações advinha da abordagem descritiva de orientação formalista de Perini (2010). As demais pesquisas no escopo do projeto CORSEL (ASSIS, 2013; GABRIEL, 2015, ARRAES, 2015, ARAÚJO, 2015), ao aplicar as etiquetas em outros tipos de legendas, começaram a refinar os padrões pelo viés funcionalista de Castilho (2012) com a justificativa de que essa vertente dos estudos linguísticos, segundo Praxedes Filho (2016) aborda a língua do ponto de vista de seu significado, de seu uso, de suas funções na interação humana, viabilizada por textos produzidos por pessoas em seus contextos de situação local ao atuarem no contexto mais amplo de cultura ou sociedade.

Em conformidade com essa abordagem linguística, a presente pesquisa faz interface com a Linguística de *Corpus*, ao utilizá-la como abordagem metodológica. Segundo Conrad e Levelle (2008, p. 539) a Linguística de *Corpus* é uma ferramenta usada na investigação do uso da língua. Emprega técnicas assistidas por computador a fim de analisar grandes conjuntos de dados de escrita ou fala transcrita, tendo em vista a descrição das escolhas linguísticas – sejam usuais ou não – que falantes e escritores fazem em circunstâncias particulares. Dessa forma, a interface entre o funcionalismo linguístico e a Linguística de *Corpus* é totalmente compatível com os objetivos, os procedimentos metodológicos e a análise deste trabalho.

As pesquisas descritivas e baseadas em *corpus*, bem como o projeto CORSEL e os pressupostos da gramática funcionalista de Castilho (2012), fundamentam a análise do *corpus* desta pesquisa e geram perspectivas quanto à identificação de padrões recorrentes nos problemas de segmentação linguística na legendagem para ouvintes de filmes em DVD, em inglês e português. Os conjuntos de etiquetas das pesquisas de Chaves (2012) e Assis (2013) subsidiaram as primeiras análises do CORSEL, porém essas etiquetas foram sendo revisadas e aperfeiçoadas de acordo com as discussões teóricas geradas a partir dos problemas de segmentação encontrados nos *corpora*.

A etiquetagem partilha a mesma nomenclatura para analisar legendas em inglês e português. Apesar de serem línguas diferentes, o conjunto de etiquetas pôde ser o mesmo porque tinha como foco a identificação de quebras nos sintagmas e nas orações, pautada, sobretudo, pelos parâmetros técnicos e linguísticos da legendagem e minimamente pela análise linguística. A identificação de quebras nos sintagmas e nas orações prescinde de análise linguística com alto nível de detalhamento.

A discussão dos problemas de segmentação do ponto de vista linguístico se dá basicamente em reconhecer as estruturas dos sintagmas e das orações, que são: núcleo, complementadores e especificadores, com o intuito de preservar a relação semântica estabelecida pela união dessas estruturas, estando ciente que, do contrário, a quebra dessa relação pode comprometer a recepção de legendas. Por essa análise, foi possível categorizar legendas em inglês e português se valendo das mesmas etiquetas.

3.5.2.2 Concepção de sintagmas e orações de Castilho (2012)

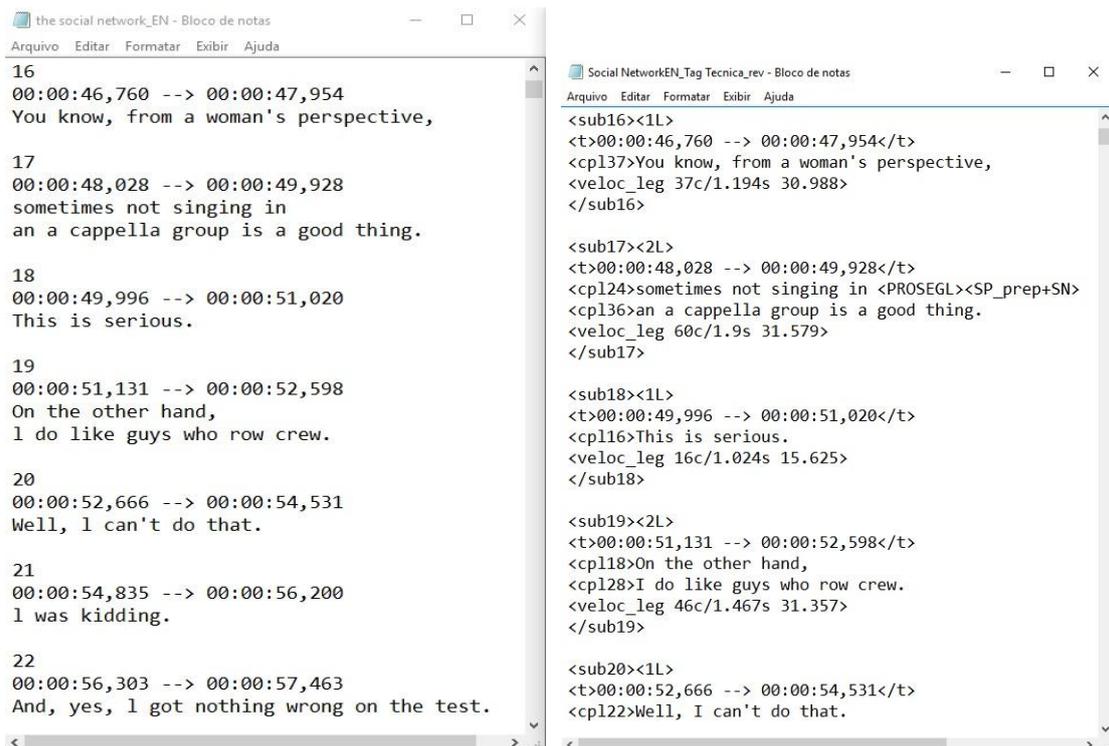
Na abordagem funcionalista de Castilho (2012, p. 249), “a estrutura sintática da sentença fundamenta-se nos arranjos lexicais de que ela é formada, os sintagmas, bem como nas funções que decorrem do relacionamento desses sintagmas”, sendo estes um somatório de constituintes que apresentam um lugar previsível dentro da estrutura da oração. Ainda segundo o autor, o sintagma tem uma estrutura composta por um “núcleo”, uma margem esquerda, preenchida ou não pelos chamados “especificadores” do sintagma, e uma margem direita, preenchida ou não pelos ditos “complementadores”.

Em geral, sobre os sintagmas e orações, Castilho (2012) explica que o SV é a construção que tem como núcleo o verbo. O SN é uma construção cuja estrutura tem por núcleo um substantivo ou um pronome, tendo por especificador o artigo e os pronomes e por complementadores os SAdjs e os SPs. O SAdj tem por núcleo o adjetivo, que é uma classe basicamente predicadora, funcionando como adjunto adnominal, enquanto constituinte do sintagma nominal, ou como predicativo, enquanto constituinte do sintagma verbal (CASTILHO, 2012, p. 516). O SAdj tem como especificador advérbios predicativos qualificadores e como complementador SPs e sentenças substantivas objetivas. O SP tem por núcleo a preposição, por especificador o advérbio, enquanto que o complementador em um SP também pode ser: (i) um verbo, (ii) um pronome, (iii) um quantificador definido, (iv) um quantificador indefinido, entre outras possibilidades de realização. O SAdv tem como núcleo o advérbio e como especificador e complementador outros advérbios. As orações são

enunciados que possuem mais de um verbo, ou seja, contêm mais de uma sentença, podendo estabelecer relação de coordenação, subordinação e correlação entre si.

A etiquetagem dos problemas de segmentação foi feita manualmente no *Bloco de Notas* a partir do arquivo srt, o qual já tinha as etiquetas de parâmetros técnicos. Terminada a etiquetagem, os arquivos ficaram conforme figura 12.

Figura 12 – Arquivos .srt não anotado e anotado



Fonte: Elaborada pela autora.

A figura 12 mostra dois arquivos de texto sem formatação (.txt). À esquerda, o arquivo gerado pelo *SubRip*, sem anotação, e à direita, o arquivo anotado com base no conjunto de etiquetas do CORSEL. Após o procedimento de anotação e revisão dessa anotação, o arquivo pode ser analisado com auxílio do *WordSmith Tools 7.0*.

3.5.3 Processamento do *corpus* via *WordSmith Tools*

A análise do *corpus* foi feita com auxílio do programa *WordSmith Tools 7.0*, especificamente o gerador de lista de palavras *WordList*, que fornece informações quantitativas e uma lista de todas as palavras do *corpus* por ordem de frequência e por ordem alfabética, e o concordanciador *Concord*, ferramenta que extrai as ocorrências de uma palavra de busca no *corpus* juntamente com seu contexto, apresentando-a destacada em posição

central, junta a seu cotexto à direita e à esquerda, em linhas de concordância. O programa forneceu dados quantificados em percentual a partir dos valores absolutos, os quais quantificam o número de vezes que determinada característica ocorreu. A partir da interação dos dados, foram feitas análises interpretativas sobre as relações exercidas pelos parâmetros e características da legendagem analisados.

A ferramenta *WordList* (vide figura 13) fornece dados como número de *tokens* (palavras corridas), número de *types* (palavras distintas), a relação entre *types* e *tokens* indicando a riqueza lexical do *corpus*, a média de palavras por sentença, a média de caracteres por palavra, entre outros dados.

Figura 13 – Tela da *WordList* com informações das legendas em inglês e português dos 5 filmes

	Overall	Social NetworkEN	Social NetworkBR	ClickEN	ClickBR	MadagascarEN	MadagascarBR	ExodusEN	ExodusBR	FlightplanEN	FlightplanBR
file size	3.961.934	691.598	599.910	447.890	443.036	341.672	336.616	326.432	324.678	225.146	224.956
tokens (running words) in text	167.734	31.031	24.294	19.828	18.240	14.915	13.739	13.395	12.771	9.998	9.523
tokens used for word list	82.963	17.109	10.856	10.677	8.922	7.787	6.467	5.865	5.171	5.311	4.798
sum of entries											
types (distinct words)	10.064	2.467	2.683	1.738	1.875	1.410	1.642	1.154	1.416	1.019	1.232
typetoken ratio (TTR)	12,13	14,42	24,71	16,28	21,02	18,11	25,39	19,68	27,38	19,19	25,68
standardised TTR	24,89	25,91	25,52	25,78	25,41	23,89	24,71	21,09	22,21	25,71	26,87
SSTR std.dev.	74,04	70,63	70,06	68,34	68,45	68,63	67,60	70,35	68,98	65,31	63,91
SSTR basis	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
mean word length (in characters)	3,75	3,79	3,78	3,69	3,78	3,64	3,82	3,64	3,74	3,72	3,89
word length std.dev.	2,08	2,08	2,28	1,95	2,09	1,88	2,15	1,94	2,14	1,97	2,23
sentences	16.796	2.639	2.241	2.136	2.145	1.823	1.619	1.149	1.121	959	964
mean (in words)	4,94	6,48	4,84	5,00	4,16	4,27	3,99	5,10	4,61	5,54	4,98
std.dev.	4,26	5,71	3,91	3,84	3,13	4,38	3,17	4,35	3,87	4,48	3,57
paragraphs	11	1	1	1	1	1	1	1	1	1	2
mean (in words)	7.542,09	17.109,00	10.856,00	10.677,00	8.922,00	7.787,00	6.467,00	5.865,00	5.171,00	5.311,00	2.399,00
std.dev.	4.270,91										797,62
headings											
mean (in words)											
std.dev.											
sections	10	1	1	1	1	1	1	1	1	1	1
mean (in words)	8.296,30	17.109,00	10.856,00	10.677,00	8.922,00	7.787,00	6.467,00	5.865,00	5.171,00	5.311,00	4.798,00
std.dev.	3.810,94										
numbers removed	84.771	13.922	13.438	9.151	9.318	7.128	7.272	7.530	7.600	4.687	4.725
stoplist tokens removed											
stoplist types removed											

Fonte: Elaborada pela autora.

A figura 13 mostra a tela da *WordList* com dados relevantes para a presente análise, a saber, na coluna à esquerda: o número de palavras corridas (*Tokens (running words) in text*) e o número de palavras excetuando os números (*Tokens used for word list*). A segunda coluna, da esquerda para a direita, *Overall*, mostra os dados numéricos referentes às duas legendagens (inglês e português) dos 5 filmes em conjunto, e as 10 colunas subsequentes mostram os dados numéricos individuais das legendas de cada filme. Cada arquivo está nomeado com o título do filme e a indicação de que está em inglês (EN) ou em português (BR). Outro dado importante fornecido pela *WordList* foi a densidade lexical, obtida pelo número de palavras distintas (*types*) dividido pelo número de palavras (*tokens used for word*

list) de uma legendagem. Pode ser visualizada na tela da *WordList*, na primeira coluna à esquerda, na opção *type/token ratio (TTR)*.

Já a ferramenta *Concord* proporcionou a análise qualitativa a partir da comparação de dados provenientes das legendas em inglês e das legendas em português, como: problemas de segmentação, velocidade e número de linhas. Essa análise avaliou, sobretudo, os parâmetros de segmentação e velocidade na legendagem para ouvintes de filmes em DVD nos contextos de línguas inglesa e portuguesa. Para extrair o número de problemas de segmentação, foram geradas linhas de concordância que mostraram as ocorrências dos problemas destacada em nóculo central e o número de ocorrências, como mostra a figura 14.

Figura 14 – Tela da *Concord* com os problemas de segmentação linguística em nóculo

The screenshot shows the Concord software interface. A callout box with the text "Problema de segmentação em destaque" points to a specific row in the concordance table. The table has columns for line number (N), time, text, and various linguistic tags. The highlighted row is line 19, which contains the text "Quero que saiba que" followed by several tags including <PROSEGL>, <SUBORD_conj/pron_rel+oração>, and <cpl22>.

N	Time	Text	Tags	Set	Tag	Word #	Seni	Senl	Para	Para	Heai	Sect	Secl
1	01:05:40,336	...e saiba que	<PROSEGL><SUBORD_conj/pron_rel+oração> <cpl22>estou com	14,556	1,414	0	14					0	14
2	01:05:27,924	...ançar quem	<PROSEGL><COORD_conj+oração> <cpl22>sabe por quanto	14,486	1,415	0	14					0	14
3	01:05:49,312	...ão podemos	<PROSEGL><SV_composto> <cpl15>brigar. Acabou. <veloc_leg	14,616	1,413	0	14					0	14
4	01:05:45,174	...ilhar e tenta	<PROSEGL><SV_composto> <cpl14>resolver isso. <veloc_leg 36c	14,592	1,414	0	14					0	14
5	01:02:25,007	...pizo meu Deus, o pato não demorou	<PROSEGL><SV_verbo+SAdv> <cpl30>nada para esquecer o	13,811	1,316	0	13					0	13
6	01:46,335	...ele tentou trepar na minha	<PROSEGL><SN_especif+SN> <cpl23>perna, e eu não deixei.	13,656	1,332	0	13					0	13
7	01:05:02,331	...Eu avisei que	<PROSEGL><SUBORD_conj/pron_rel+oração> <cpl23>não poderia	14,396	1,413	0	14					0	14
8	01:03:42,608	...Avancei o sexo uma vez, não	<PROSEGL><SV_verbo+SAdv> <cpl33>quero que seja assim para	14,082	1,416	0	14					0	14
9	02:591	...Quero ficar com você	<PROSEGL><SN_especif+SN> <cpl31>vida toda, por isso vou	14,696	1,415	0	14					0	14
10	08:33,075	...se um dia você acabasse	<PROSEGL><SV_composto> <cpl18>virando nosso CEO.	15,161	1,527	0	15					0	15
11	01:08:09,143	...Michael, Janine e eu decidimos	<PROSEGL><SV_composto> <cpl19>sair da correria... <veloc_leg	15,031	1,415	0	15					0	15
12	11:35,224	...Meu namorado, o garoto	<PROSEGL><SN_subst+SAdj> <cpl21>mais lindo da escola.	15,756	1,514	0	15					0	15
13	04:927	...Veja só quem fala! O capitão	<PROSEGL><SN_núcleo> <cpl25>Twinkie, do navio Gordão!	15,584	1,514	0	15					0	15
14	25:207	...puxando o saco e querendo	<PROSEGL><SV_composto> <cpl30>saber quem tem a melhor	14,872	1,431	0	14					0	14
15	21:537	...John, acho que não devemos	<PROSEGL><SV_composto> <cpl17>gastar energia... <veloc_leg	14,851	1,415	0	14					0	14
16	01:08:02,867	...tem uma coisa, ou melhor, uma	<PROSEGL><SN_especif+SN> <cpl27>peessoa de quem quero	15,000	1,428	0	14					0	14
17	01:07:36,719	...Esta é a maior	<PROSEGL><SN_subst+SAdj> <cpl22>maluquice que já ouvi.	14,921	1,414	0	14					0	14
18	00:54:22,993	...Você já pensou nas	<PROSEGL><SP_prep+SN> <cpl16>consequências... <veloc_leg	11,736	1,114	0	11					0	11
19	31,007	...Esqueci de dizer que sai com	<PROSEGL><SP_prep+SN> <cpl34>sua amiga Janine no fim de	11,602	1,116	0	11					0	11
20	00:54:54,224	...Ele vive atrás do pote	<PROSEGL><SN_subst+SAdj> <cpl29>de ouro, no fim do arco-iris.	11,882	1,115	0	11					0	11
21	00:54:27,998	...Sim. Acha que eu não	<PROSEGL><SV_verbo+SAdv> <cpl16>devo fazer isso?	11,764	1,116	0	11					0	11
22	00:50:00,564	...Eu sabia! Sabia que	<PROSEGL><SUBORD_conj/pron_rel+oração> <cpl24>	10,856	1,014	0	10					0	10
23	49:34,604	...Não é uma bicicleta, nem	<PROSEGL><COORD_conj+oração> <cpl36>um chapéu de caubói	10,721	1,015	0	10					0	10
24	00:53:03,809	...Eu disse: "Consiga a conta do	<PROSEGL><SP_prep+SN> <cpl28>Watsuhita e será promovido".	11,496	1,116	0	11					0	11
25	00:50:12,509	...Amor! A que estava	<PROSEGL><SV_composto> <cpl28>tocando quando nos	10,914	1,015	0	10					0	10
26	55:41,971	...Está bem. Querem que	<PROSEGL><SUBORD_conj/pron_rel+oração> <cpl17>eu veja?	11,996	1,114	0	11					0	11
27	00:50:42,111	...F agora ela profere	<PROSEGL><SV_composto> <cpl28>seu chamada de Alan	13,001	1,214	0	13					0	13

At the bottom left, a circled box contains the text "628 entries".

Fonte: Elaborada pela autora.

A parte central, em destaque, é o nóculo, ou palavra de busca, que nesse caso é a etiqueta de problema de segmentação <PROSEGL> das legendas dos cinco filmes em inglês e em português. Abaixo, no canto inferior esquerdo, circulado, é possível visualizar o número de ocorrências de problemas de segmentação, nesse caso, 628. O mesmo procedimento foi feito para extrair dados sobre as subcategorias de segmentação, sobre a velocidade, sobre o número de linhas e sobre o número de caracteres por linha.

3.6 PROCEDIMENTOS PARA ANÁLISE DOS DADOS

Após análise do *corpus*, foi empreendida a análise dos dados da pesquisa, em etapas correspondentes aos objetivos desta tese. Todas as etapas foram quanti-qualitativas, baseadas na comparação de dados entre legendas. Na primeira etapa, foi feita uma comparação da LO em línguas portuguesa e inglesa a partir da descrição dos problemas de segmentação e velocidade. Na segunda etapa, foi feita uma comparação da LO em português e da LSE em português a partir da descrição dos problemas de segmentação e velocidade. Na terceira etapa, foi feita uma descrição dos padrões de segmentação e velocidade da LO em línguas inglesa e portuguesa; e na quarta etapa, uma avaliação geral da LO em inglês e português. Esta análise dos dados, bem como a discussão dos resultados, está exposta na seção seguinte.

4 RESULTADOS E DISCUSSÃO

Esta seção apresenta os resultados da pesquisa em quatro momentos, referentes às três perguntas de pesquisa e à discussão dos resultados, respectivamente, precedidos por um panorama do *corpus*. No primeiro momento, é feita uma comparação da LO em línguas portuguesa e inglesa a partir da descrição dos problemas de segmentação e velocidade. No segundo, uma comparação da LO em português e da LSE em português a partir da descrição dos problemas de segmentação e velocidade. No terceiro, uma descrição dos padrões de segmentação, relacionados tanto à linguística quanto à distribuição do texto audiovisual em linhas de legendas, e dos padrões de velocidade da LO em línguas inglesa e portuguesa. No quarto momento, uma avaliação da LO em inglês e português, seguida da discussão dos resultados. O primeiro momento explora apenas os dados novos, da atual pesquisa: LO de filmes em DVD, em línguas inglesa e portuguesa. O segundo apresenta os dados atuais em comparação com os resultados das pesquisas descritivas (KARAMITROGLOU, 1998; PEREGO, 2008; ASSIS, 2013; GABRIEL, 2013; CHAVES; ARAÚJO, 2014; ARAÚJO; ASSIS, 2014; CHAVES; AMOREIRA, 2015; VIEIRA *et al.*, 2015; ARRAES, 2015; FRANCO, 2016; LEONARDO, 2016; ASSIS, 2016). O terceiro traz os dados da pesquisa atual em comparação com os resultados das pesquisas experimentais (D'YDEWALLE *et al.*, 1987; VIEIRA, 2016; MONTEIRO, 2016; PEREGO *et al.*, 2010). Por fim, o último momento enfoca a LO de filmes em DVD, em inglês e português, e discute os resultados, ressaltando contribuições da pesquisa para a teoria da legendagem.

4.1 PANORAMA DO *CORPUS*

Antes de discutir os dados e apresentar os resultados, é importante fazer um panorama do *corpus*, no qual são revelados dados quantitativos provenientes da *WordList* como: número de palavras (*tokens*), número de legendas, tamanho médio das palavras, número médio de palavras por sentença, densidade lexical e número de linhas. Tais características oferecem mais detalhes sobre as legendas e contribuem para uma melhor compreensão dos problemas de segmentação e da velocidade das legendas em inglês e em português. Os dados quantitativos estão apresentados no quadro 04.

Quadro 04 – Dados quantitativos provenientes da WordList

	LO dos 5 filmes em inglês		LO dos 5 filmes em português	
Número de palavras (<i>tokens</i>)	54.956		36.214	
Número de legendas	7.030		6.989	
Tamanho médio das palavras (em caracteres)	3,71		3,79	
Número médio de palavras por sentença	5,37		4,48	
Densidade lexical – <i>type/token ratio</i> (TTR)	14,42		24,71	
Número de linhas – 1L / 2L	3.871(55%)	3.149 (45%)	4.305 (62%)	2.684 (38%)

Fonte: Elaborado pela autora.

O quadro 04 reúne os dados quantitativos de ambas as legendas em colunas distintas, sendo possível comparar os contextos de língua inglesa e de língua portuguesa. Os dados provenientes da *WordList* mostram que o *corpus* de legendas dos cinco filmes em DVD, em língua inglesa e suas respectivas traduções em língua portuguesa, possui 82.963 *tokens* (palavras).

O programa oferece duas contagens de *tokens*; a primeira, “*tokens (running words) in text*” não será levada em consideração para esta análise, pois, além das palavras, inclui os números na contagem, e a segunda, “*tokens used for word list*” considera apenas as palavras, excluindo os números da contagem. Apenas a segunda opção (vide indicação 1 da figura 15) é relevante para esta análise, uma vez que a análise da segmentação se baseia apenas nos textos das legendas.

Figura 15 – Dados numéricos da WordList

text file	Overall	Social NetworkEN	Social NetworkBR	ClickEN	ClickBR	MadagascarEN	MadagascarBR	ExodusEN	ExodusBR	FlightplanEN	FlightplanBR
file size	3.961.934	691.598	599.910	447.890	443.036	341.672	336.616	326.432	324.678	225.146	224.956
tokens (running words) in text	167.734	31.031	24.294	19.828	18.240	14.915	13.739	13.395	12.771	9.998	9.523
1 tokens used for word list	82.963	17.109	10.856	10.677	8.922	7.787	6.467	5.865	5.171	5.311	4.798
sum of entries											
2 types (distinct words)	10.064	2.467	2.683	1.738	1.875	1.410	1.642	1.154	1.416	1.019	1.232
3 type/token ratio (TTR)	12,13	14,42	24,71	16,28	21,02	18,11	25,39	19,68	27,38	19,19	25,68

Fonte: Elaborada pela autora.

A figura 15 mostra o número geral (*Overall*) de *tokens* e o número parcial de *tokens* de cada um dos dez arquivos de legendas, cinco em inglês e cinco em português,

identificados pelo título do filme em inglês mais a indicação EN, para inglês, e BR, para português.

Do total de palavras, 54.956 são do inglês e 36.214 são do português, sendo o número de palavras sempre maior nas legendas em inglês. Sobre o número de legendas, a LO em inglês dos cinco filmes possui 7.030 legendas, e a LO em português dos cinco filmes possui 6.989 legendas, conforme comparativo a seguir.

	<i>Social network</i>	<i>Click</i>	<i>Madagascar</i>	<i>Exodus</i>	<i>Fightplan</i>	<i>Overall</i>
Inglês	2.291	1.521	1.186	1.254	778	7.030
Português	2.188	1.548	1.207	1.265	781	6.989

O comparativo mostra que os filmes em português possuem maior número de legendas do que os filmes em inglês, com exceção do filme *The social network*, que tem mais legendas em inglês. Esse dado já mostra que a LO em português é mais segmentada do que a LO em inglês, em outras palavras, no que diz respeito aos filmes estrangeiros em DVD, o texto audiovisual em português é dividido em maior número de legendas do que o texto audiovisual em inglês.

Sobre o tamanho médio das palavras, a LO em inglês e a LO em português têm palavras de praticamente mesmo tamanho, 3,71 e 3,79 caracteres, respectivamente. E o número médio de palavras por sentença em inglês, 5,37, é maior que em português, 4,48. Esses dados mostram que legendas em inglês apresentam palavras curtas e sentenças mais longas, enquanto legendas em português também apresentam palavras curtas, porém, sentenças mais curtas, com menos palavras que as sentenças em inglês. O número de caracteres por palavra e o número de palavras por sentença pode ser um padrão do idioma e uma característica do tipo de texto – audiovisual legendagem.

As legendas em inglês e em português são compostas de palavras com praticamente o mesmo tamanho (3,7 caracteres), porém, há mais e maiores sentenças nas legendas em inglês do que nas legendas em português. O legendista também pode influenciar a segmentação do texto audiovisual, uma vez que as escolhas são baseadas nos parâmetros técnicos e linguísticos da legendagem, nas escolhas subjetivas do tradutor, entre outros fatores. Dos cinco filmes, apenas três apresentam os tradutores responsáveis pelas legendas em português: *A rede social*, *Madagascar* e *Êxodo*.

Sobre a densidade lexical, representada pela razão de palavras distintas (*type*) pelas palavras corridas (*token*) – *type/token ratio (TTR)* –, é possível considerar que, embora possuam mais palavras, as legendas em inglês são menos densas, ou seja, possuem um léxico menos variado, por possuírem menos palavras distintas (*types*) (vide indicação 2 da figura 15). A LO dos cinco filmes em português apresenta uma densidade lexical maior (vide indicação 3 da figura 15).

Por exemplo, a maior legendagem, a do filme *The social network_EN*, contém 17.109 palavras em inglês, sendo 2.467 palavras distintas, e *The social network_BR* contém 10.856 palavras em português, sendo 2.683 palavras distintas. Há uma diferença de 6.253 palavras da legenda em inglês para a legenda em português. A densidade lexical das legendas em inglês é de 14,42, enquanto a das legendas em português é de 24,71, ou seja, 14,42% das legendas do filme em inglês correspondem a palavras distintas, enquanto 24,72% das legendas do filme em português são de palavras distintas. Isso significa que, em número de palavras, as legendas em inglês são maiores, porém, em diversidade de palavras, ou densidade lexical, nas legendas em português é maior. Essa característica se repete em todo o *corpus*, como mostra a comparação a seguir:

<i>Type/token ratio - TTR</i>	<i>Social network</i>	<i>Click</i>	<i>Madagascar</i>	<i>Exodus</i>	<i>Fightplan</i>	<i>Overall</i>
Inglês	14,42	16,28	18,11	19,68	19,19	10,27
Português	24,71	21,02	25,39	27,38	25,68	16,18

A LO dos cinco filmes em inglês é maior, porém, com riqueza lexical menor. E a LO dos cinco filmes em português é menor, porém com maior riqueza lexical. Esses dados podem ser visualizados na indicação 3 da figura 15.

Com relação ao número de linhas, das 7.030 legendas em inglês, 3.871 (55%) são de uma linha e 3.159 (45%) são de duas linhas. Das 6.989 legendas em português, 4.305 (62%) são de uma linha e 2.684 (38%) são de duas linhas. Tanto no *corpus* de legendas em inglês quanto em português, há mais legendas de uma linha. As legendas de duas linhas, presentes em menor quantidade no *corpus* da atual pesquisa, são melhor processadas que as legendas de uma linha, pois contêm conteúdo informativo maior, como afirmam D'Ydewalle e de Bruycker (2003). Embora haja problemas de segmentação entre legendas de uma linha e estas sejam maioria nos filmes em DVD, em inglês e português, nesta pesquisa, não houve ocorrência de problemas de segmentação em legendas de uma linha.

Em resumo, o panorama do *corpus* revela que a LO em inglês possui menor número de legendas, entretanto, são legendas mais volumosas, com mais palavras e sentenças mais longas. Apesar de apresentarem mais palavras, têm um léxico pobre, ou seja, as palavras se repetem muito, uma vez que há um percentual baixo (14,4%) de palavras distintas e um percentual alto de palavras repetidas (85,6%). Assim como as legendas em português, as legendas em inglês possuem palavras com tamanhos médios semelhantes, e a maioria das legendas é de uma linha. Já a legendagem em português é mais numerosa, porém com legendas menos volumosas, com menos palavras e sentenças mais curtas. Apesar de serem menos volumosas, as legendas em português têm um léxico mais rico que o do inglês, pois apresentam 24,7% de palavras distintas.

O panorama do *corpus* esclarece algumas características das legendas em inglês e português e ajuda a preparar o *corpus* para a análise. Foram analisadas legendas como um todo para verificar o padrão de segmentação e velocidade, e para investigar problemas de segmentação foram analisadas legendas de duas linhas. Sigo para a análise dos problemas de segmentação e velocidade na LO em inglês e na LO em português, com a noção de que maior conteúdo informativo é melhor recebido em legendas de duas linhas (VIEIRA, 2016) e maior conteúdo informativo pode causar desconforto quando recebido em legendas de uma linha (MONTEIRO, 2016). No presente *corpus*, apesar da pequena diferença numérica, ambas as legendas, em inglês e português, apresentam volume de informações verbais semelhante, sendo que em português as informações são mais densas.

4.2 COMPARAÇÃO DOS PROBLEMAS DE SEGMENTAÇÃO E VELOCIDADE NAS LEGENDAS DE DVD PARA OUVINTES: INGLÊS × PORTUGUÊS

Esta subseção investiga como se caracterizam os problemas de segmentação e a velocidade nas legendas para ouvintes, de filmes em DVD, nos contextos de línguas inglesa e portuguesa. Para isso, apresento os dados referentes aos problemas de segmentação e velocidade das legendas para ouvintes dos cinco filmes em DVD (*Flightplan* [2004] – suspense; *Madagascar* [2005] – animação; *Click* [2006] – comédia; *The social network* [2010] – drama; *Exodus: Gods and kings* [2015] – ação). O enfoque é a comparação de legendas para ouvintes em inglês e em português. A apresentação será feita com base na comparação das duas línguas: primeiro, comparo os dados referentes à segmentação; em seguida, os dados referentes à velocidade.

4.2.1 Problemas de segmentação nas legendas em inglês e em português

Os problemas de segmentação linguística da LO de filmes em DVD acontecem em legendas de duas linhas e têm baixa frequência no *corpus*, 4,2% em inglês e 4,7% em português. Em inglês, das 7.030 legendas, 3.871 (55%) são de uma linha e 3.159 (45%) são de duas linhas. Das 6.989 legendas em português, 4.305 (62%) são de uma linha e 2.684 (38%) são de duas linhas. As legendas de uma linha são maioria no *corpus*, o que indica que, apesar de haver poucos problemas de segmentação linguística, a própria distribuição do texto audiovisual pode ser um problema de segmentação, já que na maioria das vezes está atrelada a outros problemas de parâmetros.

A literatura sobre o processamento de legendas apresenta discussões acerca do processamento de legendas de uma e de duas linhas, a exemplo de d'Ydewalle *et al.* (1987), que verificaram que legendas de duas linhas causam menos problemas e levam menos tempo para serem processadas do que as de uma linha. D'Ydewalle e de Bruycker (2003), a partir de novo experimento, explicam que o tempo gasto na leitura de legendas é proporcional ao esforço cognitivo realizado; desse modo, quanto maior for o tempo de leitura, maior será o esforço realizado, e o esforço extra pode acarretar desconforto ou entrave no processamento de legendas. Porém, os autores explicam que, embora levem mais tempo para serem processadas, legendas de duas linhas são processadas fluentemente, por terem maior conteúdo informativo não inferido pela imagem. Já as legendas de uma linha são processadas em menos tempo, mas deixam de ser fixadas com frequência porque apresentam menor conteúdo informativo e, por isso, são processadas com menos fluência. Desse modo, os pesquisadores acreditam que o maior conteúdo informativo das legendas de duas linhas favorece mais a recepção do que legendas de uma linha.

Por outro lado, o maior conteúdo informativo expresso nas legendas rápidas de uma linha, em programas políticos, pareceu causar desconforto na recepção de surdos e ouvintes na pesquisa de Monteiro (2016), que verificou que as legendas de uma linha processadas por surdos e ouvintes em quatro condições experimentais diferentes (lenta mal segmentada, lenta bem segmentada, rápida mal segmentada e rápida bem segmentada) foram recebidas com desconforto pela maioria dos participantes. O desconforto ficou evidente pela recorrência de relatos de legendas rápidas e muito rápidas, mesmo quando as legendas eram lentas, e pela recorrência de relatos sobre a poluição visual dos programas políticos, apresentando excesso de conteúdos visuais e verbais que competiam por atenção com as legendas, além de palavras desconhecidas e conteúdos densos apresentados sempre em uma

linha. Apesar dos relatos de rapidez das legendas, Monteiro (2016) acredita que o desconforto pode não ter sido causado pela velocidade, mas sim pelo fato de serem legendas de uma linha, que apresentam menos informações verbais do que as de duas linhas, o que pode exigir mais esforço dos espectadores, acrescida da poluição visual dos programas políticos. Monteiro (2016) percebeu que as legendas de uma linha na velocidade rápida exigiram mais atenção dos espectadores quanto à completude do conteúdo, o que também pode ter causado o desconforto em relação à recepção, mesmo não tendo sido um entrave à compreensão, de acordo com os relatos da maior parte dos participantes. A maioria das legendas para ouvintes em inglês e português analisadas nesta pesquisa são de uma linha, fato que pode incidir na qualidade, independentemente da velocidade e da segmentação linguística.

No *corpus* de legendas de filmes em DVD não há problemas de segmentação linguística em legendas de uma linha, mas há em legendas de duas linhas, que, mesmo sendo processadas mais naturalmente, podem ter a qualidade da recepção comprometida, caso apresentem problemas de segmentação. De acordo com Vieira (2016, p. 97), os problemas de segmentação podem perturbar a recepção ao acentuar os movimentos regressivos para a esquerda. Os movimentos regressivos, ou regressões, podem ocorrer em um sintagma, em uma mesma palavra, em palavras diferentes ou entre o fim e o início das linhas de legendas. Vieira (2016) mostra que, na recepção de documentários legendados, ouvintes e surdos fazem mais regressões nos sintagmas de legendas mal segmentadas e, por isso, gastam mais tempo no processamento de legendas com problemas de segmentação.

Embora a LO de filmes em DVD apresente poucos problemas de segmentação linguística, há muitas legendas de uma linha, o que pode dificultar tanto a recepção quanto o processamento. Na LSE de filmes em DVD, o baixo percentual de problemas de quebra de linha (4,8%) e o expressivo número de legendas de uma linha (47,2%) também são dados que se repetem, conforme resultados de Chaves e Amoreira (2015).

A realidade da LSE *closed caption* da TV brasileira é diferente, dado que as pesquisas exploratórias (FRANCO; ARAÚJO, 2003; ARAÚJO, 2004; ARAÚJO, 2008; ARAÚJO; NASCIMENTO, 2011; ARAÚJO, 2012) já demonstraram a deficiência da LSE em relação às suas características técnicas, sobretudo velocidade e segmentação. Também é fato que há influência dos parâmetros de segmentação e velocidade sobre a recepção de surdos e ouvintes, conforme os resultados de Araújo e Nascimento (2011), Araújo (2012) e Vieira (2016). As legendas da TV brasileira seguem o modelo americano de *closed caption*, são feitas por profissionais não tradutores, correspondem, quase sempre, às transcrições das falas e são exibidas, na maioria dos casos, em letras maiúsculas e, em alguns casos, em letras

minúsculas, como os filmes. A produção das legendas também é diferente: enquanto as legendas de DVD são feitas com auxílio de programas editores de legendas que permitem etapas de marcação, redação e revisão, no *closed caption* da TV geralmente se utiliza a estenotipia ou programas de reconhecimento de voz ou refalamento para agilizar a confecção de legendas ao vivo, o que inviabiliza as etapas de marcação e de revisão. Essas diferenças reforçam a especificidade da atividade e a importância da formação na área.

Comparando LO em inglês e LO em português, os tipos e subtipos de problemas de segmentação linguística são semelhantes em ambas as línguas. Tanto em inglês quanto em português, as legendas de filmes apresentam maior frequência de problemas no sintagma verbal (SV), seguido dos sintagmas nominal (SN) e preposicional (SP), das orações subordinadas (SUBORD) e coordenadas (COORD), e dos sintagmas adjetival (SAdj) e adverbial (SAdv), conforme especificados no quadro 05.

Quadro 05 – Quantitativo de problemas de segmentação linguística nas legendas em línguas inglesa e portuguesa

PROSEGL	SV	SN	SP	OS	OC	SAdj	SAdv	Total
Legendas em inglês	140	49	43	28	13	12	9	295
Legendas em português	100	93	59	34	26	13	6	333

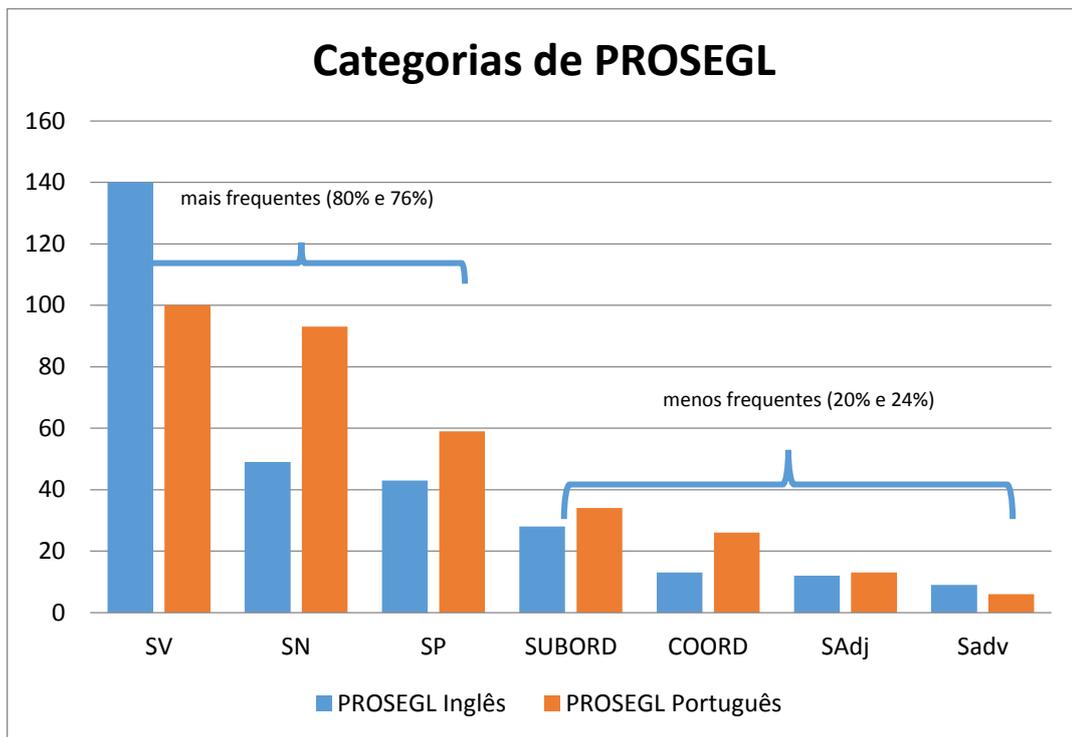
Fonte: Elaborado pela autora.

Em inglês, são mais frequentes os problemas de segmentação do sintagma verbal (140 ou 47%), seguido do sintagma nominal (49 ou 17%), sintagma preposicional (43 ou 15%), oração subordinada (28 ou 9%), oração coordenada (13 ou 4%), sintagma adjetival (12 ou 4%) e sintagma adverbial (9 ou 3%). As legendas de língua portuguesa seguem a mesma ordem de frequência, mas apresentam frequências de ocorrência diferentes: sintagma verbal (100 ou 30%), seguido do sintagma nominal (93 ou 28%), sintagma preposicional (59 ou 18%), oração subordinada (34 ou 10%), oração coordenada (26 ou 8%), sintagma adjetival (13 ou 4%) e sintagma adverbial (6 ou 2%).

Nas legendas em inglês, os problemas de segmentação no sintagma verbal representam quase a metade dos problemas de segmentação (47%); a segunda categoria mais frequente, a do sintagma nominal, ficou bem distante da primeira, com 17% de frequência. Já nas legendas em português, os percentuais de problemas nos sintagmas verbal e nominal são muito próximos: 30% e 28%, respectivamente, sendo os no sintagma nominal mais

representativos em português. Em ambas as línguas, os problemas no sintagma preposicional são semelhantes, 15% e 18%, respectivamente, sendo o terceiro mais frequente no *corpus*. As demais categorias de problemas (SUBORD, COORD, SAdj e SAdv) têm frequências de até 10%, e juntas somam 20% em inglês e 24% em português. O gráfico 01 mostra a comparação de problemas de segmentação linguística em inglês e português.

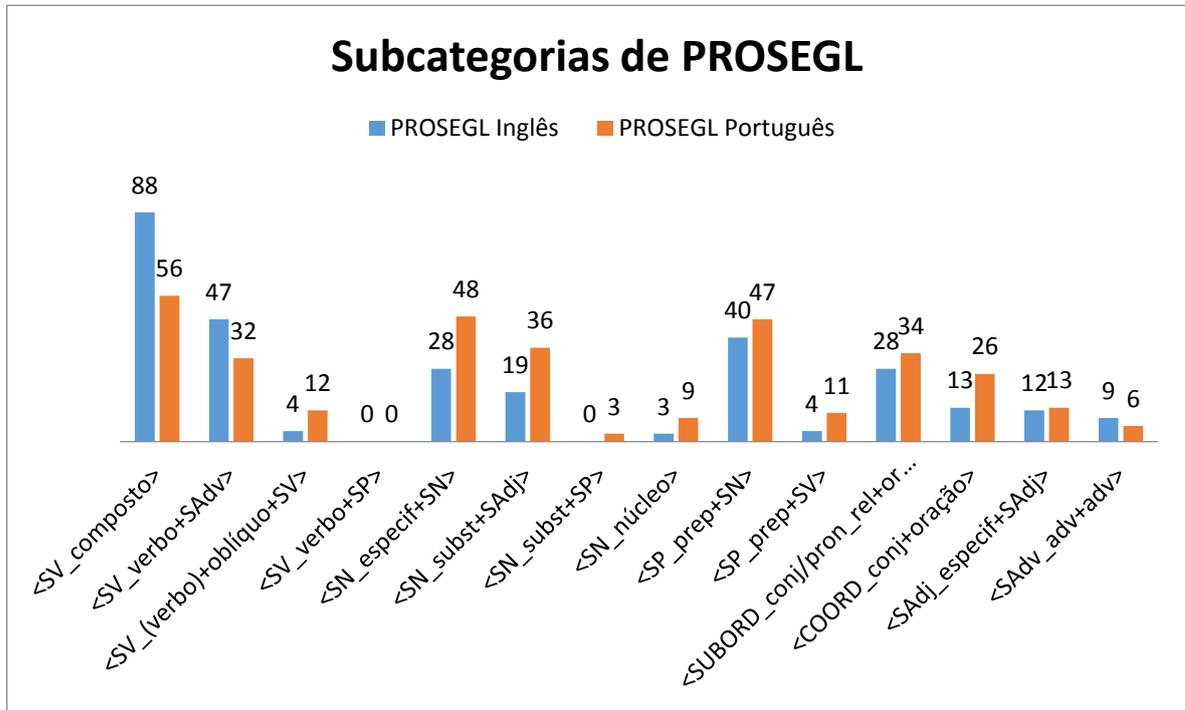
Gráfico 01 – Categorias de problemas de segmentação linguística: ocorrências em inglês e português



Fonte: Elaborado pela autora.

O gráfico 01 permite uma visualização em colunas dos problemas de segmentação linguística, que foram divididos em dois grupos de acordo com a língua: à esquerda, em colunas em azul, os dados do inglês, e à direita, em colunas laranja, os do português. As três categorias de PROSEGL mais frequentes em inglês correspondem a 80%, e em português, elas somam 76%. Antes de fazer uma discussão sobre os PROSEGL, segue o gráfico 02 com a ocorrência das 14 subcategorias de problemas de segmentação na LO em ambas as línguas.

Gráfico 02 – Subcategorias de problemas de segmentação linguística: ocorrências em inglês e português



Fonte: Elaborado pela autora.

O gráfico 02 mostra os 14 tipos ou subcategorias de problemas de segmentação linguística em comparação nas duas línguas. As sete mais frequentes, juntas, somam 89% dos problemas em inglês e 84% em português, são elas: <SV_composto>, <SV_verbo+SAdv>, <SN_especif+SN>, <SN_subst+SAdj>, <SP_prep+SN>, <SUBORD_conj/pron_rel+oração>, <COORD_conj+oração>. As sete subcategorias menos frequentes, juntas, somam 11% em inglês e 16% em português, são elas: <SV_(verbo)+oblíquo+SV>, <SV_verbo+SP>, <SN_subst+SP>, <SN_núcleo>, <SP_prep+SV>, <SAdj_especif+SAdj>, <SAdv_adv+adv>.

Os problemas de segmentação aconteceram em legendas de 1 e de 2 linhas, sendo os problemas de segmentação linguística nas legendas de 2 linhas e problemas de segmentação de texto audiovisual em legendas, em legendas de 1 linha. Segundo a literatura sobre a recepção e o processamento de legendas, legendas de 2 linhas são lidas mais naturalmente que as legendas de 1 linha (D'YDEWALLE *et al.*, 1987; D'YDEWALLE; DE BRUYCKER, 2003), e legendas de 1 linha têm sua recepção e processamento prejudicados pelo fato de causarem sensação de desconforto e de que a legenda está rápida (MONTEIRO, 2016) e pelo fato de gerarem dados de excesso de deflexão e de fixação (D'YDEWALLE *et al.*, 1987; D'YDEWALLE; DE BRUYCKER, 2003). Portanto, legendas com problemas de segmentação, sejam de uma ou duas linhas, prejudicam a recepção (MONTEIRO, 2016;

VIEIRA, 2016). No *corpus* desta pesquisa, as legendas de duas linhas não são maioria, sendo 45% em inglês e 38% em português, mas, por serem lidas mais naturalmente, expressam qualidade, exceto se estiverem em interação com problemas de segmentação linguística e/ou velocidade abaixo de 145ppm. Já as legendas de 1 linha são maioria, sendo 55% em inglês e 62% em português, dados que sugerem problemas de recepção e de processamento de legendas (D'YDEWALLE *et al.*, 1987; D'YDEWALLE; DE BRUYCKER, 2003; VIEIRA, 2016; MONTEIRO, 2016). Os problemas de segmentação linguística em legendas de 1 e de 2 linhas causam problemas de recepção e de processamento (VIEIRA, 2016; MONTEIRO, 2016), assim como os problemas de segmentação do texto audiovisual em legendas, que não são problemas linguísticos de quebra de sintagmas e de orações, mas são problemas de harmonização dos parâmetros, como velocidade e número de caracteres por linha. Ambos os problemas de número de linha e de segmentação incidem na qualidade das legendas.

Dados os percentuais de problemas de segmentação em inglês e português, passo a explorar os padrões linguísticos de cada categoria, por ordem de frequência, comparando os contextos de língua inglesa e portuguesa, buscando compreender os elementos envolvidos com os problemas e os elementos envolvidos com a solução dos problemas. Utilizo a mesma nomenclatura para analisar os problemas de segmentação nas duas línguas, pelo fato de que o conjunto de etiquetas foi o mesmo para classificar as legendas em inglês e português.

Os problemas de segmentação mais frequentes foram ilustrados com fotogramas legendados, e os sete menos frequentes não foram discutidos, já que são pouco representativos no *corpus*. Em todos os exemplos, as legendas estão identificadas por suas etiquetas de localização no *corpus*¹⁶, contendo informações de tempo e velocidade. Esses exemplos foram extraídos com auxílio da ferramenta *Concord*. As legendas estão no formato original de duas linhas e, para melhor visualização, destacadas em negrito. A localização da quebra de linha é feita pelo sinal gráfico barra lateral (/). Os exemplos em inglês estão seguidos de suas traduções conforme legendas do DVD em português.

4.2.1.1 Problemas de segmentação no sintagma verbal

No sintagma verbal, os problemas de segmentação ocorrem pela separação de seus componentes, de quatro formas: separando verbo de verbo <SV_composto>, separando verbo de advérbio ou sintagma adverbial <SV_verbo+SAdv>, separando verbo de pronome

¹⁶ As legendas com problemas de segmentação linguística em inglês e português encontram-se nos anexos A e B desta tese.

oblíquo mais verbo ou sintagma verbal <SV_(verbo)+oblíquo+SV> e separando verbo de preposição ou sintagma preposicional <SV_verbo+SP>. Das quatro subcategorias, as duas primeiras são mais frequentes no *corpus*, e a última não teve ocorrência nas duas línguas.

<SV_composto>

A quebra do sintagma verbal composto é a mais frequente. O problema ocorre quando há quebra entre verbo auxiliar e o seu especificador, o qual pode estar no infinitivo, no gerúndio ou no particípio, como mostra a figura 16.

Figura 16 – Quebra de sintagma verbal <SV_composto> em inglês (a) e português (b)



Fonte: Elaborada pela autora. Fotogramas de filmes legendados: (a) *Madagascar* (2005); (b) *A rede social* (2010).

Na figura 16 (a), a legenda em inglês *So, what are you going to do? Go / running off to the wild by yourself?*¹⁷ tem duas orações, especificamente duas interrogações. Do ponto de vista da legendagem, supostamente cada oração ocuparia uma linha de legenda, pois a pontuação por si só já consiste na pausa que delimita a quebra de linha das legendas. Porém, a segunda oração *Go / running off to the wild by yourself?* apresenta 39 caracteres, quantidade que extrapola o número máximo de 37 caracteres por linha considerado nesta pesquisa. Esta pode ter sido a razão que motivou o deslocamento do verbo *Go* para a linha de cima. Do ponto de vista linguístico, a quebra do sintagma verbal *Go running off* (minha tradução: “sair correndo”) desestabiliza o todo semântico das orações construído pela relação de dependência entre os sintagmas.

Na figura 16 (b), a legenda em português *Fiz um juramento aqui e não quero / cometer perjúrio* apresenta uma oração coordenada, composta de duas orações ligadas pelo coordenador *e*. Do ponto de vista da legendagem, supostamente cada oração ocuparia uma linha de legenda, e o coordenador *e* acompanharia a segunda oração na linha de baixo.

¹⁷ Tradução oficial para o DVD em português: *O que vai fazer? Fugir para a natureza sozinho?*

Atendendo aos parâmetros técnicos e linguísticos, a legenda estaria equilibrada da seguinte forma: *Fiz um juramento aqui / e não quero cometer perjúrio*. Na legenda oficial do DVD, a quebra de linha foi feita no meio do sintagma verbal composto **quero / cometer**, e, para manter o sintagma verbal inteiro, seria necessário deslocar a quebra de linha para a esquerda ou para a direita.

A quebra de linha à direita resultaria na legenda: *Fiz um juramento aqui e não quero cometer / perjúrio.*, sem problemas de segmentação, porém com problemas técnicos no formato por extrapolar o máximo de 37 caracteres por linha. Se o deslocamento da quebra de linha fosse feito à esquerda, resultaria na legenda: *Fiz um juramento aqui e não / quero cometer perjúrio.*, com um novo problema de segmentação, o de sintagma verbal <SV_verbo+SAdv>, em que o advérbio *não* estaria separado do sintagma verbal composto *quero cometer*. Se o deslocamento da quebra da linha fosse mais à esquerda, resultaria na legenda *Fiz um juramento aqui e / não quero cometer perjúrio.*, com um novo problema de segmentação, o de oração coordenada <COORD_conj+oração>, em que a conjunção *e* acompanharia a primeira oração, na linha de cima, ainda fora dos parâmetros de segmentação.

Nos dois exemplos apresentados, a discussão sugere que a quebra dos sintagmas desmancha o todo semântico construído pela relação entre os sintagmas e que a segmentação deve ser feita em consonância com os demais parâmetros, nesse caso: número de caracteres por linha e formato. Apresento a seguir alguns exemplos de legendas com problemas de segmentação no sintagma verbal <SV_composto>, observando que esse é o tipo mais frequente tanto em inglês quanto em português.

Madagascar
<sub219>
<veloc_leg 68c/3.595s 18.915>

So, what are you going to do? **Go/**
running off to the wild by yourself?
(O que vai fazer?/ Fugir para a natureza sozinho?)

Madagascar
<sub4>
<veloc_leg 56c/2.901s 19.304>

Come on, Marty. Just **wanted/**
to wish you a happy birthday.
(Calma, Marty. Só queria te desejar/ feliz aniversário.)

Nos exemplos em inglês, os sintagmas verbais quebrados são formados por mais de dois verbos, sejam auxiliares, modais, infinitivos, gerúndio, participio: **Go / running off; wanted / to wish**. Nesses casos, a manutenção de um sintagma na mesma linha de legenda é importante para preservar o conteúdo informativo de cada linha e/ou das duas linhas juntas,

mantendo o todo semântico da legenda. Há uma visível diferença entre a opção 1 e a opção 2 abaixo:

<sub4> opção 1

Come on, Marty. Just wanted
to wish you a happy birthday.

<sub4> opção 2

Come on, Marty.
Just wanted to wish you a happy birthday.

Na opção 1, a leitura das orações *Come on, Marty. Just wanted* (Calma, Marty. Só queria) parece truncada por um ponto final no meio da linha. Também há uma suspensão de informações que só serão recuperadas com as da linha de baixo *to wish you a happy birthday*. (te desejar feliz aniversário.), que, por sua vez, também está incompleta. Já na opção 2, a leitura da linha de cima (*Come on, Marty.*) e da linha de baixo (*Just wanted to wish you a happy birthday.*) é fluente, porque ambas são independentes sintaticamente. As duas formam um todo semântico: *Come on, Marty. Just wanted to wish you a happy birthday*. (Em português: Calma, Marty. Só queria te desejar feliz aniversário.). Outros exemplos de problema no sintagma verbal <SV_composto> ocorrem em legendas com construções linguísticas como a mostrada a seguir:

Social Network
<sub1782>
<veloc_leg 42c/2.4s 17.500>

Muda pra cá ou **vai acabar/**
ficando pra trás.

Sintagmas verbais como **vai acabar / ficando; poderia / sair errado; precisa / acabar; quero / cometer** são exemplos de quebras sintáticas que podem comprometer a legenda como um todo, já que podem distorcer os significados expressos em cada linha de legenda e/ou em cada legenda. No exemplo <sub1782>, a legenda *Muda pra cá ou vai acabar / ficando pra trás* apresenta um problema de quebra no sintagma verbal que, em princípio, poderia ser resolvido com a junção do sintagma *vai acabar ficando*. Porém, percebo que unir o sintagma verbal não é a única ação para tornar a legenda fluente, pois se este ficar na linha de cima, vai gerar problema de formato e vai exceder o número de 37 caracteres por linha. Além disso, se o mesmo ficar na linha de baixo, vai gerar um novo

problema de segmentação na oração coordenada. Nesse caso, é preciso fazer a segmentação onde as duas linhas resultantes entreguem maior conteúdo semântico por linha.

A legenda *Muda pra cá ou vai acabar / ficando pra trás* é composta de duas orações coordenadas: *Muda pra cá* e *vai acabar ficando pra trás*. As duas orações são semanticamente independentes, o que significa que cada uma tem um sentido e que uma não modifica a outra. Nesse caso, a relação entre elas é marcada pelo coordenador, a conjunção *ou*. A conjunção acompanha a segunda oração e, por isso, na segmentação da legenda, deve ficar na linha de baixo, conforme sugestão: *Muda pra cá / ou vai acabar ficando pra trás*. Assim, a legenda teria duas linhas com sintagmas inteiros e informações completas.

<SV_verbo+SAdv>

A quebra do sintagma verbal pela quebra de verbo e sintagma adverbial <SV_verbo+SAdv> é frequente em todo o *corpus*. O problema ocorre quando há quebra entre verbo e advérbio ou sintagma adverbial. Isso porque os especificadores ou complementos dos verbos, que estão à direita ou à esquerda, respectivamente, indicam as circunstâncias (advérbios) de modo, tempo, negação que norteiam o verbo e, por isso, precisam estar junto ao verbo. Os exemplos da figura 17 ilustram essa subcategoria de problema em inglês e em português.

Figura 17 – Quebra de sintagma verbal <SV_verbo+SAdv> em inglês (a) e português (b)



Fonte: Elaborada pela autora. Fotogramas de filmes legendados: (a) *Flightplan* (2004); (b) *Click* (2006).

No fotograma à esquerda, a legenda em inglês *But I tell you what. I will personally / go and speak to the captain* apresenta três orações. Na quebra de linha, temos um problema no sintagma verbal pela quebra do advérbio *personally* e o verbo *go*. O problema de quebra de linha, por si só, pode ser facilmente resolvido unindo o verbo *go* ao seu sintagma

verbal *I will personally*, na linha de cima ou de baixo. Porém, essa ação teria como consequência o uso de número de caracteres maior do que o recomendado.

Nesse caso, como a legenda tem três orações para serem apresentadas em duas linhas de no máximo 37 caracteres, as estratégias devem ir além da segmentação. A sentença 1) *But I tell you what* (mas eu te digo) tem 20 caracteres, a oração 2) *I will personally go* (eu vou pessoalmente) tem 20 caracteres, e a oração 3) *and speak to the captain* (e falo com o comandante) tem 24 caracteres. Para dar conta da segmentação e manter o máximo de 37 caracteres por linha, é preciso reduzir as orações. A sugestão que apresento preza pelo mínimo de alterações na legendagem, portanto, une as orações menores, um e dois, na linha de cima, com 41 caracteres: *But I tell you what. I will personally go*, e a oração três na linha de baixo, com 24 caracteres: *and speak to the captain*. Uma redução de 4 caracteres na linha de cima harmoniza os parâmetros. Usando o parâmetro da redução, é válido excluir palavras ou substituí-las por uma de menor tamanho, assim sendo, excluo a conjunção *But* e alcanço 37 caracteres na linha de cima, ficando da seguinte forma:

*I tell you what. I will personally go/
and speak to the captain.*

O problema de segmentação exposto e discutido demonstra que os problemas de quebra de linhas nem sempre são resolvidos com a união do sintagma quebrado. Em alguns casos, especialmente em legendas densas, com muitos caracteres por linha, a união do sintagma e seu deslocamento para uma das linhas de legenda acaba gerando novos problemas de legendagem, por isso, é preciso estar atento ao equilíbrio dos parâmetros técnicos e linguísticos da legendagem.

Os sintagmas parecem ser as menores unidades de segmentação das legendas, visto que a união de algumas palavras em um sintagma agrega valor semântico e a união de sintagmas em uma oração agrega ainda mais valor semântico. Nem sempre é possível manter uma oração por linha, pois, como no exemplo apresentado, havia três orações para duas linhas. Dessa forma, é esperado que a quebra seja feita entre os sintagmas, de modo a preservar o todo sintático. Isso torna maior o desafio de segmentar e harmonizar os parâmetros.

No fotograma *b*, a legenda em português *Não, não, não. A vida está / apenas começando*. apresenta um problema no sintagma verbal pela quebra do verbo *está* e do advérbio *apenas*. Nesse caso, a legenda possui duas orações, claramente marcadas pelo ponto

final no meio da linha, sendo uma: *Não, não, não* e a outra: *A vida está apenas começando.*, na qual se localiza um sintagma nominal (*A vida*) e um sintagma verbal (*está apenas começando*). A sugestão é que cada oração ocupe uma linha, ficando da seguinte forma:

Não, não, não.

A vida está apenas começando.

A identificação de uma oração parece ser mais intuitiva do que a identificação de um sintagma, pois geralmente apresenta pontuação de pausas mais longas, como vírgula e ponto final, que demarcam os limites sintáticos da legenda.

As legendas a seguir apresentam outros exemplos de problemas no sintagma verbal <SV_verbo+SAdv> que, ao serem disponibilizados, podem servir de subsídio para legendistas identificarem com mais facilidade os tipos de problemas em suas próprias práticas.

Social Network
<sub642>
<veloc_leg 75c/3.959s 18.944>

To the best of your knowledge, had he **even/ begun work** on Harvard Connection?
(Sabia se ele havia começado a Harvard Connection?)

Flightplan
<sub158>
<veloc_leg 65c/3.629s 17.911>

But I tell you what. I **will personally/ go** and speak to the captain
(Mas vamos fazer assim: falarei com o capitão...)

Em português, os exemplos de legendas com problemas no sintagma verbal pela quebra de verbo e advérbio também podem servir para esclarecer melhor a subcategoria.

Click
<sub1016>
<veloc_leg 60c/3.026s 19.828>

Avancei o sexo uma vez, **não/ quero** que seja assim para sempre.

Click
<sub1475>
<veloc_leg 43c/3.299s 13.034>

Não, não, não. A vida **está/ apenas** começando.

4.2.1.2 Problemas de segmentação no sintagma nominal

No sintagma nominal, os problemas de segmentação ocorrem pela separação de seus componentes, de quatro formas, podendo ocorrer devido à separação de especificador e

sintagma nominal <SN_especif+SN>; substantivo e sintagma adjetivo <SN_subst+SAdj>; substantivo e sintagma preposicional <SN_subst+SP>; e substantivos compostos <SN_núcleo>. Dos problemas de segmentação no sintagma nominal, os que tiveram mais frequência, tanto nas legendas em inglês como nas legendas em português, foram <SN_especif+SN> e <SN_subst+SAdj>.

O sintagma nominal é uma construção sintática formada por um núcleo, por especificadores à esquerda e por complementadores à direita. Castilho (2012, p. 454) observou que o núcleo pode ser preenchido por substantivo e por alguns pronomes; os especificadores podem ser preenchidos por artigos, pronomes, quantificadores, expressões qualitativas e delimitadores; e os complementadores, pelos sintagmas adjetivais e preposicionais. Para discutir melhor os problemas de segmentação no sintagma nominal, a figura 18 ilustra o tipo de problema mais frequente desse sintagma, <SN_especif+SN>, em inglês e em português.

Figura 18 – Quebra de sintagma nominal <SN_especif+SN> em inglês (a) e português (b)



Fonte: Elaborada pela autora. Fotogramas de filmes legendados: (a) *A rede social* (2011); (b) *Madagascar* (2005).

Na figura 18(a), a legenda *Tyler spelled the usual way, and my / last name is the same as my brother's.* (em português, *Tyler grafado normal e meu sobrenome é o mesmo do meu irmão.*) apresenta uma quebra no sintagma nominal, por ter separado o especificador *my* (meu) do substantivo *last name* (sobrenome). Essa é uma legenda densa, com extenso conteúdo apresentado em duas linhas de 35 e 38 caracteres, respectivamente, em apenas 3,3 segundos. Tem velocidade alta, de quase 22 caracteres por segundo. Resolver o problema de segmentação, nesse caso, não pode envolver a união do sintagma e seu deslocamento para uma das linhas, pois já atingiram a capacidade máxima de 37 caracteres. Além da junção do sintagma, outra estratégia é fazer redução de caracteres.

A legenda apresenta duas informações expressas em duas sentenças. É interessante pensar a segmentação de modo que cada informação ocupe uma linha e que os sintagmas estejam íntegros. A primeira sentença, *Tyler spelled the usual way*, tem 28 caracteres, e a segunda, *and my last name is the same as my brother's.*, tem 45 caracteres. Se cada uma ocupar uma linha, a legenda fica fora dos parâmetros técnicos de formato e de número de caracteres por linha:

(28c) *Tyler spelled the usual way,*

(45c) *and my last name is the same as my brother's.*

Com essa estratégia, o problema de segmentação é resolvido, mas outros são criados. Para sanar o problema da velocidade, sempre tendo em mente fazer o mínimo possível de alterações no texto, vejo que é possível diminuir caracteres na primeira sentença trocando *spelled the usual way* por *as usual*, que mantém o mesmo significado e reduz a velocidade da legenda:

(15c) *Tyler as usual,*

(45c) *and my last name is the same as my brother's.*

Agora, com menos caracteres na primeira linha, é possível deslocar algumas palavras para a linha de cima. Para garantir que não haja problemas de segmentação ao deslocar palavras soltas, desloco um sintagma inteiro, o sintagma nominal *and my last name*.

(32c) *Tyler as usual, and my last name*

(28c) *is the same as my brother's.*

Para sanar um problema de segmentação e manter a legenda dentro dos parâmetros técnicos e linguísticos, foi preciso utilizar estratégias de segmentação e de redução. O resultado foi uma legenda de 60 caracteres para ser exibida em 3,3s, o que confere uma velocidade de 18,1cps.

Na figura 18(b), o exemplo de quebra do sintagma nominal em português "*Paro aqui, trote ali, como **uma** / **graminha**, ando de volta para cá*". traz o especificador **uma** separado de seu núcleo, o substantivo **graminha**. A junção de especificador e núcleo é importante para reforçar o sintagma e contribuir com o todo semântico da legenda. Para juntar

o sintagma, especificador e substantivo **uma graminha** devem ficar juntos na linha de cima ou na linha de baixo. A escolha da linha para a qual se vai deslocar o sintagma deve levar em consideração os demais parâmetros da legendagem, procurando fazer o mínimo de alterações no texto.

A legenda possui quatro orações, todas separadas por vírgula: 1) *Paro aqui*, 2) *troto ali*, 3) *como uma graminha*, 4) *ando de volta para cá*. O problema de segmentação está localizado no sintagma nominal da terceira oração. Dessa forma, a sugestão que dou para sanar esse problema de segmentação é manter duas orações na linha de cima e duas na linha de baixo. Isso garante que o sintagma vai ficar íntegro, preservando sintaxe e semântica. Com tais mudanças, a linha de baixo deve sofrer redução, já que extrapola o número de caracteres. Aplicando segmentação e redução, o resultado é o seguinte: "*Paro aqui, troto ali,/ como uma graminha, volto para cá*".

Outros exemplos em inglês ilustram o problema de segmentação no sintagma nominal pela quebra de especificador e núcleo, como: **two / Twinkie burglars** (quantificador + sintagma nominal); **your first / construction order** (expressão qualitativa + sintagma nominal).

Click
<sub3>
<veloc_leg 56c/3.117s 17.966>

How can I sleep with **two/**
Twinkie burglars roaming around?
(Como posso dormir com dois ladrões de Twinkie por aí?)

Exodus
<sub549>
<veloc_leg 70c/2.4s 29.167>

it should have been **your first/**
construction order on taking the throne.
(devia ter sido seu 1º decreto quando assumiu o trono.)

Em português, os exemplos mostram quebras de especificador e núcleo, como: **uma / chance** (artigo + substantivo); **uma / graminha** (quantificador + substantivo), **uma / confusão** (artigo + substantivo); **uma / criança** (artigo + substantivo).

Madagascar
<sub28>
<veloc_leg 64c/4.835s 13.237>

"Paro aqui, troto ali, como **uma/**
graminha, ando de volta para cá".

Click
<sub906>
<veloc_leg 49c/2.264s 21.643>

Infelizmente, houve **uma/**
confusão no trabalho dele.

<SN_subst+SAdj>

Outro problema de segmentação frequente no *corpus* é a quebra do sintagma nominal pela separação de substantivo e sintagma adjetival. Isso porque, conforme Castilho (2012, p. 455), o sintagma adjetival é um dos complementadores do substantivo. A figura 19 ilustra problemas em inglês e em português.

Figura 19 – Quebra de sintagma nominal <SN_subst+SAdj> em inglês (a) e português (b)



Fonte: Elaborada pela autora. Fotogramas de filmes legendados: (a) *Click* (2006); (b) *Exodus* (2015).

Na figura 19(a), a legenda em inglês *Mr. Ammer, if you need female / companionship this weekend...* (Sr. Ammer, se precisar de companhia feminina esse fim de semana...) mostra que o problema de segmentação no sintagma nominal ocorreu pela quebra entre o adjetivo *female* e o substantivo *companionship*. A ordem do problema pode ser inversa à da etiqueta, não importando se o adjetivo vem antes ou depois. Nesse caso, o adjetivo vem antes porque, na língua inglesa, os qualificadores antecedem os substantivos (*nouns*). Mas em português o adjetivo pode preceder ou suceder o substantivo.

Já na figura 19(b), a legenda em português *travando uma luta / desesperada pela vida.* apresenta problema de segmentação pela quebra do substantivo *luta* e o adjetivo *desesperada* que formam o sintagma nominal. Em ambas as legendas, o problema pode ser sanado, em princípio, ao unir os constituintes do sintagma nominal (subst+SAdj), na linha de cima ou na linha de baixo. Mas é preciso atenção para que a segmentação seja restabelecida em harmonia com os demais parâmetros técnicos e linguísticos da legendagem.

Em inglês, os problemas de segmentação no sintagma nominal pela quebra de sintagma adjetival e substantivo ocorrem nas seguintes estruturas: **female / companionship** (companhia feminina); **this great / open plan thing** (esse ótimo espaço).

Click
<sub499>
<veloc_leg 58c/2.594s 22.359>

Mr. Ammer, if you need **female/ companionship** this weekend...
(Sr. Ammer, se precisar de companhia feminina esse fim de semana...¹⁸)

Madagascar
<sub548>
<veloc_leg 80c/5.199s 15.388>

First they tell you, "We got **this great/ open plan thing**. Let animals run wild."
(Primeiro dizem: "Temos esse ótimo espaço para soltar os animais".)

Em português, os exemplos mostram quebras de estruturas como: **cabelo / esquisito; luta / desesperada**. Para pensar em solucionar os problemas de segmentação, é preciso harmonizar os demais parâmetros, conforme discutido em exemplos anteriores.

Madagascar
<sub645>
<veloc_leg 35c/2.901s 12.065>

Tem algo naquele do **cabelo/ esquisito**

Exodus
<sub106>
<veloc_leg 39c/2.434s 16.023>

travando uma **luta/ desesperada** pela vida.

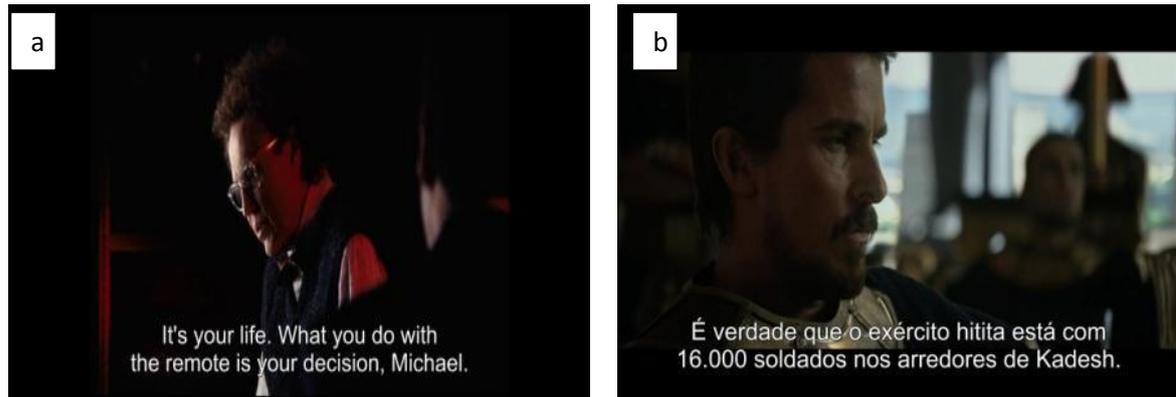
A legenda em português <sub689> *É um caminho / estreito e perigoso* apresenta problema de segmentação pela quebra do substantivo **caminho** e o sintagma adjetival que lhe complementa, **estreito e perigoso**. Para sanar o problema de segmentação linguística, basta unir o sintagma na mesma linha.

4.2.1.3 Problemas de segmentação no sintagma preposicional

No sintagma preposicional, os problemas de segmentação ocorrem pela separação de seus componentes, de duas formas: devido à separação de preposição e sintagma nominal <SP_prep+SN>; e de preposição e sintagma verbal <SP_prep+SV>. A quebra de preposição e sintagma nominal <SP_prep+SN> é bastante frequente no *corpus*, enquanto a quebra de preposição e sintagma verbal não ocorre no *corpus*. Para Castilho (2012), a estrutura desse sintagma é a mesma dos demais: (especificadores) + núcleo + (complementadores), porém, nessa subcategoria, o núcleo do sintagma é uma preposição. Os exemplos em inglês e português demonstram como a quebra do sintagma preposicional ocorre nas legendas.

¹⁸ Tradução de minha autoria. Diferente da tradução para o DVD em português: *Sr. Ammer, se precisar de companhia feminina...* que condensa o texto em português e por isso não foi utilizada para ilustrar a análise.

Figura 20 – Quebra de sintagma preposicional <SP_prep+SN> em inglês (a) e português (b)



Fonte: Elaborada pela autora. Fotogramas de filmes legendados: (a) *Click* (2006); (b) *Exodus* (2015).

<SP_prep+SN>

Na figura 20(a) *It's your life. What you do with / the remote is your decision, Michael.*, houve a quebra da preposição **with** (com) e o sintagma nominal **the remote** (o controle). Para restabelecer o sintagma e deslocá-lo para a linha de cima ou de baixo, é necessário aplicar outros parâmetros da legendagem, como a redução, já que as duas linhas de legenda estão no limite de caracteres. A linha de cima tem 32, e a linha de baixo já tem os 37. Nesse caso, é interessante acomodar o sintagma na linha em que a redução for possível. Nesse caso, avalio que a palavra *Michael*, da linha 2, pode ser excluída da legenda sem prejuízos semânticos conforme apresento:

*It's your life. What you do
with the remote is your decision.*

Outros exemplos em inglês mostram como a quebra do sintagma preposicional e seu complementador sintagma nominal ocorrem. São quebras do tipo **with / the remote** (com o controle); **out / of this dump** (sair desta porcaria).

Click
<sub836>
<veloc_leg 69c/3.629s 19.014>

It's your life. What you do **with/
the remote** is your decision, Michael.
(A vida é sua. Você decide o que fazer com o controle.)

Click
<sub739>
<veloc_leg 70c/3.117s 22.457>

That way we can get the hell **out/
of this dump**, go to T.G.I. Friday's...
(Assim, poderemos sair desta porcaria, ir ao T.G.I. Friday's...)

Em português, os exemplos mostram quebras do tipo **do / Watsuhita** (preposição + substantivo); **com / sua amiga** (preposição + sintagma nominal composto de especificador e substantivo).

Click
<sub835>
<veloc_leg 57c/2.798s 20.372>

Eu disse: "Consiga a conta **do/ Watsuhita** e será promovido".

Click
<sub842>
<veloc_leg 62c/4.3s 14.419>

Esqueci de dizer que saí **com/ sua amiga** Janine no fim de semana.

4.2.1.4 Problemas de segmentação na oração subordinada

Na oração subordinada, os problemas de segmentação são frequentes e ocorrem pela separação da conjunção ou do pronome relativo da oração que segue, ou seja, da segunda oração. Castilho (2012, p. 337) explica que “tudo o que ocorre numa sentença simples ocorre numa sentença complexa”. Desse modo, o que funciona para os sintagmas também funciona para as orações, mas com mais de um verbo. A figura 21 apresenta coincidentemente o mesmo tipo de problema na mesma legenda, em inglês (a) e em português (b).

Figura 21 – Quebra de oração subordinada <SUBORD_conj/pron_rel+oração> em inglês (a) e português (b)



Fonte: Elaborada pela autora. Fotogramas de filmes legendados: (a) *Click* (2006); (b) *Click* (2006).

Tanto o exemplo em inglês: *Don't yell at Peanut because / he found your Twinkie stash.* quanto o exemplo em português: *Não grite com Peanut porque / descobriu seu esconderijo.* mostram uma quebra da oração subordinada devido à separação da

conjunção em inglês *because* e em português *porque* e suas orações seguinte: *he found your Twinkie stash. e descobriu seu esconderijo.*

Click
<sub955>
<veloc_leg 56c/2.56s 21.875>

Don't yell at Peanut **because/**
he found your Twinkie stash.
(Não grite com Peanut porque descobriu seu esconderijo.)

Em português, há quebra da conjunção *porque* com sua oração subsequente.

Click
<sub976>
<veloc_leg 53c/2.696s 19.659>

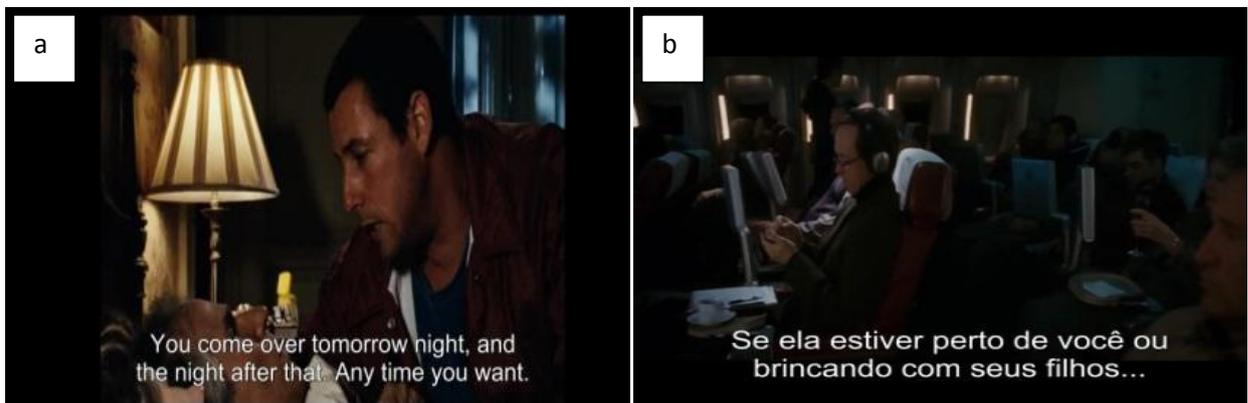
Não grite com Peanut **porque/**
descobriu seu esconderijo.

As orações são compostas de sentenças complexas e, por isso, são geralmente grandes, podendo ocupar as duas linhas. Como na oração subordinada há uma relação de dependência entre as orações e a conjunção ou o pronome relativo de fato modifica o significado da segunda oração, é importante manter a conjunção ou o pronome relativo na linha de baixo, junto à oração.

4.2.1.5 Problemas de segmentação na oração coordenada

Na oração coordenada <COORD__conj+oração>, os problemas de segmentação ocorrem pela separação da conjunção e da oração que acompanha. Diferentemente das subordinadas, as coordenadas são orações independentes, e a conjunção, ou coordenador, como designa Castilho (2012), não modifica o significado das orações, ou coordenadas, do ponto de vista semântico. Mesmo assim, o coordenador deve acompanhar a oração seguinte, ficando na linha de baixo da legenda.

Figura 22 – Quebra de oração coordenada <COORD_conj+oração> em inglês (a) e português (b)



Fonte: Elaborada pela autora. Fotogramas de filmes legendados: (a) *Click* (2006); (b) *Flightplan* (2004).

A figura 22 (a) e (b) mostra a quebra da oração em inglês *You come over tomorrow night, and / the night after that. Any time you want.* (Venham amanhã e depois de amanhã também. Quando quiserem.) porque o coordenador *and* (e) ficou na linha de cima e não acompanhou a segunda oração. A legenda em inglês é densa e já está com as linhas cheias, assim, a harmonização dos parâmetros torna-se um desafio maior. Uma sugestão seria deslocar a conjunção para a linha de baixo e reduzir a última sentença, *Any time you want* para *At any time.*

Já a oração em português: *Se ela estiver perto de você ou / brincando com seus filhos...* mostra que o coordenador *ou* ficou separado da oração seguinte: *brincando com seus filhos.* Nesse caso, como as linhas não são tão cheias quanto na legenda em inglês, é possível deslocar o coordenador *ou* sem causar problemas de outra natureza à legenda.

Em inglês, outros problemas de segmentação na oração coordenada mostram a frequência da conjunção *and* no *corpus*.

Click
</sub1463>
<veloc_leg 73c/3.31s 22.054>

You come over tomorrow night, **and/**
the night after that. Any time you want.
(Venham amanhã e depois de amanhã também. Quando quiserem.)

Madagascar
<sub1055>
<veloc_leg 73c/4.061s 17.976>

We can go back to civilization, **and/**
everything will be like it used to be.
(E podemos voltar à civilização e tudo voltará a ser como antes.)

Click
<sub79>
<veloc_leg 71c/3.197s 22.208>

Let's just take out the atrium **and/**
run with that and there's your drain.
(Espero que a ideia do átrio não tenha demorado, Michael.)

Em português, quebras de outras conjunções ou coordenadores aparecem no *corpus*: **nem e ou**.

Click
<sub778>
<veloc_leg 60c/3.868s 15.512>

Não é uma bicicleta, **nem/
um chapéu de caubói**, mas vai gostar.

Flightplan
<sub139>
<veloc_leg 59c/2.935s 20.102>

Se ela estiver perto de você **ou/
brincando com seus filhos...**

As discussões sobre os problemas de segmentação linguística tratam dos elementos envolvidos com os problemas e os envolvidos com a solução. Nos exemplos de ressegmentação, a sugestão foi fazer o deslocamento de um sintagma como um todo para uma das linhas de legenda, ou até mesmo de uma oração, desde que atenda aos demais parâmetros da legendagem, como velocidade e número de caracteres por linha. Isso porque a parte linguística por ela própria não demonstra ser capaz de suprir os problemas de segmentação. Isso ficou claro nos exemplos em que, ao unir novamente o sintagma em uma única linha de legenda, um novo problema de ordem técnica ou linguística foi gerado, sendo preciso recorrer a outros parâmetros da legendagem para adequar o produto final. Os parâmetros estão fortemente ligados, o que explica a necessidade de harmonização em prol da qualidade da legenda e da recepção.

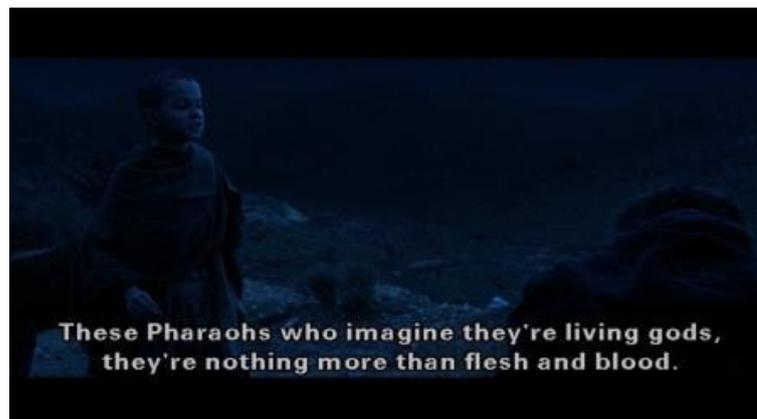
Os problemas de segmentação linguística das legendas em inglês e português são semelhantes pela baixa frequência de ocorrência, partilham os mesmos tipos de problemas e a mesma ordem de frequência. O que ainda parece ser um entrave à segmentação, mesmo em legendas feitas por profissionais legendistas, é manter íntegros sintagmas com dois ou mais verbos; sintagmas compostos por verbo e sintagma adverbial; sintagmas com especificador e sintagma nominal; sintagmas com substantivo e sintagma adjetival; sintagmas com preposição e sintagma nominal; orações coordenadas ou subordinadas íntegras, com a conjunção acompanhando a segunda oração. Os sintagmas demonstram ser as menores unidades de segmentação de uma legenda, já que a quebra deles recai sobre a qualidade da legenda e as soluções para os exemplos aqui apresentados sugerem sempre a remoção do sintagma como um todo, em vez de deslocar palavras soltas.

As questões que parecem estar mais envolvidas com os problemas de segmentação linguística em inglês e português são: a quebra de sintagmas, de conjunção/

oração e de outras informações que só serão recobradas na linha seguinte. Outra questão presente foi que uma estratégia de ressegmentação pode resolver um problema, mas gerar outro técnico ou linguístico, por isso é preciso atenção para que a segmentação seja restabelecida em harmonia com os demais parâmetros da legendagem. E as questões que parecem estar mais envolvidas com a solução desses problemas são segmentar em consonância com os demais parâmetros, como número de caracteres por linha e formato; manter um sintagma íntegro na mesma linha de legenda, tendo que deslocá-lo, fazer redução de palavras ou mesmo de caracteres para harmonizar segmentação, número de caracteres por linha e velocidade.

Os PROSEGL ocorreram em baixa frequência no *corpus*, mas não foram os únicos problemas de segmentação encontrados no *corpus*. A distribuição do texto audiovisual em legendas também não está de acordo com os parâmetros preconizados por pesquisadores da área, provocando, assim, a desarmonia dos parâmetros, como legendas muito rápidas ou muito lentas, com excesso ou falta de caracteres por linha, entre outros. Esses problemas aparecem tanto em legendas de duas quanto de uma linha, como no exemplo da figura 23 a seguir:

Figura 23 – Problema de segmentação pela distribuição de texto nas legendas de duas linhas em inglês



```
<sub980><2L>
<t>01:38:30,904 --> 01:38:35,364</t>
<cp147>These Pharaohs who imagine they're living gods,
<cp142>they're nothing more than flesh and blood.
<veloc_leg 89c/4.46s 19.955>
</sub980>
```

Fonte: Elaborada pela autora. Fotograma legendado em inglês do filme *Exodus* (2015).

Na figura 23, a legenda de duas linhas em inglês **These Pharaohs who imagine they're living gods, / they're nothing more than flesh and blood.** não apresenta PROSEGL; contudo, a falta de segmentação adequada, ou a má distribuição do texto audiovisual em legendas, contribuiu para o excesso de caracteres por linha, 47 na linha de cima e 42 na linha de baixo, extrapolando o número de 37 caracteres por linha sugerido por Díaz Cintas e Remael (2007). O excesso de caracteres exibido em 4.6s gerou uma legenda muito rápida, acima de 18cps, ou 180ppm. O problema de segmentação levou à desarmonia dos parâmetros de velocidade e número de caracteres por linha.

Uma sugestão para tal problema consiste em segmentar essa legenda em duas de duas linhas, como, por exemplo:

(14c) *These Pharaohs*

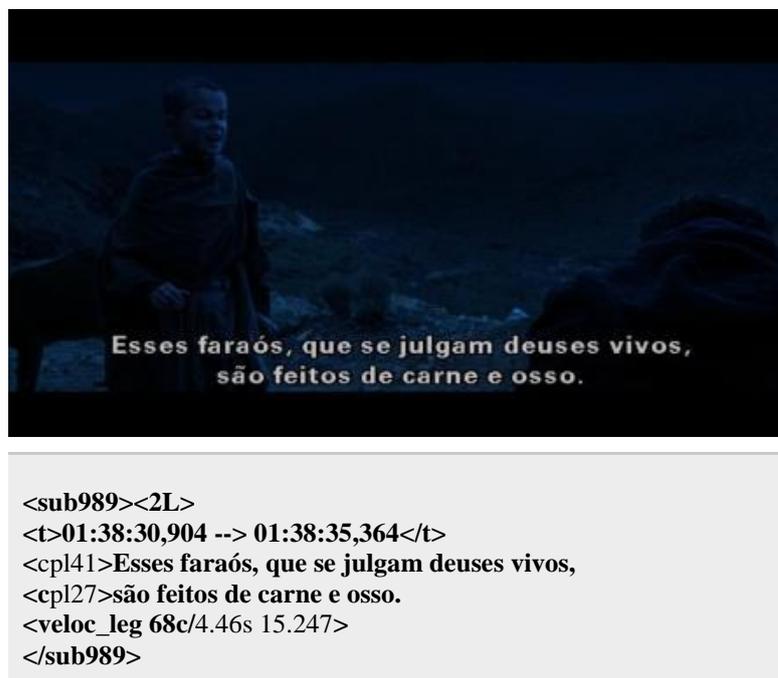
(32c) *who imagine they're living gods,*

(20c) *they're nothing more*

(22c) *than flesh and blood.*

Com essa proposta de ressegmentação, a legenda fica conforme os parâmetros preconizados pelos pesquisadores da recepção e do processamento de legendas, sem extrapolar o número de caracteres (DÍAZ CINTAS; REMAEL, 2007), sem baixar a velocidade (D'YDEWALLE *et al.*, 1987; DE LINDE; KAY, 1999) e sem causar problemas de segmentação linguística (VIEIRA, 2016; MONTEIRO, 2016). Essa foi uma estratégia que poderia ter sido adotada pelo legendista para segmentar o texto audiovisual de modo a manter os parâmetros em harmonia. Na tradução para o português, a legenda **Esses faraós, que se julgam deuses vivos, / são feitos de carne e osso.** também está fora dos parâmetros (figura 24).

Figura 24 – Problema de segmentação pela distribuição de texto nas legendas de duas linhas em português



Fonte: Elaborada pela autora. Fotograma legendado em português do filme *Exodus* (2015).

A primeira linha da legenda apresenta 41cpl. Para ressegmentá-la, a sugestão é retirar a palavra **vivos** da primeira linha, ficando da seguinte forma:

(cpl36) Esses faraós, que se julgam deuses,

(cpl26) são feitos de carne e osso.

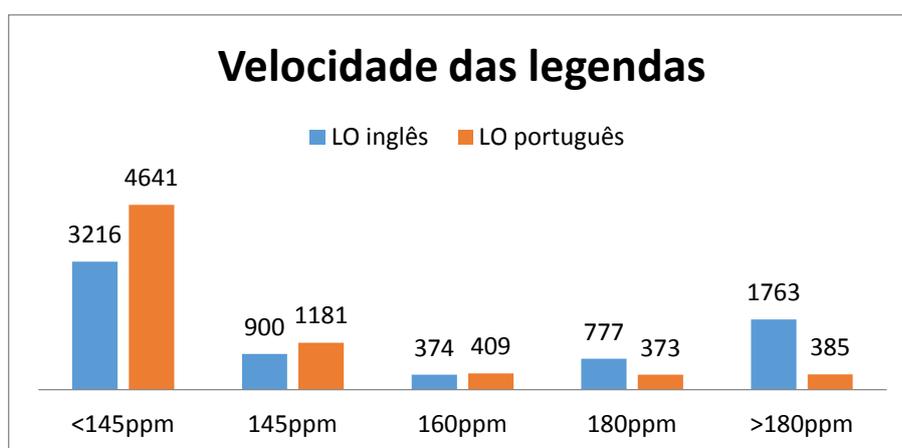
Essa estratégia poderia evitar a criação de legendas lentas, de uma linha, ou com problemas de segmentação linguística.

4.2.2 Velocidade das legendas em inglês e em português

A velocidade foi analisada nas legendas em inglês e português sob cinco parâmetros: legendas abaixo de 145ppm, legendas de 145ppm, legendas de 160ppm, legendas de 180ppm e legendas acima de 180ppm. Na literatura de processamento de legendas, foram testadas velocidades médias de 74ppm e 149ppm em De Linde e Kay (1999), e velocidades de 145ppm e 180ppm em Vieira (2016) e Monteiro (2016). A influência de outras velocidades, como as de 160ppm e acima de 180ppm, sobre o processamento de legendas não foi controlada em outros experimentos, até onde tenho conhecimento.

Em inglês, 3.216 (46%) legendas são de velocidade abaixo de 145ppm; 900 (13%) são de velocidade 145ppm; 374 (5%) são de velocidade 160ppm; 777 (11%) são de velocidade 180ppm; 1.763 (25%) são de velocidade acima de 180ppm. Em português, 4.641 (66%) são de velocidade abaixo de 145ppm; 1.181 (17%) são de velocidade 145ppm; 409 (6%) são de velocidade 160ppm; 373 (5%), de velocidade 180ppm; e 385 (6%), acima da velocidade de 180ppm. O gráfico 03 apresenta mais claramente uma comparação dessas velocidades em inglês e português.

Gráfico 03 – Velocidade das legendas em inglês e português



Fonte: Elaborado pela autora.

As primeiras colunas, da esquerda para a direita do gráfico, mostram que há uma predominância de legendas de velocidade abaixo de 145ppm em todo o *corpus*, apesar de terem número de ocorrências diferentes em inglês e português, conforme as colunas azul e vermelha, respectivamente. Já as de 145 e de 160ppm apresentam número de ocorrências semelhante em ambas as línguas. E as legendas de 180 e acima de 180ppm apresentam números de ocorrências bem distintos.

Em inglês, há claramente uma maioria de legendas de velocidade abaixo de 145ppm (46%), mas também uma expressiva ocorrência de legendas de velocidade acima de 180ppm (25%). Como nos exemplos a seguir:

Abaixo de 145ppm (inglês)	Acima de 180ppm (inglês)
<pre><sub528><1L> <t>00:34:11,800 --> 00:34:14,837</t> <cpl20> I can skip chapters. <veloc_leg 20c/3.037s 6.585> </sub528></pre>	<pre><sub737><2L> <t>01:10:15,919 --> 01:10:17,614</t> <cpl32>Is that what you are telling me? <cpl18>Are you saying no? <veloc_leg 50c/1.695s 29.499> </sub737></pre>
<pre><sub807><1L> <t>00:51:38,640 --> 00:51:40,949</t> <cpl20>I am a little tired. <veloc_leg 20c/2.309s 8.662> </sub807></pre>	<pre><sub790><2L> <t>00:41:23,961--> 00:41:25,861</t> <cpl27>He's gonna look at all this <cpl33>and if he thinks it's appropriate <veloc_leg 60c/1.9s 31.579> </sub790></pre>

As legendas em inglês estão etiquetadas de acordo com os parâmetros técnicos. À esquerda, os exemplos de legendas lentas, abaixo de 145ppm, e, à direita, os exemplos de legendas acima de 180ppm mostram problemas que, embora sejam diferentes, podem prejudicar a recepção e o processamento (D'YDEWALLE *et al.*, 1987; D'YDEWALLE; DE BRUYCKER, 2003; DE LINDE, KAY, 1999; VIEIRA, 2016; MONTEIRO, 2016), como, por exemplo, a legenda <sub528> **I can skip chapters** (tradução: Eu posso pular os capítulos), que é uma legenda lenta em inglês, assim como a legenda <sub807> **I am a little tired** (Tradução: estou um pouco cansada), abaixo de 145ppm; além disso, ambas são de uma linha. Os dois fatores podem interferir na recepção e no processamento de legendas (D'YDEWALLE *et al.*, 1987; D'YDEWALLE; DE BRUYCKER, 2003; DE LINDE; KAY, 1999; VIEIRA, 2016; MONTEIRO, 2016).

À direita, as legendas apresentam velocidade acima de 180ppm e duas linhas: <sub737> **Is that what you are telling me? / Are you saying no?** (Tradução: É isso que você está me dizendo? / Está me dizendo não?); e <sub790> **He's gonna look at all this / and if he thinks it's appropriate** (Ele está olhando pra tudo isso / e se ele achar que é apropriado). Este último parâmetro, duas linhas, favorece a recepção. Nenhuma dessas legendas em inglês excede o número de caracteres por linha (37). Sobre as velocidades acima de 180ppm, estas não foram controladas na pesquisa experimental da área, portanto não é possível concluir sobre o efeito desse padrão na recepção e no processamento de legendas. O fato de serem legendas de duas linhas que não apresentam problemas de segmentação linguística leva ao entendimento de que são melhores que as legendas lentas e de uma linha; porém não é possível ainda saber até que ponto a velocidade acima de 180ppm favorece a recepção e o processamento de legendas.

Em português, a situação é semelhante à do inglês, mas todos os problemas de parâmetros estão agravados, já que há uma grande maioria de legendas abaixo de 145ppm (66%), porém poucas legendas de velocidades mais altas, acima de 180ppm, como mostram os exemplos a seguir:

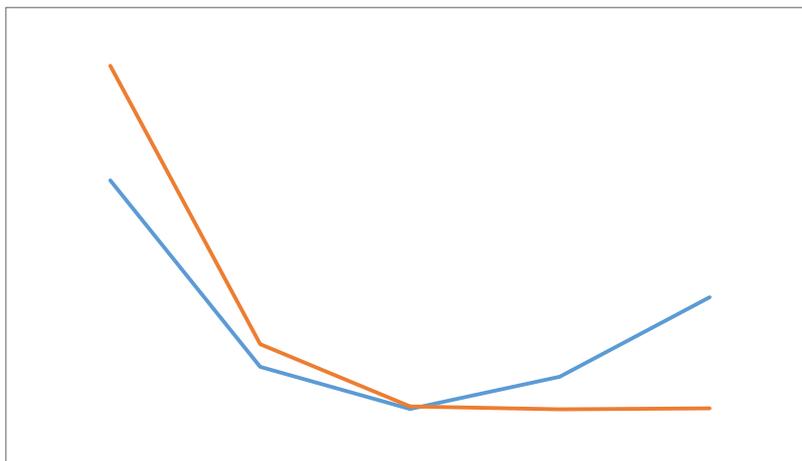
Abaixo de 145ppm (português)	Acima de 180ppm (português)
<pre><sub793><1L> <t>00:50:28,025 --> 00:50:30,323</t> <cpl15>Qual era mesmo? <veloc_leg 15c/2.298s 6.527> </sub793></pre>	<pre><sub736><1L> <t>00:47:13,997 --> 00:47:15,157</t> <cpl31>Esses babacas são muito chatos. <veloc_leg 31c/1.16s 26.724> </sub736></pre>
<pre><sub692><1L> <t>00:44:05,442 --> 00:44:09,742</t> <cpl30>A moça tinha uns peitões assim <veloc_leg 30c/4.3s 6.977> </sub692></pre>	<pre><sub959><2L> <t>01:00:23,453 --> 01:00:25,648</t> <cpl22>Isso é o que acontece. <cpl28>Ele vai atender ao telefone. <veloc_leg 50c/2.195s 22.779> </sub959></pre>

As legendas <793> e <692> são de velocidade abaixo de 145ppm e de apenas uma linha. Nenhuma delas ultrapassa o número de caracteres por linha, mas elas não atendem aos padrões de uma boa legenda, que sugeriu-se ser a rápida e bem segmentada (VIEIRA, 2016). As legendas em português à direita apresentam contextos diferentes, pois a <sub736> é rápida, acima de 180ppm, porém de uma linha, não atendendo os requisitos de Díaz Cintas e Remael (2007). Já a <sub 959> traz uma legenda rápida, segmentada em duas linhas, que não possui problemas de segmentação linguística e não ultrapassa o número de caracteres por linha. Está dentro dos padrões de uma legenda adequada, com todos os parâmetros (velocidade, número de linhas, número de caracteres e quebra de linha) em harmonia (D'YDEWALLE *et al.*, 1987; D'YDEWALLE; DE BRUYCKER, 2003; DE LINDE; KAY, 1999; VIEIRA, 2016).

Esses dados da velocidade são bastante expressivos, já que legendas abaixo de 145ppm, segundo De Linde e Kay (1999) e Vieira (2016), causaram releitura de legendas e fizeram aumentar o custo do processamento, tanto para surdos como para ouvintes. Para esta pesquisa, o alto percentual de legendas abaixo da velocidade de 145ppm pode comprometer a qualidade das legendas nas duas línguas. Entretanto, não há dados sobre o processamento de legendas de velocidade acima de 180ppm, abrindo uma lacuna para tal análise e impossibilitando concluir sobre a qualidade das legendas em inglês, que apresentam um quarto de legendas muito rápidas. Na representação gráfica da velocidade, à medida que a

velocidade vai aumentando, o número de legendas diminui como em linha decrescente, conforme representação no gráfico 04.

Gráfico 04 – Representação da velocidade em linhas



Fonte: Elaborado pela autora.

No gráfico 04, a linha azul representa a velocidade das legendas em inglês, e a linha vermelha, a velocidade das legendas em português. Enquanto, em português, a linha decresce e estabiliza, em inglês, a linha decresce e volta a crescer. O que chama atenção nessas linhas são os extremos: em inglês, há muitas legendas de velocidade abaixo de 145ppm e acima de 180ppm. Em português, há muitas legendas abaixo de 145ppm e poucas legendas rápidas (180ppm e acima de 180ppm). Em geral, a legendagem em inglês é composta de legendas lentas e rápidas, enquanto em português as legendas são predominantemente lentas. Na subseção seguinte, veremos por que esse desequilíbrio no que diz respeito à velocidade das legendas pode afetar a segmentação.

4.2.3 Problemas de segmentação e velocidade: LO em inglês × LO em português

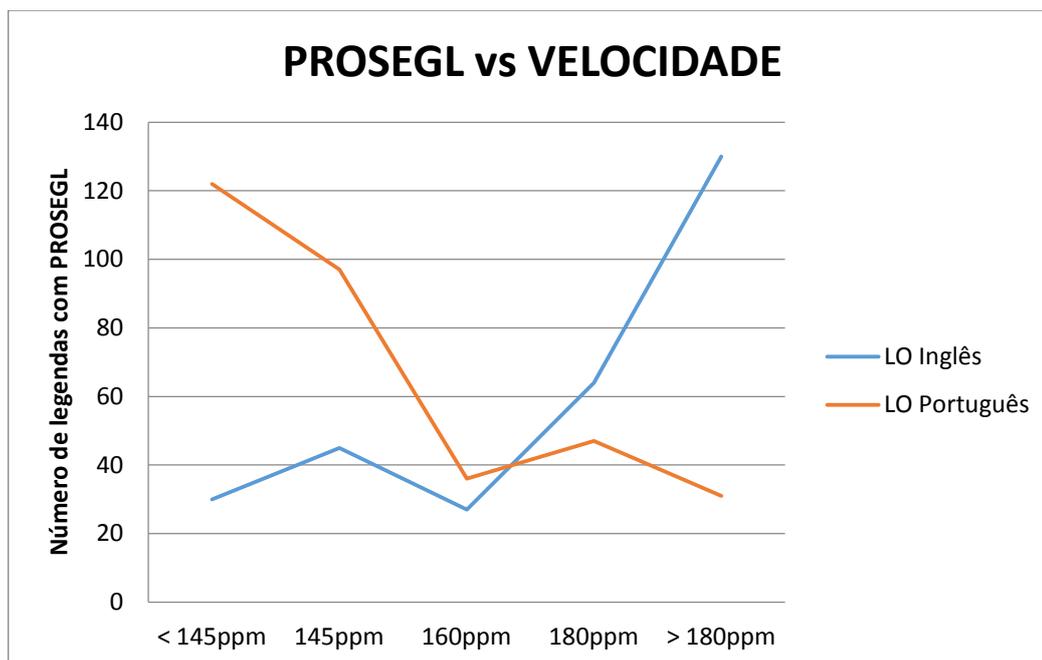
A interação entre participantes, segmentação e velocidade foi testada em Vieira (2016), com legendas de duas linhas, e em Monteiro (2016), com legendas de uma linha. Em geral, os resultados de ambas as pesquisas mostraram que os problemas de segmentação prejudicam a recepção de legendas por surdos e ouvintes. Nos dados estatísticos (ANOVA) de Vieira (2016), a interação de *grupo vs velocidade* ($F [3,60] = 1,5328, p > 0.05$), a interação *grupo vs segmentação* ($F [3,60] = 0,4085, p > 0.050$) e a interação *segmentação vs velocidade* não tiveram efeitos significativos, porém a interação *grupo vs segmentação vs velocidade*

($F [3,60] = 5,1844, p < 0.05$) teve efeito significativo, o que significa dizer que um dos grupos se comportou diferentemente em relação à segmentação e à velocidade quando tomadas em conjunto. A autora explica, em outras palavras, que quanto mais mal segmentadas forem as legendas, mais lento torna-se o processamento leitor.

Os resultados apontados em Vieira (2016) contribuem com a consolidação dos parâmetros de segmentação e velocidade e, para esta pesquisa, funcionam como um instrumento de avaliação, que pondera quais problemas de segmentação linguística podem afetar a qualidade das legendas e da recepção. Portanto, ao relacionar problemas de segmentação e velocidade nas legendas em inglês e português, verifico que, em inglês, a maioria (65%) ocorre em legendas mais rápidas, sendo 21% em legendas de 180ppm e 44% em legendas acima de 180ppm. Os demais problemas (35%) ocorrem em legendas mais lentas, 160ppm, 145ppm e abaixo de 145ppm.

Em português, o resultado é o oposto. Das 333 legendas com problemas de segmentação, a maioria (65,7%) ocorre em legendas mais lentas (145ppm e abaixo de 145ppm). De acordo com d'Ydewalle e De Bruycker (2003) e De Linde e Kay (1999), as legendas abaixo de 145ppm causam releituras e podem causar desconforto e irritação no espectador. Se essas legendas tiverem problemas de segmentação linguística (VIEIRA, 2016; MONTEIRO, 2016) e forem formadas por apenas uma linha (D'YDEWALLE *et al.*, 1987; DE LINDE; KAY, 1999; MONTEIRO, 2016), podem causar entraves na recepção e no processamento. As legendas de uma linha interferem na deflexão (movimento dos olhos que sai da legenda e vai para a imagem), aumentando a fixação nas legendas, o que pode diminuir a qualidade.

Diante desses dados, entendo que, em inglês, as legendas mal segmentadas são mais rápidas, e, em português, as legendas mal segmentadas são mais lentas. Apesar de apresentarem poucos PROSEGL (4,2% em inglês e 4,7% em português), a grande quantidade de legendas muito lentas e o excesso de distribuição do texto audiovisual em legendas de uma linha nos leva a concluir que as legendas em inglês e português estão mal segmentadas. O gráfico 05 apresenta uma comparação dos problemas de segmentação linguística, já que foram quantificados, e suas velocidades nos dois *corpora*.

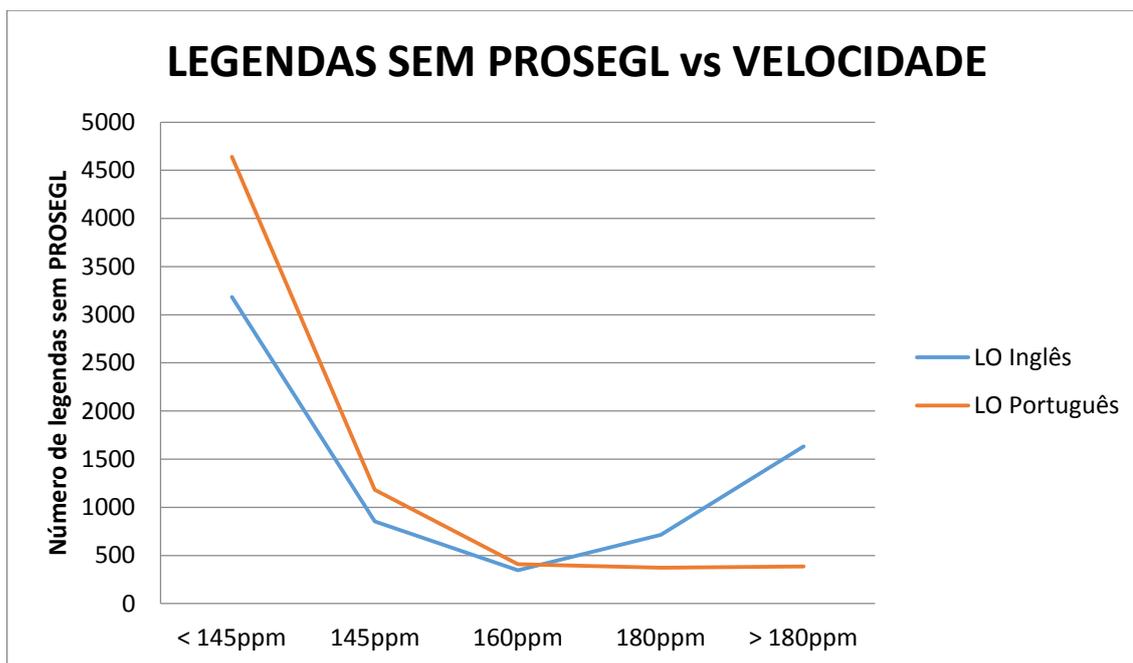
Gráfico 05 – Relação entre PROSEGL e velocidade no *corpus*

Fonte: Elaborado pela autora.

O gráfico 05 permite uma visualização em linhas da relação entre problemas de segmentação linguística e velocidade nas legendas em inglês e português. As duas linhas do gráfico, azul e vermelha, representam, em suas extensões da esquerda para a direita, a concentração dos problemas dado o aumento da velocidade. O eixo horizontal traz as cinco velocidades, e o eixo vertical traz a concentração de PROSEGL em cada uma delas. O início das linhas, na extremidade esquerda, mostra a concentração de problemas de segmentação linguística na velocidade abaixo de 145ppm. Dado o aumento da velocidade, 145, 160, 180 e acima de 180ppm, os problemas vão aumentando em inglês e vão diminuindo em português.

As demais legendas (95,8% em inglês e 95,3% em português) não apresentam PROSEGL, mas apresentam problemas de segmentação do texto audiovisual em linhas de legendas, pois são muito lentas e excessivamente segmentadas em 1 linha. O gráfico 06 apresenta a interação das legendas sem PROSEGL e as velocidades nos dois *corpora*.

Gráfico 06 – Relação entre legendas sem PROSEGL e velocidade



Fonte: Elaborado pela autora.

O gráfico 06 permite uma visualização em linhas da relação entre legendas sem problemas de segmentação linguística e velocidade. O início das linhas, à esquerda, mostra o alto número de legendas de velocidade abaixo de 145ppm tanto em inglês quanto em português. Porém, o final das linhas, à direita, mostra uma significativa concentração de legendas de alta velocidade apenas em inglês.

A média de problemas de segmentação linguística no *corpus* é baixa (4,5%), e sua maior interação se dá com velocidades abaixo de 145 e acima de 180ppm, velocidades que não foram controladas nem por Vieira (2016) nem por Monteiro (2016). No entanto, De Linde e Kay (1999) analisaram legendas com velocidade média de 74ppm e perceberam, tanto para surdos quanto para ouvintes, que a baixa velocidade causou releitura, prejudicando a recepção. Desse modo, a qualidade das legendas com velocidade acima de 180ppm continua desconhecida, possibilitando novo tema de estudo na área.

Considerando a discussão, é possível responder à primeira questão de pesquisa, que investiga como se caracterizam os problemas de segmentação e a velocidade nas legendas para ouvintes, de filmes em DVD, nos contextos de línguas inglesa e portuguesa: legendas de filmes em DVD são diferentes em inglês e português, pois, apesar de ambas serem predominantemente lentas e mal segmentadas, a LO em inglês apresenta também legendas muito rápidas. Em inglês, as legendas são rápidas e mal segmentadas, enquanto que, em português, são lentas e mal segmentadas.

4.3 COMPARAÇÃO DOS PROBLEMAS DE SEGMENTAÇÃO E VELOCIDADE NAS LEGENDAS EM PORTUGUÊS: LO × LSE

Esta subseção investiga quais as semelhanças e diferenças entre as legendas para ouvintes de filmes em DVD e as legendas para surdos de filmes em DVD e de TV. Apresento os problemas de segmentação e suas relações com a velocidade em legendas de língua portuguesa, tanto no contexto das legendas para ouvintes de filmes estrangeiros em DVD quanto no contexto das legendas para surdos de produtos audiovisuais brasileiros. O enfoque desta seção é na comparação de legendas para ouvintes e legendas para surdos em língua portuguesa. Para isso, comparo a LO dos cinco filmes em DVD analisados nessa pesquisa com a LSE de filmes em DVD e de programas de TV analisados em dois projetos do LATAV: Chaves e Amoreira (2015) e projeto CORSEL (ARAÚJO, 2015), respectivamente.

4.3.1 Problemas de segmentação e velocidade: LO × LSE

A LO de filmes em DVD teve baixa frequência de problemas de segmentação linguística em português (4,7%), os quais ocorreram, em sua maioria, nos sintagmas verbal, nominal e preposicional. A maioria dos problemas (65,7%) ocorre em legendas de velocidades mais baixas (<145ppm e 145ppm). Partindo dos dados da LO em DVD, faço uma comparação com os problemas de segmentação e a velocidade da LSE de filmes brasileiros em DVD e da LSE de programas de TV brasileiros. A comparação desses resultados mostra semelhanças e diferenças entre tais tipos de legendas.

O *corpus* de LO em português analisado nesta pesquisa corresponde às legendas de cinco filmes estrangeiros em DVD; o *corpus* de LSE de filmes analisado no relatório de pesquisa de Chaves e Amoreira (2015) corresponde às legendas de sete filmes brasileiros em DVD; o *corpus* de LSE de programas da Rede Globo de televisão analisado no projeto CORSEL (ARAÚJO, 2015) corresponde às legendas de vinte capítulos de quatro novelas¹⁹, sendo cinco de cada novela; legendas de três episódios de dois documentários²⁰; legendas de três episódios de três séries de TV²¹; legendas de um filme estrangeiro exibido pela Rede Globo²².

¹⁹ Novelas: *Amor Eterno Amor*, *Cheias de Charme*, *Chocolate com Pimenta*, *Avenida Brasil*.

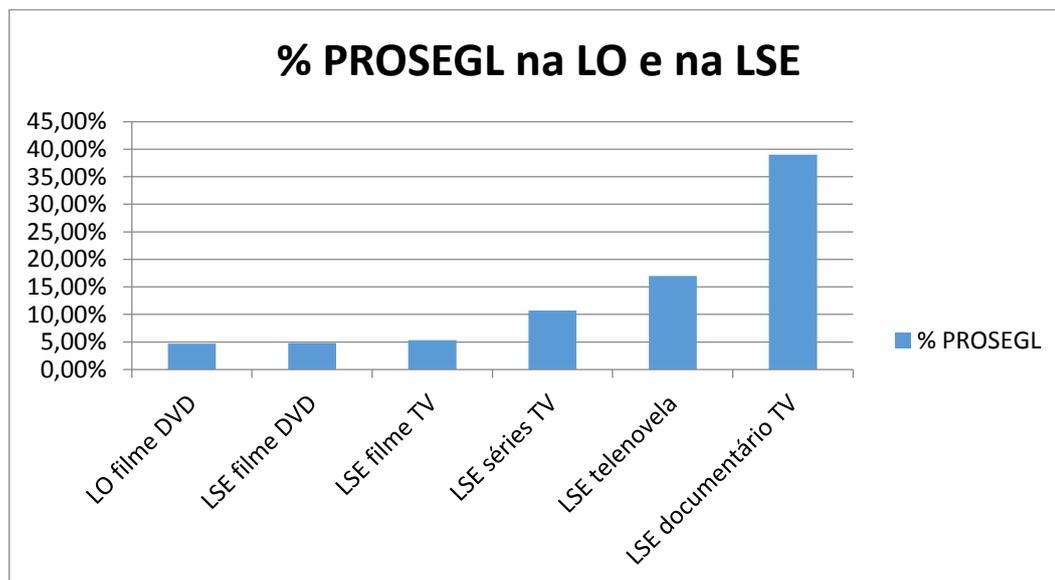
²⁰ Documentários: *Globo Repórter* e *Profissão Repórter*.

²¹ Seriados: *Tapas e Beijos*, *Pé na Cova*, *A Grande Família*.

²² Filme: *Virada Radical* (*Stick it*, 2006). Obs.: As legendas de filme não entraram para o relatório do CORSEL, mas foram analisadas no âmbito do projeto por Arraes (2015).

Com relação aos problemas de segmentação linguística, as frequências são semelhantes nas legendas de filmes (DVD e TV), mas são diferentes nas legendas de séries, telenovelas e documentários: LO de filmes em DVD (4,7%), LSE de filme em DVD (4,8%), LSE de filme de TV (5,3%), LSE de séries de TV (10,7%), LSE de telenovelas (17%) e LSE de documentários (39%). O gráfico 07 mostra uma comparação dos seis tipos de legendagem.

Gráfico 07 – Percentual de problemas de segmentação dos seis tipos de legenda

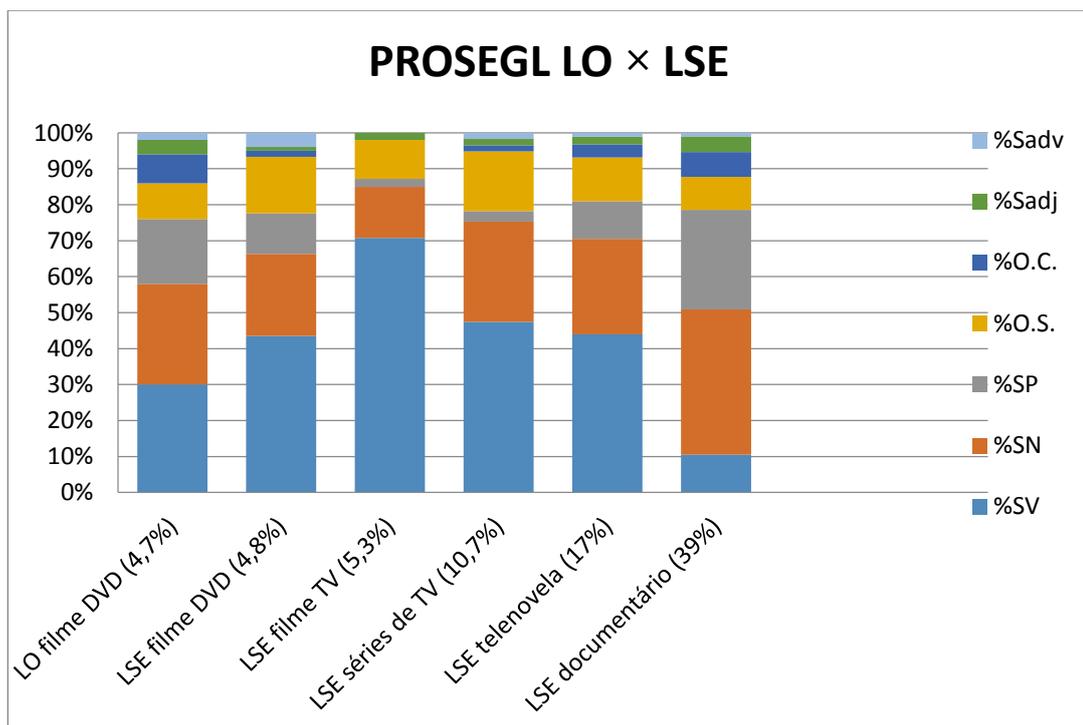


Fonte: Elaborado pela autora.

O gráfico 07 possibilita uma melhor visualização dos problemas de segmentação linguística dos seis tipos de legenda. Há semelhanças entre as legendas de filmes, tanto a LO de filmes em DVD quanto a LSE de filmes em DVD e a LSE de filme de TV; os três tipos têm baixa frequência de problemas de segmentação e maior incidência nos sintagmas verbal e nominal. A LSE de séries e a LSE de telenovelas diferem em percentual de problemas, pois apresentam percentuais maiores do que as legendas de filmes, e se assemelham em tipos de problema, nos sintagmas verbal e nominal. Já a LSE de documentários difere da LO e de todos os outros tipos de legendas tanto em percentual quanto em tipos de problema, pois apresentam 39% de problemas de segmentação linguística, e a maioria deles incide nos sintagmas nominal e preposicional.

Das sete categorias de problemas de segmentação linguística, os sintagmas verbal e nominal foram os tipos mais presentes tanto na LO quanto na LSE, exceto na LSE de documentários, cujas maiores incidências foram no nominal e no preposicional. O gráfico 08 mostra a distribuição das sete categorias de problemas nos seis tipos de legendas.

Gráfico 08 – Categorias de PROSEGL na LO e na LSE



Fonte: Elaborado pela autora.

O gráfico 08 apresenta predominância de cores azul e vermelha, que correspondem a sintagmas verbal e nominal, respectivamente. O mesmo padrão encontrado nas legendas para surdos e ensurdecidos de filmes, séries e telenovelas foi encontrado nas legendas para ouvintes, sendo, portanto, semelhantes. As legendas de filmes em DVD são feitas por profissionais da área, e as legendas de TV são feitas, em geral, por profissionais não legendistas. Pela lógica, a formação na área deve contribuir para o baixo percentual de problemas em legendas, tanto em LO quanto em LSE, devido a uma maior consciência do uso dos parâmetros da legendagem pelo legendista. No entanto, a LSE de filme da TV feita por profissionais não legendistas, assim como a LSE de séries e de telenovelas, que também apresentam baixa frequência de problemas, quando comparadas à LSE de documentários, sugerem que essa baixa frequência pode ser característica dos programas de ficção e não se repetiu nos documentários analisados no projeto CORSEL, que são programas de não ficção e apresentaram mais problemas nos sintagmas nominal e preposicional.

Pelos resultados do projeto CORSEL (ARAÚJO, 2015), os programas de ficção, como as novelas e os seriados, apresentam maioria de problemas de segmentação linguística na quebra entre os constituintes do sintagma verbal. Esse resultado se repete na LO e na LSE de filmes em DVD e na LSE de filmes de TV, sendo semelhantes, portanto, a LO e a LSE de programas de ficção de TV ou DVD e diferentes a LO e a LSE de documentários, no que diz

respeito aos problemas de segmentação. Isso responde em parte à segunda questão desta pesquisa sobre as semelhanças e diferenças dos problemas de segmentação e velocidade da LO e da LSE em português.

Sobre a velocidade das legendas, os resultados sobre a velocidade da LSE de TV com problemas de segmentação foram extraídos do relatório do CORSEL (ARAÚJO, 2015) e também das pesquisas correlatas feitas no âmbito do CORSEL, que apresentaram dados mais atualizados. A LO em inglês de filmes em DVD apresenta legendas de velocidades mais altas. A LO em português de filmes em DVD apresenta legendas de velocidades mais lentas, <145ppm e de 145ppm. A LSE de filmes brasileiros em DVD apresenta legendas mais lentas, de 145ppm (CHAVES; AMOREIRA, 2015). A LSE de filme de TV apresenta legendas de velocidade mais alta (ARRAES, 2015), a LSE de séries apresenta legendas de velocidade mais alta (ARAÚJO, 2015; FRANCO 2016), e a LSE de telenovelas apresenta legendas de velocidade mais alta (ARAÚJO, 2015; ASSIS, 2016). Os documentários apresentam legendas de velocidades mais altas (ARAÚJO, 2015; VIEIRA *et al.*, 2015; LEONARDO, 2016).

Comparando a velocidade das legendas com problema de segmentação, LO × LSE, é possível esclarecer que a LO em português de filmes em DVD se assemelha apenas à LSE em português de filmes em DVD, já que ambas apresentam legendas de velocidade baixa. Para os demais tipos de LSE, os problemas de segmentação apresentam legendas mais rápidas. Desse modo, no que diz respeito à velocidade, a LO se assemelha apenas à LSE de filmes em DVD e difere das demais LSE de TV (filme, séries, telenovelas, documentários). Esses dados revelam que as legendas de filmes em DVD são lentas e, por isso, podem causar releitura (DE LINDE; KAY, 1999; D'YDEWALLE *et al.*, 1987; VIEIRA, 2016) implicando na qualidade da legenda. Já as legendas da TV são rápidas, e esse fator isolado não tem efeito sobre o processamento de legendas, apenas quando em interação com a segmentação e o grupo (surdos ou ouvintes), sendo difícil considerar a qualidade com poucos elementos.

Ao comparar seis tipos de legendas (LO em português de filmes em DVD, LSE em português de filmes em DVD, LSE de filme de TV, LSE de séries de TV, LSE de telenovelas, LSE de documentários), os resultados mostraram que os problemas de segmentação da LO de filmes em DVD são semelhantes aos da LSE de programas de ficção (filmes de TV e DVD, séries e telenovelas) por ocorrerem mais nos sintagmas verbal e nominal, mas são diferentes dos problemas da LSE de documentários, que ocorrem nos sintagmas nominal e preposicional. A interação dos problemas de segmentação com a velocidade na LO de filmes em DVD é semelhante apenas à LSE de filmes em DVD, pois ambas concentram os problemas de segmentação em legendas de velocidades mais baixas

(145ppm), e é diferente da LSE de programas de TV, em que a concentração de problemas de segmentação está nas legendas de velocidades mais altas.

Em resumo e respondendo à segunda pergunta de pesquisa sobre as semelhanças e diferenças da LO e da LSE, a LO de filmes em DVD é semelhante à LSE de programas de ficção, no que diz respeito aos problemas de segmentação linguística, tanto em frequência de ocorrência baixa como em tipos de problema nos sintagmas verbal e nominal. Sobre a velocidade das legendas com problemas de segmentação, a LO de filmes em DVD se assemelha à LSE de filmes em DVD pelas velocidades mais baixas. Sobre as diferenças, a LO difere apenas da LSE de documentários no que diz respeito aos problemas de segmentação linguística, e, no que diz respeito à velocidade das legendas com problemas de segmentação, a LO difere da LSE de TV (filme, séries, telenovelas e documentários).

Respondendo à segunda pergunta de pesquisa sobre quais as semelhanças e diferenças entre as legendas para ouvintes de filmes em DVD em português e as legendas para surdos de filmes em DVD e de TV, chego à seguinte resposta: a LO em português apresenta problemas de segmentação, assim como a LSE de programas de ficção (filmes em DVD, filme de TV, séries de TV e telenovelas). A LO apresenta, em sua maioria, velocidade baixa, semelhante à LSE de filmes em DVD. A LO difere da LSE de documentários no que diz respeito aos problemas de segmentação linguística e difere da LSE de TV no que diz respeito à velocidade rápida que apresenta. As legendas de filmes em DVD em português (lentas mal segmentadas) são diferentes das legendas para surdos da TV, as quais são semelhantes à LO em inglês (rápidas mal segmentadas).

4.4 DESCRIÇÃO DOS PADRÕES DAS LEGENDAS PARA OUVINTES: INGLÊS × PORTUGUÊS

Esta seção investiga o padrão das legendas para ouvintes com base no padrão resultante das pesquisas sobre a recepção e sobre o processamento de legendas na LO de DVD em inglês e em português. A descrição foi feita com base nos padrões de segmentação e velocidade observados (D'YDEWALLE *et al.*, 1987; D'YDEWALLE; DE BRUYCKER, 2003; DE LINDE; KAY, 1999; VIEIRA, 2016; MONTEIRO, 2016) nos quais a legendagem pode ser: lenta mal segmentada (LMS), lenta bem segmentada (LBS), rápida mal segmentada (RMS) ou rápida bem segmentada (RBS), sendo este último padrão, RBS, aquele que apresentou melhor recepção e LMS, a pior recepção por surdos e ouvintes. Vieira (2016) mostra que problemas de segmentação atrapalham a recepção de LSE e que a boa

segmentação favorece a recepção tanto por surdos quanto por ouvintes. O enfoque desta seção é avaliar em qual desses padrões se encontram legendas em inglês e português.

4.4.1 Padrão da LO de DVD em inglês e português

Em geral, as legendas em inglês e português são muito lentas (português) ou muito rápidas (inglês), mal segmentadas, de apenas uma linha e com poucos problemas de segmentação linguística. O fato de as legendas serem feitas por profissionais leva a crer que as legendas de DVD atendem ao padrão de conforto resultante das pesquisas experimentais (D'YDEWALLE *et al.*, 1987; D'YDEWALLE; DE BRUYCKER, 2003; DE LINDE; KAY, 1999; VIEIRA, 2016; MONTEIRO, 2016), entretanto, os dados revelam problemas nas legendas de DVD. Levando em consideração a interação entre os problemas de segmentação linguística e velocidade nas legendas para ouvintes de filmes em DVD, apresento os seguintes dados sobre as legendas em inglês:

Quadro 06 – Características da LO em inglês de filmes em DVD

VELOCIDADE DAS LEGENDAS (INGLÊS)	COM PROSEGL (INGLÊS) 295 (4,2%)	SEM PROSEGL (INGLÊS) 6.735 (95,8%)
< 145 (46%)	30 (0,4%)	3186 (45,3%)
145 (13%)	45 (0,6%)	855 (12,1%)
160 (5%)	27 (0,3%)	347 (4,9%)
180 (11%)	64 (0,9%)	713 (10,1%)
>180 (25%)	130 (1,8%)	1633 (23,2%)

Fonte: Elaborado pela autora.

O quadro 06 mostra três colunas, a primeira com cinco padrões de velocidade, a segunda com as frequências de legendas com PROSEGL em cada velocidade, e a terceira com a frequência de legendas sem PROSEGL em cada velocidade. Em destaque, estão os dados das legendas com e sem PROSEGL nas velocidades de 145 e 180ppm. Embora Vieira (2016) e Monteiro (2016) tenham chamado de mal e bem segmentadas as legendas que apresentavam apenas problemas de segmentação linguística e tenhamos utilizado os padrões das autoras (LMS, LBS, RMS e RBS), nesta pesquisa fazemos distinção entre problemas de segmentação linguística e problemas de segmentação de texto audiovisual em legendas. Sobre a velocidade das legendas em inglês sem problemas de segmentação linguística, 3.186 (45,3%) estão abaixo de 145ppm, 855 (12,1%) são de 145ppm, 347 (4,9%) são de 160ppm, 713 (10,1%) são de velocidade de 180ppm e 1.633 (23,2%) estão acima de 180ppm.

Ao avaliar os padrões de segmentação e velocidade das legendas em inglês, é possível dizer que as legendas em inglês não atendem ao padrão de conforto resultante do projeto EXLEG, pois, apesar de bem segmentadas linguisticamente, são de velocidades abaixo da lenta, quando em inglês, acima da rápida. A recepção de legendas acima de 180ppm continua sendo inconclusiva, pois não tenho dados de base científica, advindos do rastreamento ocular ou da recepção de legendas, para avaliar qualidade de legendas com velocidade acima de 180ppm e sua interação com a segmentação, mas todas as discussões sobre o efeito da velocidade sobre o processamento sugerem que a velocidade de 180ppm não causou entrave no processamento de surdos e ouvintes, tanto na leitura de legendas de duas linhas (VIEIRA, 2016) quanto na leitura de legendas de uma linha (MONTEIRO, 2016). A recepção de legendas com velocidade abaixo de 145ppm foi verificada em De Linde e Kay (1999) como sendo problemática por ter causado releituras, tanto para surdos como para ouvintes, e, por isso, é possível avaliar que a LO em inglês está fora dos padrões de qualidade esperados para uma boa legendagem e recepção.

Em português, as legendas são bem segmentadas linguisticamente, mas a sua qualidade está mais comprometida que a das legendas em inglês, pois apresentam dois terços de legendas lentas. O quadro 07 mostra a frequência de cada um dos padrões nas legendas em português:

Quadro 07 – Características da LO em português de filmes em DVD

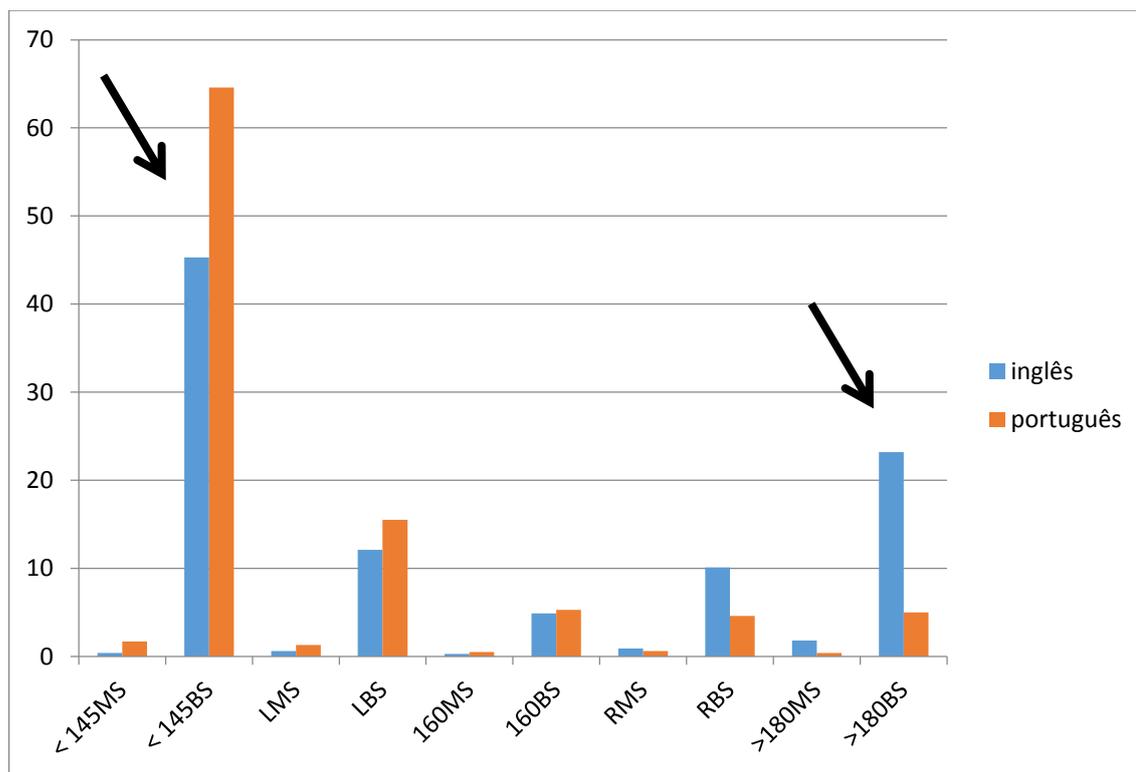
VELOCIDADE DAS LEGENDAS (PORTUGUÊS)	COM PROSEGL (PORTUGUÊS) 333 (4,7%)	SEM PROSEGL (PORTUGUÊS) 6.656 (95,3%)
< 145 (66%)	122 (1,7%)	4519 (64,6%)
145 (17%)	97 (1,3%)	1084 (15,5%)
160 (6%)	36 (0,5%)	373 (5,3%)
180 (5%)	47 (0,6%)	326 (4,6%)
>180 (6%)	31 (0,4%)	354 (5%)

Fonte: Elaborado pela autora.

Sobre a velocidade das legendas em português, 4.641 (66%) estão abaixo de 145ppm, 1.181 (17%) são de 145ppm, 409 (6%) são de 160ppm, 373 (5%) são de 180ppm e 385 (6%) estão acima de 180ppm.

Tanto em inglês quanto em português, os padrões de segmentação e velocidade predominante <145BS são avaliados como aqueles que podem causar entraves à recepção, por proporcionar releituras. Os padrões emergentes da LO em inglês e da LO em português estão em destaque, com setas indicativas, no gráfico 09.

Gráfico 09 – Interação segmentação e velocidade nas legendas em inglês e português



Fonte: Elaborado pela autora.

No gráfico 09, as maiores colunas à esquerda, no padrão <145BS, indicadas por seta, mostram o alto percentual de legendas com velocidade abaixo de 145ppm em inglês (46%) e português (66%). As colunas à direita, no padrão >180BS, mostram que esse padrão é expressivo apenas em inglês (23,2%).

Para ouvintes, conforme dados de Vieira (2016), LSE bem segmentadas em ambas as velocidades (145ppm e 180ppm) são bem recebidas, mas padrões diferentes, como os encontrados nesta pesquisa nas legendas para ouvintes em inglês e português (<145BS e >180BS), como dito anteriormente, não foram analisados, não sendo possível conjecturar sobre a recepção desses padrões sem investigação prévia. As legendas averiguadas por Vieira (2016) foram LSE de documentários, que, comparadas à LO de filmes, demonstraram ser diferentes em problemas de segmentação e em velocidade. Legendas de filmes em DVD, feitas por profissionais, com rigor de mercado, parecem ter sua qualidade questionada após os resultados desta pesquisa. Contudo, acredito que novas pesquisas possam discutir a recepção desse padrão das legendas de filmes em DVD, pois ainda que converse com os resultados de Perego *et al.* (2010) sobre o processamento de legendas de filmes, os resultados dos autores são questionáveis quanto ao desenho metodológico (VIEIRA, 2016). Perego *et al.* (2010), diferentemente de Vieira (2016) e Monteiro (2016), encontraram que os problemas de

segmentação não exercem influência sobre o processamento de legendas, mas, para esta pesquisa, esses dados são insuficientes para embasar a avaliação de legendas, uma vez que os autores desconsideraram a totalidade dos problemas de segmentação que uma legenda pode proporcionar em todos os seus sintagmas e orações, ao testar apenas problemas no sintagma nominal, presentes em um número reduzido de 28 legendas de um total de 171.

Os padrões encontrados na LO de filmes em DVD não condizem com o padrão de legendas rápidas, bem segmentadas linguisticamente e de maioria de duas linhas. Em geral, esse padrão existe na LO em inglês e português, mas em baixa frequência. O padrão predominante em inglês é de legendas mal segmentadas e acima de 180ppm. O padrão predominante em português é de legendas mal segmentadas e abaixo de 145ppm.

Respondendo à terceira pergunta de pesquisa: com base no padrão de conforto para ouvintes resultante das pesquisas sobre a recepção e o processamento de legendas (D'YDEWALLE *et al.*, 1987; D'YDEWALLE; DE BRUYCKER, 2003; DE LINDE; KAY, 1999; VIEIRA, 2016; MONTEIRO, 2016), representado por legendas rápidas e bem segmentadas de duas linhas e sem problemas de segmentação linguística, a LO de filmes em DVD em inglês e português atendem a esse padrão? Tanto em inglês quanto em português, a LO de DVD é mal segmentada e de uma linha, sendo a legenda em inglês expressivamente rápida e em português lenta, portanto não atende ao padrão das pesquisas experimentais. As legendas para ouvintes em inglês e em português não atendem ao padrão de legendas rápidas bem segmentadas (RBS) e de duas linhas.

4.5 LO DE FILMES EM DVD EM INGLÊS E PORTUGUÊS: SÍNTESE DOS RESULTADOS

Nesta subseção, faço uma síntese dos resultados discutindo sobre os achados desta pesquisa em três contextos distintos, correspondentes aos três objetivos específicos: LO (inglês e português) de filmes em DVD; LSE em português de filmes brasileiros em DVD e de programas da TV brasileira; e recepção e processamento da LSE (português) por surdos e ouvintes. A discussão desses resultados culmina com o objetivo geral desta pesquisa, que avaliou em que medidas as legendas para ouvintes, de filmes em DVD, nos contextos de língua inglesa e portuguesa, dialogam com os padrões de segmentação e velocidade encontrados na LSE de filmes em DVD e de programas de TV, e com a recepção e processamento de LSE por surdos e ouvintes, conforme resultados das pesquisas sobre a recepção e o processamento de legendas (D'YDEWALLE *et al.*, 1987; DE LINDE; KAY, 1999; D'YDEWALLE; DE BRUYCKER, 2003; VIEIRA, 2016; MONTEIRO, 2016).

Legendas para ouvintes de filmes em DVD em inglês e português são diferentes: lentas mal segmentadas em português e rápidas mal segmentadas em inglês. A desarmonia dos parâmetros de velocidade e número de linhas torna as legendas mal segmentadas, pois, quando lentas (abaixo de 145ppm), prejudicam recepção e processamento (D'YDEWALLE *et al.*, 1987; DE LINDE; KAY, 1999; VIEIRA, 2016), causando percepção de desconforto e gerando dados de excesso de releitura e excesso de deflexão. Quando segmentadas em uma linha, também prejudicam recepção e processamento (D'YDEWALLE; DE BRUYCKER, 2003; MONTEIRO, 2016), causando percepção de desconforto e a sensação de que a legenda está rápida e gerando dados de excesso de deflexão e excesso de fixação. Além dos problemas de segmentação gerados pela desarmonia entre velocidade e número de linhas, ambas apresentaram problemas de segmentação linguística – em inglês, 4,2%, em português, 4,7%. Embora o percentual de problemas de segmentação linguística seja baixo, sua interação com velocidades muito lentas ou muito rápidas e com o alto número de legendas de uma linha pode prejudicar a qualidade das legendas.

Essa desarmonia está muito mais presente nas legendas em português do que nas legendas em inglês, já que apresentam todos os problemas em maior quantidade. Dessa forma, ambas têm a qualidade comprometida, mas as legendas em português são ainda mais problemáticas do que as legendas em inglês. Tal resultado surpreende, pois além de serem feitas por legendistas profissionais, visualmente as legendas de DVD parecem atender aos parâmetros preconizados pelas pesquisas experimentais (D'YDEWALLE *et al.*, 1987; DE LINDE; KAY, 1999; D'YDEWALLE; DE BRUYCKER, 2003; VIEIRA, 2016; MONTEIRO, 2016).

Enxergar essa desarmonia só foi possível pela metodologia baseada em *corpus*, que, ao processar grandes volumes de textos, revela dados e padrões impossíveis de serem vistos numa análise manual, como, por exemplo, que as legendas em inglês apresentam 45% de legendas com velocidade abaixo de 145ppm, 55% de legendas de uma linha e 4,2% de problemas de segmentação linguística, enquanto as legendas em português apresentam 66% de legendas com velocidade abaixo de 145ppm, 62% de legendas de uma linha e 4,7% de problemas de segmentação linguística.

Embora as legendas de TV sejam rápidas (CHAVES, 2009; ARAÚJO; ASSIS, 2014; ARAÚJO, 2015) e isoladamente esse parâmetro não seja prejudicial ao processamento (VIEIRA, 2016), quando em interação com os demais parâmetros problemáticos, promove a desarmonia, prejudicando a qualidade. Novamente, a linguística de *corpus*, por meio da

metodologia baseada em *corpus*, foi fundamental para revelar esses dados, uma vez que essa comparação não seria viável por análise manual.

Legendas de DVD são rápidas e mal segmentadas em inglês, e lentas e mal segmentadas em português. Desse modo, ambas as legendas têm sua qualidade comprometida pela desarmonia dos parâmetros, e as legendas em português são ainda mais problemáticas que as legendas em inglês. A desarmonia expressa nas legendas de DVD é corroborada pelas pesquisas sobre a recepção e o processamento de legendas usadas como referência nesta tese (D'YDEWALLE *et al.*, 1987; D'YDEWALLE; DE BRUYCKER, 2003; DE LINDE; KAY, 1999; PEREGO *et al.*, 2010; VIEIRA, 2016; MONTEIRO, 2016).

O resultado dessa avaliação sugere que as legendas para ouvintes, de filmes em DVD, em inglês e português não condizem com os parâmetros preconizados nas pesquisas experimentais na área (D'YDEWALLE *et al.*, 1987; D'YDEWALLE; DE BRUYCKER, 2003; DE LINDE; KAY, 1999; PEREGO *et al.*, 2010; VIEIRA, 2016; MONTEIRO, 2016).

5 CONSIDERAÇÕES FINAIS

Esta tese teve como objetivo avaliar em que medidas as legendas para ouvintes, de filmes em DVD, nos contextos de língua inglesa e de língua portuguesa, dialogam com os padrões de segmentação e velocidade encontrados nas pesquisas descritivas e experimentais da legendagem. De um lado, as pesquisas descritivas e baseadas em *corpus* da legendagem (ARAÚJO, 2015; ASSIS, 2013; GABRIEL, 2013; NASCIMENTO, 2013; AGUIAR, 2013; CHAVES; ARAÚJO, 2014; ARAÚJO; ASSIS, 2014; VIEIRA *et al.*, 2015; ARRAES, 2015; FRANCO, 2016; LEONARDO, 2016; ASSIS, 2016;) revelaram os padrões de vários tipos de legendagem, fornecendo informações sobre seus parâmetros, como: legendas de filmes em DVD são lentas e mal segmentadas, legendas de TV são rápidas e mal segmentadas, legendas de programas de ficção (filmes, séries e telenovelas) apresentam problemas de segmentação linguística nos sintagmas verbal e nominal, e as legendas de programas de não ficção (documentários) apresentam problemas de segmentação nos sintagmas nominal e preposicional. Esses padrões dizem respeito às características das legendas e não sugerem por si sós a influência que exercem sobre a recepção. Por outro lado, as pesquisas experimentais sobre a recepção e o processamento de legendas (D'YDEWALLE *et al.*, 1987; D'YDEWALLE; DE BRUYCKER, 2003; DE LINDE; KAY, 1999; PEREGO *et al.*, 2010; VIEIRA, 2016; MONTEIRO, 2016) são capazes de fornecer dados sobre essa influência e sugerir, por exemplo, que legendas mal segmentadas, lentas ou de uma linha causam entraves no processamento.

Para avaliar os padrões das legendas para ouvintes de DVD, além do suporte teórico da legendagem e das pesquisas descritivas e experimentais da área recém mencionado, foi necessário buscar suporte metodológico na abordagem baseada em *corpus* oferecida pela Linguística de *Corpus*, que vem demonstrando ser essencial para a discussão teórica da legendagem, na medida em que revela padrões e relaciona-os aos diferentes tipos de legendas analisadas, contribuindo para elucidar os parâmetros da legendagem. Como, por exemplo, constatar que as legendas de DVD do Brasil são de velocidade lenta, que as legendas de TV são de velocidade alta, que as legendas de DVD são, em sua maioria, de uma linha, que as legendas de TV são de 2 e de 3 linhas. Constatar que as legendas de programas de ficção (filmes, séries e telenovelas) têm menos problemas de segmentação linguística e que estes estão concentrados nos sintagmas verbal e nominal, diferentemente dos programas de não ficção, como os documentários, que têm muitos problemas de segmentação linguística, em sua maioria nos sintagmas nominal e preposicional. Dentre outros esclarecimentos que não seriam viáveis por uma análise quantitativa manual.

A Linguística de *Corpus*, enquanto metodologia descritiva, ajudou a revelar padrões e constatar características marcantes da legendagem em si, elucidando suas questões teóricas a partir de procedimentos como os utilizados nesta pesquisa: extração das legendas de DVD, etiquetagem dos parâmetros da legendagem a partir do quadro de etiquetas do projeto CORSEL (ARAÚJO, 2015), processamento e análise com auxílio das ferramentas do programa de análise linguística *WordSmith Tools 7.0*.

A literatura sobre a recepção de legendas ajudou a discutir as questões cognitivas sobre a percepção de legendas, envolvendo análise da percepção e das sensações ao assistir produtos audiovisuais legendados, obtida por relatos (protocolos verbais ou escritos) de espectadores. A literatura sobre o processamento de legendas a partir de rastreamento ocular ajudou a discutir sobre questões fisiológicas do comportamento ocular, envolvendo análise dos efeitos de fatores das legendas, experimentalmente controlados ou não, sobre o processamento de legendas, obtidos por medidas quantitativas da movimentação ocular geradas pelos registros do rastreador. Tais vertentes, tomadas como base nesta pesquisa, sustentaram as conclusões desta tese de que as legendas de filmes em DVD são muito problemáticas.

Tal resultado acaba frustrando a expectativa inicial de que legendas de DVD atendem aos padrões de qualidade de uma legenda, uma vez que são feitas por profissionais legendistas e tendo em vista que o DVD é um tipo de mídia que, quando lançado por produtoras e/ou distribuidoras, sob as exigências de mercado, imprime credibilidade. É fácil acreditar que as legendas de DVD apresentam qualidade superior às demais, e avaliar os padrões de algo que já é considerado de qualidade pode gerar perspectivas de transformações de realidades contrárias àquelas que estivessem relacionadas a um problema de fato, mas os resultados desta tese lançaram luz para um novo dado sobre esse tipo de legenda, o que tem reflexo sobre o exercício da prática profissional e a discussão teórica da área.

Para avaliar os padrões das legendas para ouvintes em DVD com base nas pesquisas descritivas e experimentais da legendagem, foram comparadas diferentes legendas, e as características resultantes de cada uma foram avaliadas com base no padrão de recepção e processamento das pesquisas exploratórias experimentais. Dessa forma, parti de três questionamentos que correspondem aos objetivos específicos desta tese: 1 - Como se caracterizam os problemas de segmentação e a velocidade nas legendas para ouvintes, em filmes em DVD, nos contextos de línguas inglesa e portuguesa? 2 - Quais as semelhanças e diferenças entre as legendas para ouvintes em filmes em DVD em português e as legendas para surdos de filmes em DVD e programas de TV em língua portuguesa? 3 - Qual o padrão das legendas para ouvintes em

DVD em inglês e português com base no padrão resultante das pesquisas sobre a recepção e o processamento de legendas?

A comparação entre legendas de DVD em inglês e português mostrou que, mesmo em línguas diferentes, as legendas são semelhantes por serem lentas e mal segmentadas, mas também rápidas em inglês. As duas legendagens possuem poucos problemas de segmentação linguística e apresentaram um grande percentual de legendas de uma linha. Embora ambas tenham esse padrão, as legendas em português são mais problemáticas que as legendas em inglês, pois apresentam todos os problemas em maior quantidade. Além disso, as legendas em inglês apresentam mais legendas rápidas do que as que em português. Esse percentual é muito representativo, mas não há dados sobre a recepção e o processamento de legendas de velocidade acima de 180ppm que possibilitem concluir sobre os efeitos das legendas acima de 180ppm. Com isso, abre-se uma lacuna para novas investigações sobre a recepção e o processamento de legendas muito rápidas, acima de 180ppm.

A comparação entre legendas para ouvintes (DVD) e legendas para surdos (DVD e TV) mostra que há semelhança entre as legendas de DVD (LO e LSE): ambas são mal segmentadas, possuem muitas legendas de uma linha e percentuais e tipos de problemas de segmentação linguística semelhantes. Os padrões encontrados na LO de DVD podem causar algum tipo de desconforto no espectador, se levarmos em conta que nenhuma delas atendeu ao padrão de rápida e bem segmentada encontrado por Vieira (2016).

Segundo os resultados da pesquisa, o que diminui a qualidade das legendas é a desarmonia dos parâmetros, que pode ser ocasionada por problemas de segmentação linguística e de distribuição do texto audiovisual, velocidade lenta (abaixo de 145ppm) e legendas de uma linha. E o que contribui para a qualidade das legendas é a harmonia dos parâmetros, dada pela boa segmentação linguística, de modo a preservar os sintagmas, velocidade de 180ppm e maioria de legendas de 2 linhas.

Tais resultados convergem para o argumento de que a harmonia dos parâmetros é essencial para a qualidade das legendas. O fato de as legendas de filmes em DVD serem feitas por legendistas profissionais gerou expectativas quanto à qualidade, porque, em geral, os profissionais e as empresas de legendagem traduzem com base em parâmetros seguidos por profissionais da área e aceitos por espectadores. Porém, os resultados questionaram o consenso em torno da qualidade das legendas de DVD, sugerindo que são problemáticas. Portanto, esta tese abre espaço para novos estudos em busca de avaliar recepção e processamento das legendas de DVD.

REFERÊNCIAS

- ALVARENGA, L. Subtíler: legendador ou legendista. In: CONGRESSO IBERO-AMERICANO DE TRADUÇÃO E INTERPRETAÇÃO (CIATI): TRADUÇÃO, INTEPRETAÇÃO E CULTURA NA ERA DA GLOBALIZAÇÃO, 1, 1998, São Paulo. **Anais...** São Paulo: UNIBERO, 1998. p. 214-216.
- ARAÚJO, V. L. S. Closed subtitling in Brazil. In: **Topics in audiovisual translation**. Amsterdam: John Benjamins Publishing Company, 2004. v. 1. p. 199-212.
- _____. A legendagem para surdos no Brasil. In: LIMA, P. L. C.; ARAÚJO, A. D. (Org.). **Questões de Linguística Aplicada**: miscelânea. Fortaleza: EdUECE, 2005. p. 163-188.
- _____. Subtitling for the deaf and hard-of-hearing in Brazil. In: ORERO, P.; REMAEL, A. (Org.). **Media for All: Subtitling for the Deaf, Audio Description and Sign Language**. Kenilworth: Nova Jersey, EUA: Rodopi, 2007. v. 30. p. 99-107.
- _____. Por um modelo de legendagem para surdos no Brasil. **Tradução e Comunicação – Revista Brasileira de Tradutores**, São Paulo, n. 17, p. 59-76, 2008.
- _____. In Search of SDH Parameters for Brazilian Party Political Broadcasts. **The Sign Language Translator and Interpreter**, Manchester, v. 3, n. 2, p. 157-167, 2009.
- _____. **Legendagem para surdos**: em busca de um modelo para o Brasil. Relatório Técnico n. 306948/2008-7. Fortaleza: CNPq. Fev./2012.
- _____. **A segmentação na legendagem para surdos e ensurdecidos (LSE)**: um estudo baseado em *corpus* (CORSEL). Relatório Técnico n. 306441/2011-0. Fortaleza: CNPq. Fev./2015.
- 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**, Uberlândia, v. 30, n. 2, p. 156-184, 2014. Disponível em: <<http://www.seer.ufu.br/index.php/letraseletras>>. Acesso em: 30 dez. 2014.
- ARAÚJO, V. L. S.; CHAVES, É. G. Guia de legendagem para surdos e ensurdecidos. In: ARAÚJO, V. L. S.; ALVES, S. F.; MAUCH, C.; NEVES, S. B. **Guia para Produções Audiovisuais Acessíveis**. Brasília: Ministério da Cultura/Secretaria do Audiovisual, 2016. p. 42-81. Disponível em: <<https://grupoleaduece.blogspot.com.br/p/guia-para-producoes-audiovisuais.html>>. Acesso em: 07 fev. 2017.
- ARAÚJO, V. L. S.; NASCIMENTO, A. K. P. Investigando parâmetros de legendas para Surdos e Ensurdecidos no Brasil. **Tradução em Revista**, v. 2, p. 1-18, 2011. Disponível em: <<http://www.maxwell.lambda.ele.puc-rio.br/18862/18862>>. Acesso em: 01 fev. 2012.
- ARRAES, D. A. **Legendagem para surdos e ensurdecidos**: análise baseada em *corpus* da segmentação linguística do filme “Virada Radical”. 2015. 70 f. Monografia (Bacharelado em Letras Inglês) – Universidade Estadual do Ceará, Fortaleza-CE, 2015.

ASSIS, I. A. P. **A segmentação na LSE de “Amor Eterno Amor”**: uma análise baseada em *corpus*. 2013. 71 f. Monografia (Bacharelado em Letras Inglês) – Universidade Estadual do Ceará, Fortaleza-CE, 2013.

_____. **Legendagem para surdos e ensurdecidos (LSE)**: análise baseada em *corpus* da segmentação linguística em “Amor Eterno Amor”. 2016. 108 f. Dissertação (Mestrado) – Programa de Pós-Graduação em Linguística Aplicada, Universidade Estadual do Ceará, Fortaleza-CE, 2016.

BAKER, M. *Corpus Linguistics and Translation Studies: Implications and Applications*. In: BAKER, M.; FRANCIS, G.; TOGNINI-BONELLI, E. (Org.). **Text and technology**: In honour of John Sinclair. Philadelphia: John Benjamins, 1993. p. 233-250.

_____. *Corpora in translation studies: an overview and some suggestions for future research*. **Target**, v. 7, n. 3, p. 223-243, 1995.

_____. *Corpus-based translation studies: the challenges that lie ahead*. In: SOMERS, H. (Ed.). **Terminology, LSP and translation**. Amsterdã: John Benjamins, 1996. p. 175-187.

BARTOLL, E. The surtitling in Catalan of classic foreign theatre plays. In: INCALTELLA, L.; BISCIO, L.; MHAINNIN, M. A. N. **Audiovisual Translation: Subtitles and Subtitling**. Oxford: Peter Lang, 2011. p. 87-108.

BRASIL. Lei nº 10.098, de 19 de dezembro de 2000. Estabelece normas gerais e critérios básicos para a promoção da acessibilidade das pessoas portadoras de deficiência ou com mobilidade reduzida, e dá outras providências. **D.O.U.**, 20 dez. 2000. Disponível em: <<http://www3.dataprev.gov.br/SISLEX/paginas/42/2000/10098.htm>>. Acesso em: 10 jul. 2011.

CASTILHO, A. **Nova Gramática do Português Brasileiro**. São Paulo: Contexto, 2012.

CHAVES, É. G. **Legenda para Surdos no Brasil**: uma análise baseada em *corpus*. 2009. 52 f. Monografia (Bacharelado em Letras Inglês) – Universidade Estadual do Ceará, Fortaleza-CE, 2009.

_____. **Legendagem para Surdos e Ensurdidos**: Um estudo baseado em *corpus* da segmentação nas legendas de filmes brasileiros em DVD. 2012. 130 f. Dissertação (Mestrado) – Universidade Estadual do Ceará, Fortaleza-CE, 2012.

CHAVES, É. G.; AMOREIRA, S. N. **Corpus de Legendagem para Surdos e Ensurdidos (LSE)**: Um estudo da segmentação nas legendas de filmes brasileiros em DVD. Relatório de pesquisa de Iniciação Científica. Fortaleza: IC-UECE. Ago/2015.

CHAVES, É. G.; ARAÚJO, V. L. S. Segmentation Tags: A Proposal for the Analysis of Subtitles. In: ALUISIO, S. M.; TAGNIN, Stella E. O. **New Language Technologies and Linguistic Research: A Two-Way Road**. Newcastle upon Tyne, UK: Cambridge Scholars Publishing, 2014. Disponível em: <<http://lib.myilibrary.com/Open.aspx?id=586208>>. Acesso em: 20 dez. 2014.

DE LINDE, Z.; KAY, N. **The semiotics of subtitling**. Manchester: St. Jerome Publishing, 1999.

DÍAZ CINTAS, J. Audiovisual Translation Today: A question of accessibility for all. **Translating Today**, v. 4, p. 3-5, 2005.

DÍAZ CINTAS, J.; ANDERMAN, G. **New Trends in Audiovisual Translation**. Bristol, UK: Multilingual Matters, 2009.

DÍAZ CINTAS, J.; REMAEL, A. **Audiovisual Translation: Subtitling**. Manchester, UK: St. Jerome Publishing, 2007.

DÍAZ CINTAS, J.; SANCHÉZ, P. M. Fansubs: Audiovisual Translation in an Amateur Environment. **The Journal of specialized translation**, v. 6, p. 37-52, 2006.

DINIZ, N. S. **A segmentação na legendagem para surdos e ensurdecidos**: um estudo baseado em *corpus*. 2012. 150 f. Dissertação (Mestrado) – Universidade Federal de Minas Gerais, Belo Horizonte-MG, 2012.

D'YDEWALLE, G.; DE BRUYCKER, W. Reading native and foreign language television subtitles in children and adults. In: HYONA, J.; RADACH, R.; DEUBEL, H. (Org.). **The mind's eye**: cognitive and applied aspects of eye movements. Oxford, UK: Research Lirias, 2003. p. 671-684.

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. (Ed.). **Eye Movements**: from physiology to cognition. Amsterdam: Elsevier Science Publishers, 1987. p. 313-321.

FRANCO, R. P. **Problemas de segmentação linguística na legendagem para surdos e ensurdecidos (lse) de séries da TV brasileira**: um estudo baseado em *corpus*. 2016. 63 f. Monografia. (Bacharelado em Letras Inglês) – Universidade Estadual do Ceará, Fortaleza-CE, 2016.

FRANCO, E.; ARAÚJO, V. L. S. Reading Television: Checking deaf people's reactions to Closed Subtitling in Fortaleza, Brazil. **The Translator**, v. 9, n. 2, p. 249-267, 2003.

_____. Questões terminológicas e conceituais. **Tradução em Revista**, v. 2, p. 1-18, 2011. Disponível em: <http://www.maxwell.lambda.ele.puc-rio.br/trad_em_revista.php?strSecao=input0>. Acesso em: 30 out. 2012.

FREDDI, M.; LURAGHI, S. Titling for the opera house: a test case for universals of translation? In: INCALTERRA, L.; BISCIO, L.; MHAINNIN, M. A. N. **Audiovisual translation**: subtitles and subtitling. Oxford: Peter Lang, 2011. p. 55-86.

GABRIEL, M. H. C. **A segmentação linguística na legendagem para surdos e ensurdecidos (LSE)**: uma análise baseada em *corpus*. 2013. 59 f. Monografia (Especialização em Estudos da Tradução) – Universidade Federal do Ceará, Fortaleza-CE, 2013.

_____. **Problemas de segmentação linguística na legendagem para surdos e ensurdecidos (LSE) de “cheias de charme”**: uma análise baseada em corpus. 2015. 108 f. Dissertação (Mestrado em Linguística Aplicada) – Universidade Estadual do Ceará, Fortaleza-CE, 2015.

GAMBIER, Y. Introduction: screen transadaptation: perception and reception. **The Translator**, v. 9, n. 2, p. 191-205, 2003.

GOTTLIEB, Henrik. Subtitling: diagonal translation. **Perspectives: Studies in Translatology**, v. 2, n. 1, p. 101-121, 1994.

HOLMES, J. The name and nature of translation studies. In: VENUTI, L. (Ed.). **The Translation Studies Reader**. London: Routledge, 2000. p. 172-185.

IVARSSON, J.; CARROLL, M. **Subtitling**. Simrishamm, Suécia: TransEditHB, 1998.

JAKOBSON, R. On linguistic aspects of translation. In: BROWER, R. A. (Ed.). **On Translation**. Cambridge, MA: Harvard University Press, 1959. p. 232-239.

KALANTZI, D. **Subtitling for the Deaf and Hard of Hearing: A corpus-based methodology for the analysis of subtitles with a focus on segmentation and deletion**. 2008. 366 f. Tese (Doutorado) – School of Languages, Linguistics and Cultures of the University of Manchester, UK, 2008.

KARAMITROGLOU, F. A proposed set of subtitling standards in Europe. **Translation Journal**, v. 2, n. 2, p. 1-15, 1998. Disponível em: <<http://translationjournal.net/journal//04stndrd.htm>>. Acesso em: 07 jul. 2017.

LEECH, G. Adding linguistic annotation. In: WYNNE, M. **Developing Linguistic Corpora: a guide to good practice**. Oxford: Oxbrow Books, 2005. p. 17-29. Disponível em: <<http://ahds.ac.uk/linguistic-corpora/>>. Acesso em: 07 de julho de 2017.

LEMLE, M. **Análise Sintática: teoria geral e descrição do português**. 2. ed. São Paulo: Ática, 1989. 212p.

LEONARDO, G. M. N. **A Segmentação na legendagem para surdos e ensurdecidos em documentário da TV brasileira aberta: uma análise baseada em corpus**. 2016. 69 f. Monografia (Bacharelado em Letras Inglês) – Universidade Estadual do Ceará, Fortaleza-CE, 2016.

MCENERY, T.; XIAO, R. Y. **Corpus-based language studies: an advanced resource book**. London: Routledge, 2006.

MONTEIRO, S. M. M. **Legendagem para surdos e ensurdecidos (LSE) e legendagem para ouvintes: Um estudo sobre a segmentação e a velocidade na legendagem da campanha política de 2010**. 2016. 232 f. Tese (Doutorado) – Programa em Pós-Graduação em Linguística Aplicada, Universidade Estadual do Ceará, Fortaleza-CE, 2016.

NASCIMENTO, A. K. P. **Linguística de corpus e legendagem para surdos e ensurdecidos (LSE): uma análise baseada em corpus da tradução de efeitos sonoros na legenda de filmes**

brasileiros em DVD. 2013. 109 f. Dissertação (Mestrado em Linguística Aplicada) – Universidade Estadual do Ceará, Fortaleza-CE, 2013.

O'DONNELL, M. B. The use of annotated corpora for New Testament discourse analysis: a survey of current practice and future prospects. In: PORTER, S. E.; REED, J. T. (Ed.). **Discourse Analysis and the New Testament: results and applications**. Sheffield: Sheffield Academic Press, 1999. p. 71-117.

PEREGO, E. What would we read best? Hypotheses and suggestions for the location of line breaks in film subtitles. **The Sign Language Translator and Interpreter**, Manchester, v. 2, n. 1, p. 35-63, 2008.

PEREGO, E.; DEL MISSIER, F.; PORTA, M.; MOSCONI, M. The cognitive effectiveness of subtitle processing. **Media Psychology**, Philadelphia, v. 13, p. 243-272, 2010. Disponível em: <<http://www2.units.it/delmisfa/papers/SubtitlesProcessing2010.pdf>>. Acesso em: 21 maio 2014.

PERINI, M. A. **gramática do português brasileiro**. São Paulo: Parábola Editorial, 2010.

PRAXEDES FILHO, P. H. L. A reconciliação entre LT e LA: o caso da pesquisa no Laboratório de Tradução Audiovisual do Programa de Pós-Graduação em Linguística Aplicada – UECE. In: JORDÃO, C. M. **A Linguística Aplicada no Brasil: rumos e passagens**. Campinas: Pontes, 2016. p. 121-138.

REID, H. Literature on the screen: subtitle translation for public broadcasting. In: BART, W.; D'HAEN, T. (Ed.). **Something understood**. Studies in Anglo-Dutch literary translation. Amsterdã: Rodopi, 1990. p. 97-107.

SELVATICI, C. **Closed Caption: conquistas e questões**. 2010. 80 f. Dissertação (Mestrado) – Pontífica Universidade Católica do Rio, 2010.

SILVA, D. A. **A explicitação na legendagem do filme O nascimento de Cristo: um estudo baseado em corpus**. 2013. 137 f. Dissertação (Mestrado em Linguística Aplicada) – Universidade Estadual do Ceará, Fortaleza-CE, 2013.

SUBRIP 1.50 Beta 4. 2006. Programa disponível em: <www.divx-digest.com/software/subrip.html>.

VIEIRA, P. A. **A influência da segmentação e da velocidade na recepção de legendas para surdos e ensurdecidos (LSE)**. 2016. 244 f. Tese (Doutorado) – Programa em Pós-Graduação em Linguística Aplicada, Universidade Estadual do Ceará, Fortaleza-CE, 2016.

VIEIRA, P. A.; ARAÚJO, V. L. S.; CHAVES, É. G.; MONTEIRO, S. M. M. A segmentação na legendagem do documentário Globo Repórter: uma análise baseada em *corpus*. In: PONTES, V. O. *et al.* (Org.). **A tradução e suas interfaces: múltiplas perspectivas**. Curitiba: CRV, 2015. p. 65-85.

WORDSMITH Tools 7.0. Disponível em: <<http://www.lexically.net/wordsmith/index.html>>. Acesso em: 10 mar. 2017.

ANEXOS

ANEXO A – Problemas de segmentação linguística nas legendas em inglês

Filme: *Click* (2006)

<sub3><2L>

<t>00:02:49,560 --> 00:02:52,677</t>

<cpl24>How can I sleep with two <PROSEGL><SN_especif+SN>

<cpl32>Twinkie burglars roaming around?

<veloc_leg 56c/3.117s 17.966>

</sub3>

<sub7><2L>

<t>00:03:15,280 --> 00:03:16,998</t>

<cpl23>Honey, I think you just <PROSEGL><SV_verbo+SAdv>

<cpl18>opened the garage.

<veloc_leg 41c/1.718s 23.865>

</sub7>

<sub12><2L>

<t>00:03:29,440 --> 00:03:31,954</t>

<cpl16>The O'Doyles got <PROSEGL>

<cpl27>a universal remote control.

<veloc_leg 43c/2.514s 17.104>

</sub12>

<sub21><2L>

<t>00:03:53,440 --> 00:03:58,036</t>

<cpl43>All right. I want to, it's just I got a lot <PROSEGL><SAdv_adv+adv>

<cpl37>of things going on at work right now.

<veloc_leg 80c/4.596s 17.406>

</sub21>

<sub25><2L>

<t>00:04:10,640 --> 00:04:14,269</t>

<cpl30>That's something you shouldn't <PROSEGL><SV_verbo+SAdv>

<cpl38>know about for another 10 to 30 years.

<veloc_leg 68c/3.629s 18.738>

</sub25>

<sub58><2L>

<t>00:05:57,440 --> 00:05:59,635</t>

<cpl30>But he is one of our brightest <PROSEGL><SAdj_especif+SAdj>

<cpl17>young architects.

<veloc_leg 47c/2.195s 21.412>

</sub58>

<sub78><2L>

<t>00:06:54,480 --> 00:06:56,436</t>

<cpl18>Boob water's gotta <PROSEGL><SV_composto>

<cpl22>go someplace, Michael.

<veloc_leg 40c/1.956s 20.450>
</sub78>

<sub79><2L>
<t>00:06:56,640 --> 00:06:59,837</t>
<cpl34>Let's just take out the atrium and <PROSEGL><COORD_conj+oração>
<cpl37>run with that and there's your drain.
<veloc_leg 71c/3.197s 22.208>
</sub79>

<sub89><2L>
<t>00:07:24,560 --> 00:07:25,913</t>
<cpl17>Prince Habeeboo's <PROSEGL><SV_verbo+SAdv>
<cpl16>not chewing gum.
<veloc_leg 33c/1.353s 24.390>
</sub89>

<sub92><2L>
<t>00:07:33,360 --> 00:07:36,158</t>
<cpl35>I hope that atrium idea didn't take <PROSEGL><SV_verbo+SAdv>
<cpl31>too long to configure, Michael.
<veloc_leg 66c/2.798s 23.588>
</sub92>

<sub97><2L>
<t>00:07:49,840 --> 00:07:54,675</t>
<cpl30>I just need someone to come up <PROSEGL><SV_composto>
<cpl34>with a design proposal. Intrigued?
<veloc_leg 64c/4.835s 13.237>
</sub97>

<sub132><2L>
<t>00:09:45,960 --> 00:09:49,111</t>
<cpl33>You master that, people are gonna <PROSEGL><SV_composto>
<cpl38>think you're half a dolphin out there.
<veloc_leg 71c/3.151s 22.533>
</sub132>

<sub196><2L>
<t>00:12:52,480 --> 00:12:54,948</t>
<cpl25>...if you collapse trying <PROSEGL><SV_composto>
<cpl32>to put up the tent this weekend.
<veloc_leg 57c/2.468s 23.096>
</sub196>

<sub205><2L>
<t>00:13:14,880 --> 00:13:17,394</t>
<cpl24>Make sure you don't keep <PROSEGL><SV_composto>
<cpl31>disappointing the wrong people.
<veloc_leg 55c/2.514s 21.877>

</sub205>

<sub222><2L>

<t>00:14:22,080 --> 00:14:24,878</t>

<cpl26>You guys got any universal <PROSEGL><SAdj_especif+SAdj>

<cpl25>remote controls in there?

<veloc_leg 51c/2.798s 18.227>

</sub222>

<sub236><2L>

<t>00:15:44,640 --> 00:15:46,835</t>

<cpl24>You guys got a universal <PROSEGL><SAdj_especif+SAdj>

<cpl25>remote control back here?

<veloc_leg 49c/2.195s 22.323>

</sub236>

<sub241><2L>

<t>00:15:59,920 --> 00:16:02,957</t>

<cpl32>But the way my life's been going <PROSEGL><SV_verbo+SAdv>

<cpl36>lately, that ain't such a bad thing.

<veloc_leg 68c/3.037s 22.391>

</sub241>

<sub276><2L>

<t>00:18:17,440 --> 00:18:20,591</t>

<cpl17>So I'm gonna have <PROSEGL><SV_composto>

<cpl23>to just give it to you.

<veloc_leg 40c/3.151s 12.694>

</sub276>

<sub294><2L>

<t>00:19:37,680 --> 00:19:40,399</t>

<cpl32>They're very understanding about <PROSEGL><SP_prep+SV>

<cpl31>not going camping this weekend.

<veloc_leg 63c/2.719s 23.170>

</sub294>

<sub297><2L>

<t>00:19:45,800 --> 00:19:48,314</t>

<cpl16>Thanks for being <PROSEGL><SV_verbo+SAdv>

<cpl23>so cool about this too.

<veloc_leg 39c/2.514s 15.513>

</sub297>

<sub300><2L>

<t>00:19:55,040 --> 00:19:57,235</t>

<cpl28>...do you think you're gonna <PROSEGL><SV_composto>

<cpl24>have more time for us...

<veloc_leg 52c/2.195s 23.690>

</sub300>

<sub344><2L>
 <t>00:23:33,440 --> 00:23:37,069</t>
 <cpl31>You've already cheated on three <PROSEGL><SAdj_especif+SAdj>
 <cpl39>different husbands with their brothers.
 <veloc_leg 70c/3.629s 19.289>
 </sub344>

<sub364><2L>
 <t>00:24:36,320 --> 00:24:41,030</t>
 <cpl35>A two-run shot into the right-field <PROSEGL><SAdj_especif+SAdj>
 <cpl24>seats for Hideki Matsui.
 <veloc_leg 59c/4.71s 12.527>
 </sub364>

<sub470><2L>
 <t>00:30:20,880 --> 00:30:24,236</t>
 <cpl27>Plus, it's more fun to have <PROSEGL><SV_composto>
 <cpl33>to keep each other warm at night.
 <veloc_leg 60c/3.356s 17.878>
 </sub470>

<sub490><2L>
 <t>00:31:42,240 --> 00:31:45,152</t>
 <cpl33>Michael. Look, I'm heading out to <PROSEGL><SP_prep+SN>
 <cpl36>the Hamptons in a couple of minutes.
 <veloc_leg 69c/2.912s 23.695>
 </sub490>

<sub499><2L>
 <t>00:32:10,840 --> 00:32:13,434</t>
 <cpl29>Mr. Ammer, if you need female <PROSEGL><SN_subst+SAdj>
 <cpl29>companionship this weekend...
 <veloc_leg 58c/2.594s 22.359>
 </sub499>

<sub515><2L>
 <t>00:33:23,880 --> 00:33:26,189</t>
 <cpl30>But did you see her Kate Spade <PROSEGL><SN_subst+SAdj>
 <cpl25>handbag? It was gorgeous.
 <veloc_leg 55c/2.309s 23.820>
 </sub515>

<sub608><2L>
 <t>00:38:32,920 --> 00:38:36,151</t>
 <cpl21>It's simple. You must <PROSEGL><SV_composto>
 <cpl23>have been on autopilot.
 <veloc_leg 44c/3.231s 13.618>
 </sub608>

<sub663><2L>

<t>00:41:46,960 --> 00:41:49,394</t>
 <cpl31>Are you gonna have to work late <PROSEGL><SAdv_adv+adv>
 <cpl14>again tonight?
 <veloc_leg 45c/2.434s 18.488>
 </sub663>

<sub665><2L>
 <t>00:41:54,240 --> 00:41:57,437</t>
 <cpl36>Well, I'm gonna be partner soon, and <PROSEGL><COORD_conj+oração>
 <cpl36>I'm gonna need new building designs.
 <veloc_leg 72c/3.197s 22.521>
 </sub665>

<sub699><2L>
 <t>00:44:38,160 --> 00:44:41,789</t>
 <cpl34>Let me try to give you a real-life <PROSEGL><SN_especif+SN>
 <cpl38>example because I'm sure I've done it.
 <veloc_leg 72c/3.629s 19.840>
 </sub699>

<sub704><2L>
 <t>00:45:03,160 --> 00:45:05,799</t>
 <cpl22>Now, this is like when <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl31>one dude comes up to another...
 <veloc_leg 53c/2.639s 20.083>
 </sub704>

<sub713><2L>
 <t>00:45:52,440 --> 00:45:55,750</t>
 <cpl33>Well, if you guys love Ichiro and <PROSEGL><COORD_conj+oração>
 <cpl38>Matsui so much, let's order them both.
 <veloc_leg 71c/3.31s 21.450>
 </sub713>

<sub721><2L>
 <t>00:46:29,960 --> 00:46:32,030</t>
 <cpl18>They make me wanna <PROSEGL><SV_composto>
 <cpl21>chop my own dick off.
 <veloc_leg 39c/2.07s 18.841>
 </sub721>

<sub732><2L>
 <t>00:46:58,960 --> 00:47:02,669</t>
 <cpl30>Before we order, I was looking <PROSEGL><SV_verbo+SP>
 <cpl31>at our proposal. You know what?
 <veloc_leg 61c/3.709s 16.446>
 </sub732>

<sub739><2L>
 <t>00:47:22,200 --> 00:47:25,317</t>

<cpl32>That way we can get the hell out <PROSEGL><SP_prep+SN>
 <cpl38>of this dump, go to T.G.I. Friday's...
 <veloc_leg 70c/3.117s 22.457>
 </sub739>

<sub836><2L>
 <t>00:53:27,360 --> 00:53:30,989</t>
 <cpl32>It's your life. What you do with <PROSEGL><SP_prep+SN>
 <cpl37>the remote is your decision, Michael.
 <veloc_leg 69c/3.629s 19.014>
 </sub836>

<sub908><2L>
 <t>00:57:34,480 --> 00:57:37,233</t>
 <cpl33>Are you kidding me? You guys have <PROSEGL><SV_composto>
 <cpl32>been dating a couple months now?
 <veloc_leg 65c/2.753s 23.611>
 </sub908>

<sub937><2L>
 <t>00:59:01,320 --> 00:59:04,153</t>
 <cpl37>And when you are, you're just sitting <PROSEGL><SV_verbo+SAdv>
 <cpl29>there staring off into space.
 <veloc_leg 66c/2.833s 23.297>
 </sub937>

<sub955><2L>
 <t>00:59:51,440 --> 00:59:54,000</t>
 <cpl28>Don't yell at Peanut because <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl28>he found your Twinkie stash.
 <veloc_leg 56c/2.56s 21.875>
 </sub955>

<sub1014><2L>
 <t>01:03:21,680 --> 01:03:23,432</t>
 <cpl15>Yeah? I'm gonna <PROSEGL><SV_composto>
 <cpl27>take my clothes off, then--
 <veloc_leg 42c/1.752s 23.973>
 </sub1014>

<sub1044><2L>
 <t>01:05:40,240 --> 01:05:43,915</t>
 <cpl37>Your first day as a partner, you come <PROSEGL><SV_composto>
 <cpl31>to work in a fricking bathrobe?
 <veloc_leg 68c/3.675s 18.503>
 </sub1044>

<sub1046><2L>
 <t>01:05:46,680 --> 01:05:49,990</t>
 <cpl27>John, I just feel we should <PROSEGL><SV_composto>

<cpl26>stop wasting our energy...
 <veloc_leg 53c/3.31s 16.012>
 </sub1046>

<sub1047><2L>
 <t>01:05:50,200 --> 01:05:53,351</t>
 <cpl32>...on corporate brown-nosing and <PROSEGL><COORD_conj+oração>
 <cpl35>worrying who's got the better suit.
 <veloc_leg 67c/3.151s 21.263>
 </sub1047>

<sub1058><2L>
 <t>01:06:29,840 --> 01:06:32,354</t>
 <cpl29>Michael, Janine and I decided <PROSEGL><SV_composto>
 <cpl30>to drop out of the rat race...
 <veloc_leg 59c/2.514s 23.469>
 </sub1058>

<sub1078><2L>
 <t>01:07:30,000 --> 01:07:32,230</t>
 <cpl20>... or you'll end up <PROSEGL><SV_composto>
 <cpl21>killing yourself too.
 <veloc_leg 41c/2.23s 18.386>
 </sub1078>

<sub1127><2L>
 <t>01:10:42,000 --> 01:10:44,639</t>
 <cpl29>Hey, it's just that you don't <PROSEGL><SV_verbo+SAdv>
 <cpl29>mess with another man's wife!
 <veloc_leg 58c/2.639s 21.978>
 </sub1127>

<sub1139><2L>
 <t>01:11:17,160 --> 01:11:19,435</t>
 <cpl22>No. No, I didn't wanna <PROSEGL><SV_composto>
 <cpl20>waste all that time.
 <veloc_leg 42c/2.275s 18.462>
 </sub1139>

<sub1145><2L>
 <t>01:11:30,920 --> 01:11:34,629</t>
 <cpl37>The remote is lie-proof, so you can't <PROSEGL><SV_composto>
 <cpl29>change what already happened.
 <veloc_leg 66c/3.709s 17.795>
 </sub1145>

<sub1304><2L>
 <t>01:21:35,960 --> 01:21:38,394</t>
 <cpl22>I thought an angel was <PROSEGL><SV_composto>
 <cpl27>supposed to protect people.

<veloc_leg 49c/2.434s 20.131>
</sub1304>

<sub1366><2L>
<t>01:26:05,720 --> 01:26:10,430</t>
<cpl41>Okay, well, if that's your way of telling <PROSEGL><SV_(verbo)+oblíquo+SV>
<cpl38>me that I look pretty, then thank you.
<veloc_leg 79c/4.71s 16.773>
</sub1366>

<sub1376><2L>
<t>01:27:05,880 --> 01:27:09,839</t>
<cpl17>Why'd you make me <PROSEGL><SV_composto>
<cpl21>waste my entire life?
<veloc_leg 38c/3.959s 9.598>
</sub1376>

<sub1390><2L>
<t>01:27:55,840 --> 01:28:00,311</t>
<cpl32>One of them told me he was gonna <PROSEGL><SV_composto>
<cpl31>live till he was 200. Remember?
<veloc_leg 63c/4.471s 14.091>
</sub1390>

<sub1391><2L>
<t>01:28:02,160 --> 01:28:04,469</t>
<cpl20>You both really need <PROSEGL><SV_composto>
<cpl18>to get some sleep.
<veloc_leg 38c/2.309s 16.457>
</sub1391>

<sub1400><2L>
<t>01:28:25,960 --> 01:28:28,599</t>
<cpl29>Guys, I'm sorry. You're gonna <PROSEGL><SV_composto>
<cpl33>have to leave. He needs his rest.
<veloc_leg 62c/2.639s 23.494>
</sub1400>

<sub1463><2L>
<t>01:33:54,360 --> 01:33:57,670</t>
<cpl33>You come over tomorrow night, and <PROSEGL><COORD_conj+oração>
<cpl40>the night after that. Any time you want.
<veloc_leg 73c/3.31s 22.054>
</sub1463>

<sub1465><2L>
<t>01:34:00,960 --> 01:34:03,679</t>
<cpl28>Thank you so much for giving <PROSEGL><SN_núcleo>
<cpl34>birth to me, I know it hurt. Okay?
<veloc_leg 62c/2.719s 22.803>

</sub1465>

<sub1493><2L>

<t>01:35:38,360 --> 01:35:40,191</t>

<cpl21>So tomorrow I'm gonna <PROSEGL><SV_composto>

<cpl19>teach you calculus.

<veloc_leg 40c/1.831s 21.846>

</sub1493>

<sub1521><2L>

<t>01:38:07,040 --> 01:38:09,474</t>

<cpl14>Hey, who wants <PROSEGL><SV_composto>

<cpl23>to have a pillow fight?

<veloc_leg 37c/2.434s 15.201>

</sub1521>

Filme: *Exodus* (2015)

<sub3><2L>

<t>00:02:19,680 --> 00:02:22,740</t>

<cpl23>The Hittites are trying <PROSEGL><SV_composto>

<cpl20>to cross the border.

<veloc_leg 43c/3.06s 14.052>

</sub3>

<sub364><2L>

<t>00:32:59,418 --> 00:33:02,148</t>

<cpl29>And then, he told me the most <PROSEGL><SAdj_especif+SAdj>

<cpl19>unbelievable story.

<veloc_leg 48c/2.73s 17.582>

</sub364>

<sub440><2L>

<t>00:37:16,975 --> 00:37:19,842</t>

<cpl39>You want to believe it, because it's an <PROSEGL><SN_especif+SN>

<cpl31>opportunity to be rid of him...

<veloc_leg 70c/2.867s 24.416>

</sub440>

<sub449><2L>

<t>00:38:06,725 --> 00:38:11,628</t>

<cpl38>It isn't every man who would protect a <PROSEGL><SN_especif+SN>

<cpl44>servant who's nothing more to him than that.

<veloc_leg 82c/4.903s 16.724>

</sub449>

<sub547><2L>

<t>00:51:56,087 --> 00:52:00,114</t>

<cpl38>I want to discuss why you seem to want <PROSEGL><SV_composto>

<cpl27>to discuss it all the time.

<veloc_leg 65c/4.027s 16.141>
</sub547>

<sub549><2L>
<t>00:52:01,459 --> 00:52:03,859</t>
<cpl30>it should have been your first <PROSEGL><SN_especif+SN>
<cpl40>construction order on taking the throne.
<veloc_leg 70c/2.4s 29.167>
</sub549>

<sub714><2L>
<t>01:09:16,226 --> 01:09:19,354</t>
<cpl22>Things are better than <PROSEGL><SUBORD_conj/pron_rel+oração>
<cpl27>they ever have been, Moses.
<veloc_leg 49c/3.128s 15.665>
</sub714>

<sub747><2L>
<t>01:11:02,265 --> 01:11:04,756</t>
<cpl20>You're not seriousIy <PROSEGL><SV_verbo+SAdv>
<cpl25>considering his proposal.
<veloc_leg 45c/2.491s 18.065>
</sub747>

<sub841><2L>
<t>01:20:51,479 --> 01:20:53,140</t>
<cpl26>Still, if you're not going <PROSEGL><SV_composto>
<cpl18>to listen to me...
<veloc_leg 44c/1.661s 26.490>
</sub841>

<sub972><2L>
<t>01:38:07,814 --> 01:38:10,715</t>
<cpl29>the People who I grew up with <PROSEGL><SP_prep+SV>
<cpl20>suffering this much.
<veloc_leg 49c/2.901s 16.891>
</sub972>

<sub1062><2L>
<t>01:51:51,737 --> 01:51:54,297</t>
<cpl17>if we're going to <PROSEGL><SV_composto>
<cpl27>recapture 400,000 slaves...
<veloc_leg 44c/2.56s 17.188>
</sub1062>

<sub1208><2L>
<t>02:15:22,914 --> 02:15:24,745</t>
<cpl23>You don't think they'll <PROSEGL><SV_composto>
<cpl24>let us settle in Canaan?
<veloc_leg 47c/1.831s 25.669>

</sub1208>

Filme: *Flightplan* (2004)

<sub3><2L>

<t>00:03:37,330 --> 00:03:41,426</t>

<cpl19>avid, could we just <PROSEGL><SV_verbo+SAdv>

<cpl34>sit in the courtyard for a minute?

<veloc_leg 53c/4.096s 12.939>

</sub3>

<sub28><2L>

<t>00:08:57,317 --> 00:09:02,152</t>

<cpl32>Unattended luggage is subject to <PROSEGL><SP_prep+SN>

<cpl40>collection and search by airport police.

<veloc_leg 72c/4.835s 14.891>

</sub28>

<sub90><2L>

<t>00:14:51,070 --> 00:14:55,370</t>

<cpl36>They have to nick the ice off or the <PROSEGL><SN_especif+SN>

<cpl38>air won't pass over the wing properly.

<veloc_leg 74c/4.3s 17.209>

</sub90>

<sub107><2L>

<t>00:19:05,224 --> 00:19:09,092</t>

<cpl33>Course at 36,000ft you can't just <PROSEGL><SV_verbo+SAdv>

<cpl31>up and walk out of the theater.

<veloc_leg 64c/3.868s 16.546>

</sub107>

<sub158><2L>

<t>00:24:27,679 --> 00:24:31,308</t>

<cpl38>But I tell you what. I will personally <PROSEGL><SV_verbo+SAdv>

<cpl27>go and speak to the captain

<veloc_leg 65c/3.629s 17.911>

</sub158>

<sub196><2L>

<t>00:27:25,157 --> 00:27:29,389</t>

<cpl39>Ma'am, we don't have any record of your <PROSEGL><SN_especif+SN>

<cpl35>daughter ever having been on board.

<veloc_leg 74c/4.232s 17.486>

</sub196>

<sub198><2L>

<t>00:27:34,132 --> 00:27:39,434</t>

<cpl33>The departure gate say they never <PROSEGL><SV_verbo+SAdv>

<cpl38>checked a Julia Pratt onto this plane.

<veloc_leg 71c/5.302s 13.391>
</sub198>

<sub236><2L>
<t>00:30:32,944 --> 00:30:37,108</t>
<cpl38>It was under her blanket. She wouldn't <PROSEGL><SV_composto>
<cpl28>have walked away without it.
<veloc_leg 66c/4.164s 15.850>
</sub236>

<sub237><2L>
<t>00:30:38,650 --> 00:30:43,485</t>
<cpl38>Have you had anything to drink on this <PROSEGL><SN_especif+SN>
<cpl34>flight? Anything alcoholic at all?
<veloc_leg 72c/4.835s 14.891>
</sub237>

<sub258><2L>
<t>00:32:14,713 --> 00:32:17,204</t>
<cpl16>and that I'm now <PROSEGL><SV_verbo+SAdv>
<cpl27>pretending to look for her.
<veloc_leg 43c/2.491s 17.262>
</sub258>

<sub262><2L>
<t>00:32:27,826 --> 00:32:32,559</t>
<cpl38>If a pilot has an inaccurate passenger <PROSEGL><SN_subst+SAdj>
<cpl33>manifest, he must land the plane.
<veloc_leg 71c/4.733s 15.001>
</sub262>

<sub264><2L>
<t>00:32:36,234 --> 00:32:40,364</t>
<cpl37>I'm telling you it's inaccurate. That <PROSEGL><SUBORD_conj/pron_rel+oração>
<cpl35>means you have to search the plane,
<veloc_leg 72c/4.13s 17.433>
</sub264>

<sub265><2L>
<t>00:32:40,438 --> 00:32:44,306</t>
<cpl22>just like you would if <PROSEGL><SUBORD_conj/pron_rel+oração>
<cpl39>it was your own child that was missing.
<veloc_leg 61c/3.868s 15.770>
</sub265>

<sub269><2L>
<t>00:32:58,023 --> 00:33:02,585</t>
<cpl39>But she did get in that car. She got on <PROSEGL><SP_prep+SN>
<cpl40>this plane like every other person here.

<veloc_leg 79c/4.562s 17.317>
</sub269>

<sub282><2L>
<t>00:34:08,159 --> 00:34:13,290</t>
<cpl29>but the captain requests that <PROSEGL><SUBORD_conj/pron_rel+oração>
<cpl37>all passengers return to their seats.
<veloc_leg 66c/5.131s 12.863>
</sub282>

<sub314><2L>
<t>00:37:12,777 --> 00:37:17,009</t>
<cpl28>No. It's just that I'm dying <PROSEGL><SV_composto>
<cpl38>to watch you crawl through that hatch.
<veloc_leg 66c/4.232s 15.595>
</sub314>

<sub330><2L>
<t>00:38:32,790 --> 00:38:36,658</t>
<cpl34>There is no way a child could have <PROSEGL><SV_composto>
<cpl34>gained access to any of the holds.
<veloc_leg 68c/3.868s 17.580>
</sub330>

<sub342><2L>
<t>00:39:18,403 --> 00:39:23,170</t>
<cpl36>Mr. Carson, I am responsible for the <PROSEGL><SN_especific+SN>
<cpl40>safety of every passenger on this plane.
<veloc_leg 76c/4.767s 15.943>
</sub342>

<sub343><2L>
<t>00:39:23,241 --> 00:39:27,268</t>
<cpl39>Even the delusional ones and I will not <PROSEGL><SV_verbo+SAdv>
<cpl32>have a passenger out of her seat
<veloc_leg 71c/4.027s 17.631>
</sub343>

<sub352><2L>
<t>00:40:02,046 --> 00:40:05,675</t>
<cpl30>because every one of my flight <PROSEGL><SN_núcleo>
<cpl34>attendants are looking for a child
<veloc_leg 64c/3.629s 17.636>
</sub352>

<sub417><2L>
<t>00:44:38,889 --> 00:44:43,155</t>
<cpl37>Fine, let's arrest them, just as soon <PROSEGL><SAdv_adv+adv>
<cpl32>as you answer one more Question.
<veloc_leg 69c/4.266s 16.174>

</sub417>

<sub423><2L>

<t>00:45:09,086 --> 00:45:12,647</t>

<cpl32>Your husband's death is starting <PROSEGL><SV_composto>

<cpl31>to make a lot more sense to me.

<veloc_leg 63c/3.561s 17.692>

</sub423>

<sub460><2L>

<t>00:48:06,163 --> 00:48:09,963</t>

<cpl24>Our air marshal is going <PROSEGL><SV_composto>

<cpl32>to escort you back to your seat.

<veloc_leg 56c/3.8s 14.737>

</sub460>

<sub499><2L>

<t>00:51:41,522 --> 00:51:46,221</t>

<cpl20>When something's too <PROSEGL><SAdj_especif+SAdj>

<cpl26>overwhelming, too painful,

<veloc_leg 46c/4.699s 9.789>

</sub499>

<sub502><1L>

<t>00:51:53,867 --> 00:51:57,359</t>

<cpl29>So you can walk with avid and <PROSEGL><COORD_conj+oração>

<veloc_leg 29c/3.492s 8.305>

</sub502>

<sub512><2L>

<t>00:52:44,752 --> 00:52:50,384</t>

<cpl36>They'll always be there. And you can <PROSEGL><SV_composto>

<cpl34>swim to them any time you need to.

<veloc_leg 70c/5.632s 12.429>

</sub512>

<sub579><2L>

<t>01:01:50,063 --> 01:01:53,931</t>

<cpl35>My daughter's alive. You know she's <PROSEGL><SN_subst+SAdj>

<cpl36>alive. You know she's on this plane.

<veloc_leg 71c/3.868s 18.356>

</sub579>

<sub610><2L>

<t>01:07:49,722 --> 01:07:53,556</t>

<cpl34>The plane's to be deboarded on the <PROSEGL><SAdv_adv+adv>

<cpl36>tarmac, passengers first, then crew.

<veloc_leg 70c/3.834s 18.258>

</sub610>

<sub634><2L>
 <t>01:09:25,118 --> 01:09:29,646</t>
 <cpl40>Then why are you panicking? Everything's <PROSEGL><SV_composto>
 <cpl32>working out how we wanted it to.
 <veloc_leg 72c/4.528s 15.901>
 </sub634>

<sub637><2L>
 <t>01:09:36,829 --> 01:09:40,629</t>
 <cpl36>What difference does it make? She'll <PROSEGL><SV_composto>
 <cpl34>be dead ten minutes after we land.
 <veloc_leg 70c/3.8s 18.421>
 </sub637>

<sub659><2L>
 <t>01:11:11,290 --> 01:11:14,817</t>
 <cpl23>I've notified them that <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl35>we have a missing person situation.
 <veloc_leg 58c/3.527s 16.445>
 </sub659>

<sub669><2L>
 <t>01:13:00,433 --> 01:13:05,336</t>
 <cpl39>All units, the suspect will be escorted <PROSEGL><SV_composto>
 <cpl34>off the plane by a flight marshal.
 <veloc_leg 73c/4.903s 14.889>
 </sub669>

Filme: *Madagascar* (2005)

<sub4><2L>
 <t>00:01:40,063 --> 00:01:42,964</t>
 <cpl27>Come on, Marty. Just wanted <PROSEGL><SV_composto>
 <cpl29>to wish you a happy birthday.
 <veloc_leg 56c/2.901s 19.304>
 </sub4>

<sub10><2L>
 <t>00:01:58,615 --> 00:02:03,643</t>
 <cpl34>Please hop on top of my sterilized <PROSEGL><SN_subst+SAdj>
 <cpl30>examination table, if you may.
 <veloc_leg 64c/5.028s 12.729>
 </sub10>

<sub30><2L>
 <t>00:03:02,112 --> 00:03:04,808</t>
 <cpl26>You just need to break out <PROSEGL><SP_prep+SN>
 <cpl23>of that boring routine.
 <veloc_leg 49c/2.696s 18.175>
 </sub30>

<sub78><2L>
 <t>00:06:23,947 --> 00:06:26,745</t>
 <cpl15>Yeah! You don't <PROSEGL><SV_verbo+SAdv>
 <cpl26>see that on Animal Planet.
 <veloc_leg 41c/2.798s 14.653>
 </sub78>

<sub101><2L>
 <t>00:07:41,958 --> 00:07:45,985</t>
 <cpl18>We're going to the <PROSEGL><SN_especif+SN>
 <cpl31>wide-open spaces of Antarctica.
 <veloc_leg 49c/4.027s 12.168>
 </sub101>

<sub115><2L>
 <t>00:08:25,535 --> 00:08:29,164</t>
 <cpl28>You guys are great. You're a <PROSEGL><SN_especif+SN>
 <cpl36>great crowd. Give yourselves a hand.
 <veloc_leg 64c/3.629s 17.636>
 </sub115>

<sub166><2L>
 <t>00:11:15,438 --> 00:11:18,839</t>
 <cpl16>This is a highly <PROSEGL><SV_verbo+SAdv>
 <cpl26>refined type of food thing
 <veloc_leg 42c/3.401s 12.349>
 </sub166>

<sub168><2L>
 <t>00:11:20,810 --> 00:11:24,041</t>
 <cpl28>You ever thought there might <PROSEGL><SV_composto>
 <cpl33>be more to life than steak, Alex?
 <veloc_leg 61c/3.231s 18.880>
 </sub168>

<sub186><2L>
 <t>00:13:01,044 --> 00:13:05,708</t>
 <cpl34>I mean, I tell you, it just doesn't <PROSEGL><SV_verbo+SAdv>
 <cpl35>get any better than this, you know?
 <veloc_leg 69c/4.664s 14.794>
 </sub186>

<sub191><2L>
 <t>00:13:20,997 --> 00:13:25,093</t>
 <cpl38>Everybody has days when they think the <PROSEGL><SN_especif+SN>
 <cpl38>grass might be greener somewhere else.
 <veloc_leg 76c/4.096s 18.555>
 </sub191>

<sub219><2L>
 <t>00:14:42,679 --> 00:14:46,274</t>
 <cpl32>So, what are you going to do? Go <PROSEGL><SV_composto>
 <cpl36>running off to the wild by yourself?
 <veloc_leg 68c/3.595s 18.915>
 </sub219>

<sub224><2L>
 <t>00:14:52,722 --> 00:14:56,158</t>
 <cpl20>It's a straight shot <PROSEGL><SV_verbo+SAdv>
 <cpl35>down Fifth Avenue to Grand Central.
 <veloc_leg 55c/3.436s 16.007>
 </sub224>

<sub230><2L>
 <t>00:15:14,143 --> 00:15:16,873</t>
 <cpl25>Don't do that. You really <PROSEGL><SV_verbo+SAdv>
 <cpl21>had me worried there.
 <veloc_leg 46c/2.73s 16.850>
 </sub230>

<sub232><2L>
 <t>00:15:20,984 --> 00:15:24,545</t>
 <cpl23>Yeah, me too. I'll need <PROSEGL><SV_composto>
 <cpl30>to rest my voice for tomorrow.
 <veloc_leg 53c/3.561s 14.883>
 </sub232>

<sub251><2L>
 <t>00:16:51,941 --> 00:16:56,310</t>
 <cpl34>I got up to pee and looked over at <PROSEGL><SP_prep+SN>
 <cpl38>Marty's pen, which I usually don't do.
 <veloc_leg 72c/4.369s 16.480>
 </sub251>

<sub266><2L>
 <t>00:17:33,249 --> 00:17:36,980</t>
 <cpl34>We've got a lost zebra probably on <PROSEGL><SP_prep+SN>
 <cpl38>the way to Connecticut, and we need...
 <veloc_leg 72c/3.731s 19.298>
 </sub266>

<sub284><2L>
 <t>00:18:53,095 --> 00:18:56,258</t>
 <cpl21>Well, of course we're <PROSEGL><SV_composto>
 <cpl26>going to throw poo at him.
 <veloc_leg 47c/3.163s 14.859>
 </sub284>

<sub289><2L>

<t>00:19:13,382 --> 00:19:17,113</t>
 <cpl31>That room has some nifty little <PROSEGL><SN_subst+SAdj>
 <cpl35>sinks you can wash up in, and look!
 <veloc_leg 66c/3.731s 17.690>
 </sub289>

<sub310><2L>
 <t>00:20:30,693 --> 00:20:34,151</t>
 <cpl28>Did that say "Grand Central <PROSEGL><SN_núcleo>
 <cpl40>Station" or "my aunt's constipation"?
 <veloc_leg 68c/3.458s 19.665>
 </sub310>

<sub322><2L>
 <t>00:21:26,515 --> 00:21:29,416</t>
 <cpl20>Looks like I'll have <PROSEGL><SV_composto>
 <cpl27>to take the Stamford local.
 <veloc_leg 47c/2.901s 16.201>
 </sub322>

<sub348><2L>
 <t>00:22:51,834 --> 00:22:54,394</t>
 <cpl24>You're not good with the <PROSEGL><SN_especific+SN>
 <cpl24>"putting words together
 <veloc_leg 48c/2.56s 18.750>
 </sub348>

<sub359><2L>
 <t>00:23:23,866 --> 00:23:27,996</t>
 <cpl36>We'll take my little friend home and <PROSEGL><COORD_conj+oração>
 <cpl37>forget this ever happened. All right?
 <veloc_leg 73c/4.13s 17.676>
 </sub359>

<sub367><2L>
 <t>00:24:26,737 --> 00:24:30,605</t>
 <cpl35>is an example of what animal rights <PROSEGL><SN_subst+SAdj>
 <cpl36>wackos have been shouting for years:
 <veloc_leg 71c/3.868s 18.356>
 </sub367>

<sub378><2L>
 <t>00:25:12,349 --> 00:25:15,944</t>
 <cpl18>Oh, no, they can't <PROSEGL><SV_verbo+SAdv>
 <cpl20>transfer me. Not me!
 <veloc_leg 38c/3.595s 10.570>
 </sub378>

<sub400><2L>
 <t>00:26:12,509 --> 00:26:15,103</t>

<cpl18>No other zoo could <PROSEGL><SV_composto>
 <cpl23>afford my medical care.
 <veloc_leg 41c/2.594s 15.806>
 </sub400>

<sub405><2L>
 <t>00:26:26,557 --> 00:26:30,084</t>
 <cpl21>Because of me? I fail <PROSEGL><SV_composto>
 <cpl28>to see how this is my fault.
 <veloc_leg 49c/3.527s 13.893>
 </sub405>

<sub408><2L>
 <t>00:26:36,233 --> 00:26:40,761</t>
 <cpl36>"I don't know who I am. I got to go <PROSEGL><SV_composto>
 <cpl38>find myself in the wild." Oh, please.
 <veloc_leg 74c/4.528s 16.343>
 </sub408>

<sub519><2L>
 <t>00:34:22,999 --> 00:34:27,163</t>
 <cpl26>cleverly simulated natural <PROSEGL><SN_subst+SAdj>
 <cpl34>environment, wide-open enclosures.
 <veloc_leg 60c/4.164s 14.409>
 </sub519>

<sub523><2L>
 <t>00:34:34,411 --> 00:34:38,643</t>
 <cpl21>San Diego? What could <PROSEGL><SV_composto>
 <cpl24>be worse than San Diego?
 <veloc_leg 45c/4.232s 10.633>
 </sub523>

<sub530><2L>
 <t>00:34:54,697 --> 00:34:57,165</t>
 <cpl21>Now, look. We're just <PROSEGL><SV_verbo+SAdv>
 <cpl25>going to find the people,
 <veloc_leg 46c/2.468s 18.639>
 </sub530>

<sub542><2L>
 <t>00:35:30,600 --> 00:35:33,728</t>
 <cpl23>Marty, look, you've got <PROSEGL><SV_composto>
 <cpl35>to be just a little bit more und...
 <veloc_leg 58c/3.128s 18.542>
 </sub542>

<sub548><2L>
 <t>00:35:50,720 --> 00:35:55,919</t>
 <cpl40>First they tell you, "We got this great <PROSEGL><SN_subst+SAdj>

<cpl40>open plan thing. Let animals run wild."
 <veloc_leg 80c/5.199s 15.388>
 </sub548>

<sub591><2L>
 <t>00:38:32,215 --> 00:38:35,548</t>
 <cpl26>Because we've been sitting <PROSEGL><SV_verbo+SAdv>
 <cpl35>on that beach back there for hours,
 <veloc_leg 61c/3.333s 18.302>
 </sub591>

<sub607><2L>
 <t>00:39:30,640 --> 00:39:34,406</t>
 <cpl20>Get up, Mort. Do not <PROSEGL><SV_verbo+SAdv>
 <cpl30>be near the king's feet, okay?
 <veloc_leg 50c/3.766s 13.277>
 </sub607>

<sub622><2L>
 <t>00:40:11,814 --> 00:40:15,545</t>
 <cpl14>to see whether <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl25>these are savage killers.
 <veloc_leg 39c/3.731s 10.453>
 </sub622>

<sub655><2L>
 <t>00:42:33,089 --> 00:42:37,355</t>
 <cpl34>Shame on you, Maurice. Can you not <PROSEGL><SV_verbo+SAdv>
 <cpl37>see that you have insulted the freak?
 <veloc_leg 71c/4.266s 16.643>
 </sub655>

<sub662><2L>
 <t>00:42:57,547 --> 00:42:59,811</t>
 <cpl20>Is this some sort of <PROSEGL><SP_prep+SN>
 <cpl19>inbreeding program?
 <veloc_leg 39c/2.264s 17.226>
 </sub662>

<sub663><2L>
 <t>00:42:59,882 --> 00:43:02,976</t>
 <cpl30>I say we just got to ask these <PROSEGL><SN_especific+SN>
 <cpl27>bozos where the people are.
 <veloc_leg 57c/3.094s 18.423>
 </sub663>

<sub689><2L>
 <t>00:44:24,600 --> 00:44:27,899</t>
 <cpl23>Look. There's obviously <PROSEGL><SV_verbo+SAdv>
 <cpl27>just been a little mistake.

<veloc_leg 50c/3.299s 15.156>
</sub689>

<sub690><2L>
<t>00:44:27,970 --> 00:44:30,530</t>
<cpl17>The people didn't <PROSEGL><SV_verbo+SAdv>
<cpl24>dump us here on purpose.
<veloc_leg 41c/2.56s 16.016>
</sub690>

<sub695><2L>
<t>00:44:50,059 --> 00:44:54,621</t>
<cpl22>Well, boys, it's going <PROSEGL><SV_composto>
<cpl35>to be ice-cold sushi for breakfast.
<veloc_leg 57c/4.562s 12.495>
</sub695>

<sub697><2L>
<t>00:45:01,837 --> 00:45:05,898</t>
<cpl22>Well, since I'm doomed <PROSEGL><SV_composto>
<cpl31>to die on this forsaken island,
<veloc_leg 53c/4.061s 13.051>
</sub697>

<sub721><2L>
<t>00:46:19,548 --> 00:46:23,507</t>
<cpl35>This is the good side for those who <PROSEGL><SUBORD_conj/pron_rel+oração>
<cpl40>love New York and care about going home.
<veloc_leg 75c/3.959s 18.944>
</sub721>

<sub730><2L>
<t>00:46:47,376 --> 00:46:50,903</t>
<cpl27>That side stinks! You're on <PROSEGL><SP_prep+SN>
<cpl33>the Jersey side of this cesspool!
<veloc_leg 60c/3.527s 17.012>
</sub730>

<sub736><2L>
<t>00:47:25,114 --> 00:47:28,174</t>
<cpl18>I've been standing <PROSEGL><SV_verbo+SAdv>
<cpl20>here for hours, man.
<veloc_leg 38c/3.06s 12.418>
</sub736>

<sub737><2L>
<t>00:47:28,250 --> 00:47:30,912</t>
<cpl18>How long do I have <PROSEGL><SV_composto>
<cpl18>to pose like this?
<veloc_leg 36c/2.662s 13.524>

</sub737>

<sub747><2L>

<t>00:48:00,749 --> 00:48:04,742</t>

<cpl27>That's wild fire. We're not <PROSEGL><SV_verbo+SAdv>

<cpl32>using wild fire on Lady Liberty.

<veloc_leg 59c/3.993s 14.776>

</sub747>

<sub785><2L>

<t>00:50:06,508 --> 00:50:10,672</t>

<cpl27>Come on, y'all. They're not <PROSEGL><SV_verbo+SAdv>

<cpl31>attacking us this very instant.

<veloc_leg 58c/4.164s 13.929>

</sub785>

<sub798><2L>

<t>00:50:54,323 --> 00:50:57,588</t>

<cpl26>I tell you, that dude just <PROSEGL><SV_verbo+SAdv>

<cpl31>gives me the heebiedabajebies.

<veloc_leg 57c/3.265s 17.458>

</sub798>

<sub810><2L>

<t>00:52:01,824 --> 00:52:04,816</t>

<cpl23>I don't know, you might <PROSEGL><SV_verbo+SAdv>

<cpl20>even enjoy yourself.

<veloc_leg 43c/2.992s 14.372>

</sub810>

<sub822><2L>

<t>00:52:59,248 --> 00:53:03,981</t>

<cpl22>But I've been thinking <PROSEGL><SV_verbo+SAdv>

<cpl35>about what you said, and I'm sorry.

<veloc_leg 57c/4.733s 12.043>

</sub822>

<sub829><2L>

<t>00:53:41,823 --> 00:53:45,623</t>

<cpl35>Oh, you don't swallow it. It's just <PROSEGL><SV_verbo+SAdv>

<cpl35>temporary till the plumbing's done.

<veloc_leg 70c/3.8s 18.421>

</sub829>

<sub840><2L>

<t>00:54:21,863 --> 00:54:25,094</t>

<cpl20>Well, maybe it could <PROSEGL><SV_composto>

<cpl19>use a little lemon.

<veloc_leg 39c/3.231s 12.071>

</sub840>

<sub852><2L>
 <t>00:55:07,876 --> 00:55:10,310</t>
 <cpl20>I think I'm going to <PROSEGL><SV_composto>
 <cpl17>hit the sack too.
 <veloc_leg 37c/2.434s 15.201>
 </sub852>

<sub861><2L>
 <t>00:56:23,318 --> 00:56:26,253</t>
 <cpl20>How can you have the <PROSEGL><SN_especif+SN>
 <cpl28>heebie-jeebies for Mr. Alex?
 <veloc_leg 48c/2.935s 16.354>
 </sub861>

<sub866><2L>
 <t>00:56:38,600 --> 00:56:42,764</t>
 <cpl19>Soon we will put my <PROSEGL><SN_especif+SN>
 <cpl25>excellent plan to action.
 <veloc_leg 44c/4.164s 10.567>
 </sub866>

<sub889><2L>
 <t>00:58:24,739 --> 00:58:28,641</t>
 <cpl37>That's the thing that you were always <PROSEGL><SV_verbo+SAdv>
 <cpl36>looking at, but it's actually there.
 <veloc_leg 73c/3.902s 18.708>
 </sub889>

<sub907><2L>
 <t>00:59:10,886 --> 00:59:12,877</t>
 <cpl13>You just have <PROSEGL><SV_composto>
 <cpl27>to let out that inner lion.
 <veloc_leg 40c/1.991s 20.090>
 </sub907>

<sub935><2L>
 <t>01:01:02,964 --> 01:01:05,865</t>
 <cpl16>Oh. No, I really <PROSEGL><SV_verbo+SAdv>
 <cpl28>don't think I could... Okay.
 <veloc_leg 44c/2.901s 15.167>
 </sub935>

<sub942><2L>
 <t>01:01:30,325 --> 01:01:34,193</t>
 <cpl34>See, if he is the king, then where <PROSEGL><SV_verbo+SAdv>
 <cpl31>is his crown? I've got a crown.
 <veloc_leg 65c/3.868s 16.805>
 </sub942>

<sub959><2L>
 <t>01:02:55,777 --> 01:03:00,237</t>
 <cpl37>Come, come, Maurice. What is a simple <PROSEGL><SN_subst+SAdj>
 <cpl35>bite on the buttocks among friends?
 <veloc_leg 72c/4.46s 16.143>
 </sub959>

<sub970><2L>
 <t>01:03:25,907 --> 01:03:30,708</t>
 <cpl29>The foosas will come back and <PROSEGL><COORD_conj+oração>
 <cpl38>gobble us with their mouths because...
 <veloc_leg 67c/4.801s 13.955>
 </sub970>

<sub978><2L>
 <t>01:03:48,429 --> 01:03:51,262</t>
 <cpl24>What? Come on, do I look <PROSEGL><SV_composto>
 <cpl20>like a steak to you?
 <veloc_leg 44c/2.833s 15.531>
 </sub978>

<sub980><2L>
 <t>01:03:52,634 --> 01:03:55,535</t>
 <cpl23>See, I told you I don't <PROSEGL><SV_verbo+SAdv>
 <cpl34>look like a... Wait. What you say?
 <veloc_leg 57c/2.901s 19.648>
 </sub980>

<sub1003><2L>
 <t>01:07:30,284 --> 01:07:33,219</t>
 <cpl20>And we are not going <PROSEGL><SV_composto>
 <cpl36>to lay down like a bunch of Melmans.
 <veloc_leg 56c/2.935s 19.080>
 </sub1003>

<sub1038><2L>
 <t>01:09:29,237 --> 01:09:32,400</t>
 <cpl15>Well, boys, our <PROSEGL><SN_especif+SN>
 <cpl33>monochromatic friend's in danger.
 <veloc_leg 48c/3.163s 15.175>
 </sub1038>

<sub1041><2L>
 <t>01:09:37,378 --> 01:09:40,905</t>
 <cpl31>Kowalski. We'll need to win the <PROSEGL><SN_especif+SN>
 <cpl36>hearts and the minds of the natives.
 <veloc_leg 67c/3.527s 18.996>
 </sub1041>

<sub1055><2L>

<t>01:10:36,170 --> 01:10:40,231</t>
 <cpl35>We can go back to civilization, and <PROSEGL><COORD_conj+oração>
 <cpl38>everything will be like it used to be.
 <veloc_leg 73c/4.061s 17.976>
 </sub1055>

<sub1070><2L>
 <t>01:12:23,678 --> 01:12:26,340</t>
 <cpl20>Please don't make me <PROSEGL><SV_(verbo)+oblíquo+SV>
 <cpl20>sing this by myself.
 <veloc_leg 40c/2.662s 15.026>
 </sub1070>

<sub1071><2L>
 <t>01:12:26,414 --> 01:12:29,611</t>
 <cpl21>You really don't want <PROSEGL><SV_composto>
 <cpl31>to hear me sing this by myself.
 <veloc_leg 52c/3.197s 16.265>
 </sub1071>

<sub1115><2L>
 <t>01:15:47,348 --> 01:15:51,409</t>
 <cpl18>I never, ever want <PROSEGL><SV_composto>
 <cpl28>to see you on my turf again.
 <veloc_leg 46c/4.061s 11.327>
 </sub1115>

<sub1158><2L>
 <t>01:18:03,317 --> 01:18:07,879</t>
 <cpl28>After much deep and profound <PROSEGL><SN_subst+SAdj>
 <cpl28>brain things inside my head,
 <veloc_leg 56c/4.562s 12.275>
 </sub1158>

<sub1184><2L>
 <t>01:19:10,517 --> 01:19:14,112</t>
 <cpl30>Don't you think we should tell <PROSEGL><SV_(verbo)+oblíquo+SV>
 <cpl32>them that the boat's out of gas?
 <veloc_leg 62c/3.595s 17.246>
 </sub1184>

Filme: *Social network* (2010)

<sub17><2L>
 <t>00:00:48,028 --> 00:00:49,928</t>
 <cpl24>sometimes not singing in <PROSEGL><SP_prep+SN>
 <cpl36>an a cappella group is a good thing.
 <veloc_leg 60c/1.9s 31.579>
 </sub17>

<sub37><2L>
 <t>00:01:24,631 --> 00:01:27,828</t>
 <cpl36>No. It's just since the beginning of <PROSEGL><SP_prep+SN>
 <cpl35>the conversation about finals club,
 <veloc_leg 71c/3.197s 22.208>
 </sub37>

<sub39><2L>
 <t>00:01:30,570 --> 00:01:33,733</t>
 <cpl42>There are really more people in China with <PROSEGL><SP_prep+SN>
 <cpl40>genius IQs than the entire population...
 <veloc_leg 82c/3.163s 25.925>
 </sub39>

<sub45><2L>
 <t>00:01:45,218 --> 00:01:46,685</t>
 <cpl22>who want to party with <PROSEGL><SP_prep+SN>
 <cpl22>the next Fed chairman?
 <veloc_leg 44c/1.467s 29.993>
 </sub45>

<sub46><2L>
 <t>00:01:46,753 --> 00:01:49,483</t>
 <cpl18>So you can see why <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl28>it's so important to get in.
 <veloc_leg 46c/2.73s 16.850>
 </sub46>

<sub52><2L>
 <t>00:02:03,236 --> 00:02:05,204</t>
 <cpl33>The ability to make money doesn't <PROSEGL><SV_verbo+SAdv>
 <cpl28>impress anybody around here.
 <veloc_leg 61c/1.968s 30.996>
 </sub52>

<sub73><2L>
 <t>00:02:50,750 --> 00:02:53,275</t>
 <cpl32>And there's a difference between <PROSEGL><SP_prep+SV>
 <cpl35>being obsessed and being motivated.
 <veloc_leg 67c/2.525s 26.535>
 </sub73>

<sub89><2L>
 <t>00:03:23,016 --> 00:03:25,280</t>
 <cpl30>and tell you I think you might <PROSEGL><SV_composto>
 <cpl36>want to be a little more supportive.
 <veloc_leg 66c/2.264s 29.152>
 </sub89>

<sub125><2L>

<t>00:04:43,463 --> 00:04:46,296</t>

<cpl35>I am sorry you are not sufficiently <PROSEGL><SN_subst+SAdj>

<cpl28>impressed with my education.

<veloc_leg 63c/2.833s 22.238>

</sub125>

<sub139><2L>

<t>00:08:22,615 --> 00:08:25,812</t>

<cpl28>or do you think it's because <PROSEGL><SUBORD_conj/pron_rel+oração>

<cpl25>all BU girls are bitches?

<veloc_leg 53c/3.197s 16.578>

</sub139>

<sub158><2L>

<t>00:09:29,749 --> 00:09:32,274</t>

<cpl31>than, say, choosing a number to <PROSEGL><SP_prep+SN>

<cpl32>represent each person's hotness,

<veloc_leg 63c/2.525s 24.950>

</sub158>

<sub166><2L>

<t>00:10:01,881 --> 00:10:04,873</t>

<cpl35>They keep everything open and allow <PROSEGL><SV_composto>

<cpl38>indexes in their Apache configuration.

<veloc_leg 73c/2.992s 24.398>

</sub166>

<sub181><2L>

<t>00:11:03,710 --> 00:11:06,873</t>

<cpl41>and I'm gonna go ahead and say they don't <PROSEGL><SV_verbo+SAdv>

<cpl42>have access to the main FAS user database,

<veloc_leg 83c/3.163s 26.241>

</sub181>

<sub190><2L>

<t>00:11:27,734 --> 00:11:30,328</t>

<cpl31>and if your search returns more <PROSEGL><SAdv_adv+adv>

<cpl39>than 20 matches, nothing gets returned.

<veloc_leg 70c/2.594s 26.985>

</sub190>

<sub211><2L>

<t>00:12:31,831 --> 00:12:33,765</t>

<cpl32>No, I need the algorithm you use <PROSEGL><SV_composto>

<cpl22>to rank chess players.

<veloc_leg 54c/1.934s 27.921>

</sub211>

<sub276><2L>

<t>00:16:21,394 --> 00:16:23,021</t>
 <cpl27>You weren't called in front <PROSEGL><SP_Prep+SN>
 <cpl28>of the Administrative Board?
 <veloc_leg 55c/1.627s 33.805>
 </sub276>

<sub282><2L>
 <t>00:16:33,906 --> 00:16:37,433</t>
 <cpl35>You think if I know she can make me <PROSEGL><SV_(verbo)+oblíquo+SV>
 <cpl48>look like a jerk, I'll be more likely to settle.
 <veloc_leg 83c/3.527s 23.533>
 </sub282>

<sub285><2L>
 <t>00:16:40,880 --> 00:16:44,543</t>
 <cpl39>And frankly, you did spend an awful lot <PROSEGL><SN_especif+SN>
 <cpl35>of time embarrassing Mr. Zuckerberg
 <veloc_leg 74c/3.663s 20.202>
 </sub285>

<sub309><2L>
 <t>00:18:17,176 --> 00:18:19,701</t>
 <cpl21>A three-hour low-rate <PROSEGL><SAdj_especif+SAdj>
 <cpl31>technical row before breakfast.
 <veloc_leg 52c/2.525s 20.594>
 </sub309>

<sub326><2L>
 <t>00:18:54,413 --> 00:18:57,746</t>
 <cpl35>Tyler spelled the usual way, and my <PROSEGL><SN_especif+SN>
 <cpl38>last name is the same as my brother's.
 <veloc_leg 73c/3.333s 21.902>
 </sub326>

<sub331><2L>
 <t>00:19:12,064 --> 00:19:14,589</t>
 <cpl24>You're also charged with <PROSEGL><SP_prep+SV>
 <cpl40>being in violation of university policy,
 <veloc_leg 64c/2.525s 25.347>
 </sub331>

<sub364><2L>
 <t>00:20:36,515 --> 00:20:40,042</t>
 <cpl30>How do you do this thing where <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl39>you manage to get all girls to hate us?
 <veloc_leg 69c/3.527s 19.563>
 </sub364>

<sub429><2L>
 <t>00:23:26,318 --> 00:23:28,718</t>

<cpl35>Harvard.edu is the most prestigious <PROSEGL><SN_subst+SAdj>
 <cpl35>e-mail address in the country, man.
 <veloc_leg 70c/2.4s 29.167>
 </sub429>

<sub434><2L>
 <t>00:23:37,429 --> 00:23:40,921</t>
 <cpl38>The main difference between what we're <PROSEGL><SV_composto>
 <cpl40>talking about and MySpace or Friendster,
 <veloc_leg 78c/3.492s 22.337>
 </sub434>

<sub441><2L>
 <t>00:23:52,277 --> 00:23:55,212</t>
 <cpl24>Women's groups are ready <PROSEGL><SV_composto>
 <cpl19>to declare a fatwa.
 <veloc_leg 43c/2.935s 14.651>
 </sub441>

<sub442><2L>
 <t>00:23:55,280 --> 00:23:57,714</t>
 <cpl30>But, you know, this could help <PROSEGL><SV_composto>
 <cpl24>rehabilitate your image.
 <veloc_leg 54c/2.434s 22.186>
 </sub442>

<sub447><2L>
 <t>00:24:10,362 --> 00:24:13,559</t>
 <cpl39>We would need you to build the site and <PROSEGL><COORD_conj+oração>
 <cpl44>write the code, and we'll provide all the...
 <veloc_leg 83c/3.197s 25.962>
 </sub447>

<sub501><2L>
 <t>00:26:33,005 --> 00:26:35,064</t>
 <cpl20>which has absolutely <PROSEGL><SAdv_adv+adv>
 <cpl33>nothing to do with the Caribbean.
 <veloc_leg 53c/2.059s 25.741>
 </sub501>

<sub504><2L>
 <t>00:26:49,454 --> 00:26:51,046</t>
 <cpl21>But it wasn't because <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl31>they saw pictures of hot girls.
 <veloc_leg 52c/1.592s 32.663>
 </sub504>

<sub513><2L>
 <t>00:27:08,573 --> 00:27:13,977</t>
 <cpl42>I'm talking about taking the entire social <PROSEGL><SN_subst+SAdj>

<cpl44>experience of college and putting it online.
 <veloc_leg 86c/5.404s 15.914>
 </sub513>

<sub527><2L>
 <t>00:27:46,078 --> 00:27:49,536</t>
 <cpl21>See, in a world where <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl32>social structure was everything,
 <veloc_leg 53c/3.458s 15.327>
 </sub527>

<sub531><2L>
 <t>00:27:56,688 --> 00:27:59,316</t>
 <cpl24>so I wondered why he was <PROSEGL><SV_composto>
 <cpl35>coming to me and not his roommates,
 <veloc_leg 59c/2.628s 22.451>
 </sub531>

<sub572><2L>
 <t>00:29:41,793 --> 00:29:43,886</t>
 <cpl20>But Mark didn't want <PROSEGL><SV_composto>
 <cpl25>to rehabilitate anything.
 <veloc_leg 45c/2.093s 21.500>
 </sub572>

<sub573><2L>
 <t>00:29:43,962 --> 00:29:46,760</t>
 <cpl31>With Facemash, he'd hacked into <PROSEGL><SP_prep+SN>
 <cpl22>the Harvard computers,
 <veloc_leg 53c/2.798s 18.942>
 </sub573>

<sub589><2L>
 <t>00:30:25,003 --> 00:30:27,528</t>
 <cpl32>"And it seems like it shouldn't <SV_verbo+SAdv>
 <cpl27>take too long to implement.
 <veloc_leg 59c/2.525s 23.366>
 </sub589>

<sub601><2L>
 <t>00:31:01,339 --> 00:31:04,069</t>
 <cpl32>but we do have to make sure that <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl42>we meet up before we all go off for break.
 <veloc_leg 74c/2.73s 27.106>
 </sub601>

<sub642><2L>
 <t>00:33:41,366 --> 00:33:45,325</t>
 <cpl42>To the best of your knowledge, had he even <PROSEGL><SV_verbo+SAdv>
 <cpl33>begun work on Harvard Connection?

<veloc_leg 75c/3.959s 18.944>
</sub642>

<sub675><2L>
<t>00:35:18,596 --> 00:35:22,327</t>
<cpl29>People don't walk around with <PROSEGL><SP_prep+SN>
<cpl27>a sign on them that says...
<veloc_leg 56c/3.731s 15.009>
</sub675>

<sub687><2L>
<t>00:36:30,702 --> 00:36:34,570</t>
<cpl35>And at its center, you know, that's <PROSEGL><SUBORD_conj/pron_rel+oração>
<cpl36>what The Facebook is gonna be about.
<veloc_leg 71c/3.868s 18.356>
</sub687>

<sub688><2L>
<t>00:36:34,639 --> 00:36:36,869</t>
<cpl31>People are gonna log on because <PROSEGL><SUBORD_conj/pron_rel+oração>
<cpl34>after all the cake and watermelon,
<veloc_leg 65c/2.23s 29.148>
</sub688>

<sub756><2L>
<t>00:40:05,349 --> 00:40:08,580</t>
<cpl40>"Everyone's been talking a lot about a <PROSEGL><SN_especific+SN>
<cpl46>universal facebook within Harvard,' he says."
<veloc_leg 86c/3.231s 26.617>
</sub756>

<sub764><2L>
<t>00:40:21,299 --> 00:40:25,030</t>
<cpl34>"over 650 students had registered <PROSEGL><SV_composto>
<cpl23>to use thefacebook.com.
<veloc_leg 57c/3.731s 15.277>
</sub764>

<sub766><2L>
<t>00:40:28,840 --> 00:40:32,037</t>
<cpl37>Yeah, Divya was just reading that 650 <PROSEGL><SN_especific+SN>
<cpl43>students signed up for it on the first day.
<veloc_leg 80c/3.197s 25.023>
</sub766>

<sub767><2L>
<t>00:40:32,143 --> 00:40:36,375</t>
<cpl38>God, if I was a drug dealer I couldn't <PROSEGL><SV_verbo+SAdv>
<cpl41>give free drugs to 650 people in one day.
<veloc_leg 79c/4.232s 18.667>

</sub767>

<sub771><2L>

<t>00:40:43,688 --> 00:40:45,383</t>

<cpl23>Well, you won't be able <PROSEGL><SV_composto>

<cpl30>to go on the website yourself.

<veloc_leg 53c/1.695s 31.268>

</sub771>

<sub788><2L>

<t>00:41:18,055 --> 00:41:21,456</t>

<cpl42>I don't think you need to school me in the <PROSEGL><SN_especif+SN>

<cpl45>importance of getting there first. Thank you.

<veloc_leg 87c/3.401s 25.581>

</sub788>

<sub821><2L>

<t>00:42:32,830 --> 00:42:37,199</t>

<cpl32>This is Harvard, where you don't <PROSEGL><SV_verbo+SAdv>

<cpl39>plant stories and you don't sue people.

<veloc_leg 71c/4.369s 16.251>

</sub821>

<sub831><2L>

<t>00:43:01,492 --> 00:43:03,392</t>

<cpl21>I had no idea whether <PROSEGL><SUBORD_conj/pron_rel+oração>

<cpl28>they came from money or not.

<veloc_leg 49c/1.9s 25.789>

</sub831>

<sub835><2L>

<t>00:43:10,334 --> 00:43:14,065</t>

<cpl38>Howard Winklevoss founded a firm whose

<PROSEGL><SUBORD_conj/pron_rel+oração>

<cpl39>assets are in the hundreds of millions.

<veloc_leg 77c/3.731s 20.638>

</sub835>

<sub846><2L>

<t>00:43:38,362 --> 00:43:41,422</t>

<cpl41>I went to my friend for the money because

<PROSEGL><SUBORD_conj/pron_rel+oração>

<cpl40>that's who I wanted to be partners with.

<veloc_leg 81c/3.06s 26.471>

</sub846>

<sub847><2L>

<t>00:43:41,532 --> 00:43:44,023</t>

<cpl28>Eduardo was the president of <PROSEGL><SP_prep+SN>

<cpl34>the Harvard Investors Association,

<veloc_leg 62c/2.491s 24.890>
</sub847>

<sub860><2L>
<t>00:44:13,364 --> 00:44:17,232</t>
<cpl33>And Mark was the biggest thing on <PROSEGL><SP_prep+SN>
<cpl42>a campus that included 19 Nobel laureates,
<veloc_leg 75c/3.868s 19.390>
</sub860>

<sub873><2L>
<t>00:44:52,870 --> 00:44:57,136</t>
<cpl33>That kit computer on the cover of <PROSEGL><SP_prep+SN>
<cpl42>that magazine had an 8080 processor in it.
<veloc_leg 75c/4.266s 17.581>
</sub873>

<sub874><2L>
<t>00:44:57,208 --> 00:45:01,304</t>
<cpl25>Unless you paid extra for <PROSEGL><SP_prep+SN>
<cpl38>a 1 K memory board, you got 256 bytes.
<veloc_leg 63c/4.096s 15.381>
</sub874>

<sub936><2L>
<t>00:47:26,423 --> 00:47:29,756</t>
<cpl21>Okay. But they manage <PROSEGL><SV_composto>
<cpl28>to make money selling pants.
<veloc_leg 49c/3.333s 14.701>
</sub936>

<sub945><2L>
<t>00:47:46,277 --> 00:47:48,643</t>
<cpl31>I find that to be a little more <PROSEGL><SAdv_adv+adv>
<cpl21>than mildly annoying.
<veloc_leg 52c/2.366s 21.978>
</sub945>

<sub962><2L>
<t>00:48:21,078 --> 00:48:22,705</t>
<cpl20>to everyone who ever <PROSEGL><SV_verbo+SAdv>
<cpl24>has built a chair, okay?
<veloc_leg 44c/1.627s 27.044>
</sub962>

<sub999><2L>
<t>00:50:02,546 --> 00:50:05,606</t>
<cpl36>I think if your clients wanna sit on <PROSEGL><SP_prep+SN>
<cpl38>my shoulders and call themselves tall,
<veloc_leg 74c/3.06s 24.183>

</sub999>

<sub1006><2L>

<t>00:50:21,165 --> 00:50:24,498</t>

<cpl32>are intellectually or creatively <PROSEGL><SAdj_especif+SAdj>

<cpl17>capable of doing.

<veloc_leg 49c/3.333s 14.701>

</sub1006>

<sub1031><2L>

<t>00:52:33,497 --> 00:52:35,522</t>

<cpl31>I don't know if you heard about <PROSEGL><SP_prep+SN>

<cpl28>this new website I launched.

<veloc_leg 59c/2.025s 29.136>

</sub1031>

<sub1100><2L>

<t>00:55:29,039 --> 00:55:31,564</t>

<cpl25>They are suing me because <PROSEGL><SUBORD_conj/pron_rel+oração>

<cpl34>for the first time in their lives,

<veloc_leg 59c/2.525s 23.366>

</sub1100>

<sub1108><2L>

<t>00:55:51,762 --> 00:55:55,061</t>

<cpl44>I wanna hire a lawyer to file for injunctive <PROSEGL><SN_subst+SAdj>

<cpl39>relief and get the site taken down now.

<veloc_leg 83c/3.299s 25.159>

</sub1108>

<sub1133><2L>

<t>00:56:50,420 --> 00:56:53,878</t>

<cpl21>I want a meeting with <PROSEGL><SP_prep+SN>

<cpl37>the goddamn president of this school.

<veloc_leg 58c/3.458s 16.773>

</sub1133>

<sub1167><2L>

<t>00:58:12,202 --> 00:58:16,832</t>

<cpl40>Well, I founded an Internet company that

<PROSEGL><SUBORD_conj/pron_rel+oração>

<cpl44>let folks download and share music for free.

<veloc_leg 84c/4.63s 18.143>

</sub1167>

<sub1181><2L>

<t>00:58:41,932 --> 00:58:43,593</t>

<cpl18>Broke. There's not <PROSEGL><SAdv_adv+adv>

<cpl29>a lot of money in free music,

<veloc_leg 47c/1.661s 28.296>

</sub1181>

<sub1182><2L>

<t>00:58:43,700 --> 00:58:47,033</t>

<cpl35>even less when you're being sued by

<PROSEGL><SUBORD_conj/pron_rel+oração>

<cpl40>everyone who's ever been to the Grammys.

<veloc_leg 75c/3.333s 22.502>

</sub1182>

<sub1204><2L>

<t>00:59:45,962 --> 00:59:48,760</t>

<cpl35>Yeah, I went to check my e-mail and <PROSEGL><COORD_conj+oração>

<cpl40>there's a website open on your computer.

<veloc_leg 75c/2.798s 26.805>

</sub1204>

<sub1225><2L>

<t>01:00:49,926 --> 01:00:52,690</t>

<cpl31>From the looks of it, they want <PROSEGL><SV_composto>

<cpl39>to sell me a Brooks Brothers franchise.

<veloc_leg 70c/2.764s 25.326>

</sub1225>

<sub1233><2L>

<t>01:01:05,408 --> 01:01:08,673</t>

<cpl33>Well, we came up with an idea for <PROSEGL><SP_prep+SN>

<cpl38>a website called Harvard Connection...

<veloc_leg 71c/3.265s 21.746>

</sub1233>

<sub1236><2L>

<t>01:01:13,550 --> 01:01:16,041</t>

<cpl31>And I'm asking what you want me <PROSEGL><SV_composto>

<cpl15>to do about it.

<veloc_leg 46c/2.491s 18.466>

</sub1236>

<sub1253><2L>

<t>01:02:00,764 --> 01:02:02,561</t>

<cpl27>Have you tried dealing with <PROSEGL><SP_prep+SN>

<cpl27>the other student directly?

<veloc_leg 54c/1.797s 30.050>

</sub1253>

<sub1257><2L>

<t>01:02:08,671 --> 01:02:10,502</t>

<cpl25>and the closest I've come <PROSEGL><SV_composto>

<cpl32>to dealing with him face-to-face

<veloc_leg 57c/1.831s 31.131>

</sub1257>

<sub1263><2L>

<t>01:02:23,486 --> 01:02:26,944</t>

<cpl36>You enter into a code of ethics with <PROSEGL><SP_prep+SN>

<cpl36>the university, not with each other.

<veloc_leg 72c/3.458s 20.821>

</sub1263>

<sub1264><2L>

<t>01:02:27,791 --> 01:02:30,988</t>

<cpl38>I'm sorry, President Summers, but what

<PROSEGL><SUBORD_conj/pron_rel+oração>

<cpl42>you just said makes no sense to me at all.

<veloc_leg 80c/3.197s 25.023>

</sub1264>

<sub1266><2L>

<t>01:02:32,328 --> 01:02:34,990</t>

<cpl28>What my brother means is, if <PROSEGL><SUBORD_conj/pron_rel+oração>

<cpl41>Mark Zuckerberg walked into our dorm room

<veloc_leg 69c/2.662s 25.920>

</sub1266>

<sub1280><2L>

<t>01:03:03,793 --> 01:03:05,226</t>

<cpl26>Yes. Everyone at Harvard's <PROSEGL><SV_composto>

<cpl20>inventing something.

<veloc_leg 46c/1.433s 32.100>

</sub1280>

<sub1281><2L>

<t>01:03:05,328 --> 01:03:09,958</t>

<cpl35>Harvard undergraduates believe that

<PROSEGL><SUBORD_conj/pron_rel+oração>

<cpl45>inventing a job is better than finding a job.

<veloc_leg 80c/4.63s 17.279>

</sub1281>

<sub1306><2L>

<t>01:04:14,731 --> 01:04:17,894</t>

<cpl26>Eduardo, spring break, you <PROSEGL><COORD_conj+oração>

<cpl43>and Mr. Zuckerberg took a trip to New York.

<veloc_leg 69c/3.163s 21.815>

</sub1306>

<sub1311><2L>

<t>01:04:29,412 --> 01:04:33,246</t>

<cpl39>At this point, your \$1,000 was the only <PROSEGL><SN_especif+SN>

<cpl41>money that had been put into the company?

<veloc_leg 80c/3.834s 20.866>
</sub1311>

<sub1332><2L>
<t>01:05:24,267 --> 01:05:25,666</t>
<cpl21>It was set up through <PROSEGL><SP_prep+SN>
<cpl26>my girlfriend at the time.
<veloc_leg 47c/1.399s 33.595>
</sub1332>

<sub1333><2L>
<t>01:05:25,835 --> 01:05:28,827</t>
<cpl27>Would you say that Mark was <PROSEGL><SV_composto>
<cpl27>excited about this meeting?
<veloc_leg 54c/2.992s 18.048>
</sub1333>

<sub1350><2L>
<t>01:06:05,675 --> 01:06:09,167</t>
<cpl41>Well, I didn't know him at all, but I had <PROSEGL><SV_composto>
<cpl35>done a search and I'd asked around.
<veloc_leg 76c/3.492s 21.764>
</sub1350>

<sub1378><2L>
<t>01:07:35,098 --> 01:07:38,329</t>
<cpl39>We tried to sell the company to pay the <PROSEGL><SP_prep+SN>
<cpl42>35 million they said we owed in royalties,
<veloc_leg 81c/3.231s 25.070>
</sub1378>

<sub1383><2L>
<t>01:07:50,747 --> 01:07:52,942</t>
<cpl22>And then he went on to <PROSEGL><SP_prep+SN>
<cpl28>his second business venture,
<veloc_leg 50c/2.195s 22.779>
</sub1383>

<sub1384><2L>
<t>01:07:53,483 --> 01:07:56,646</t>
<cpl32>which was an online Rolodex that <PROSEGL><SUBORD_conj/pron_rel+oração>
<cpl36>he got thrown out of by Case Equity.
<veloc_leg 68c/3.163s 21.499>
</sub1384>

<sub1387><2L>
<t>01:08:02,859 --> 01:08:05,794</t>
<cpl34>So let me tell you what happens to <PROSEGL><SP_prep+SN>
<cpl42>a 20-year-old at the top of a hot dot-com.
<veloc_leg 76c/2.935s 25.894>

</sub1387>

<sub1396><2L>

<t>01:08:26,949 --> 01:08:29,816</t>

<cpl39>And if somehow, somehow, you've managed <PROSEGL><SV_composto>

<cpl38>to live your life like the Dalai Lama,

<veloc_leg 77c/2.867s 26.857>

</sub1396>

<sub1451><2L>

<t>01:10:35,578 --> 01:10:37,944</t>

<cpl36>But we all know marlins don't really <PROSEGL><SV_verbo+SAdv>

<cpl26>weigh 3,000 pounds, right?

<veloc_leg 62c/2.366s 26.205>

</sub1451>

<sub1460><2L>

<t>01:10:57,333 --> 01:11:01,235</t>

<cpl34>But not before he made his biggest <PROSEGL><SAdj_especif+SAdj>

<cpl28>contribution to the company.

<veloc_leg 62c/3.902s 15.889>

</sub1460>

<sub1466><2L>

<t>01:11:22,024 --> 01:11:24,959</t>

<cpl28>That's gotta be some kind of <PROSEGL><SP_prep+SN>

<cpl30>land speed record for talking.

<veloc_leg 58c/2.935s 19.761>

</sub1466>

<sub1487><2L>

<t>01:12:15,311 --> 01:12:19,839</t>

<cpl39>You never did anything to embarrass the <PROSEGL><SP_prep+SN>

<cpl40>company or even seriously jeopardize it?

<veloc_leg 79c/4.528s 17.447>

</sub1487>

<sub1507><2L>

<t>01:13:12,468 --> 01:13:15,232</t>

<cpl22>I was having dinner in <PROSEGL><SP_prep+SN>

<cpl35>the Kirkland dining hall with Mark,

<veloc_leg 57c/2.764s 20.622>

</sub1507>

<sub1532><2L>

<t>01:14:06,856 --> 01:14:08,118</t>

<cpl27>I'm supposed to write about <PROSEGL><SP_prep+SN>

<cpl21>these four paintings.

<veloc_leg 48c/1.262s 38.035>

</sub1532>

<sub1551><2L>
 <t>01:14:54,403 --> 01:14:55,495</t>
 <cpl20>and we're gonna have <PROSEGL><SV_composto>
 <cpl22>to pay them something.
 <veloc_leg 42c/1.092s 38.462>
 </sub1551>

<sub1557><2L>
 <t>01:15:04,313 --> 01:15:07,476</t>
 <cpl30>Was it somewhere in the middle <PROSEGL><SAdv_adv+adv>
 <cpl32>of the Sean Parker Variety Hour?
 <veloc_leg 62c/3.163s 19.602>
 </sub1557>

<sub1566><2L>
 <t>01:15:26,769 --> 01:15:29,363</t>
 <cpl33>And as somebody who's just really <PROSEGL><SAdj_especific+SAdj>
 <cpl39>embarrassed the company in a bad way...
 <veloc_leg 72c/2.594s 27.756>
 </sub1566>

<sub1572><2L>
 <t>01:15:41,917 --> 01:15:45,751</t>
 <cpl35>You don't think it was strange that <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl38>he was followed by private detectives?
 <veloc_leg 73c/3.834s 19.040>
 </sub1572>

<sub1624><2L>
 <t>01:18:25,547 --> 01:18:30,075</t>
 <cpl36>But mostly I figured, how much could <PROSEGL><SV_verbo+SAdv>
 <cpl34>possibly go wrong in three months?
 <veloc_leg 70c/4.528s 15.459>
 </sub1624>

<sub1719><2L>
 <t>01:23:27,115 --> 01:23:30,778</t>
 <cpl19>And the water under <PROSEGL><SP_prep+SN>
 <cpl33>the Golden Gate is freezing cold.
 <veloc_leg 52c/3.663s 14.196>
 </sub1719>

<sub1732><2L>
 <t>01:24:07,422 --> 01:24:09,720</t>
 <cpl29>I think you should definitely <PROSEGL><SV_composto>
 <cpl22>come and live with us.
 <veloc_leg 51c/2.298s 22.193>
 </sub1732>

<sub1738><2L>
 <t>01:24:56,071 --> 01:25:00,531</t>
 <cpl28>At the barrier, Hollandia is <PROSEGL><SV_composto>
 <cpl41>leading Harvard University by one length.
 <veloc_leg 69c/4.46s 15.471>
 </sub1738>

<sub1755><2L>
 <t>01:26:39,574 --> 01:26:41,974</t>
 <cpl30>Mile-and-a-half races are more <PROSEGL><SN_subst+SAdj>
 <cpl37>commonly won by a boat length or two.
 <veloc_leg 67c/2.4s 27.917>
 </sub1755>

<sub1784><2L>
 <t>01:27:48,310 --> 01:27:50,403</t>
 <cpl25>And apparently Oxford and <PROSEGL><COORD_conj+oração>
 <cpl31>the London School of Economics.
 <veloc_leg 56c/2.093s 26.756>
 </sub1784>

<sub1799><2L>
 <t>01:28:27,349 --> 01:28:30,944</t>
 <cpl35>we tried the Ad Board, and we tried <PROSEGL><SV_composto>
 <cpl43>talking to the president of the university.
 <veloc_leg 78c/3.595s 21.697>
 </sub1799>

<sub1849><2L>
 <t>01:30:52,594 --> 01:30:55,222</t>
 <cpl31>Forget the Wall, tell him about <PROSEGL><SP_prep+SN>
 <cpl28>the meeting I've got set up.
 <veloc_leg 59c/2.628s 22.451>
 </sub1849>

<sub1944><2L>
 <t>01:35:40,848 --> 01:35:43,874</t>
 <cpl37>No, you didn't change it so you could <PROSEGL><SV_composto>
 <cpl32>screw those Silicon Valley sluts
 <veloc_leg 69c/3.026s 22.802>
 </sub1944>

<sub1946><2L>
 <t>01:35:45,253 --> 01:35:48,245</t>
 <cpl33>Not even remotely true, and I can <PROSEGL><SV_composto>
 <cpl41>promise you that the Silicon Valley sluts
 <veloc_leg 74c/2.992s 24.733>
 </sub1946>

<sub1947><2L>

<t>01:35:48,356 --> 01:35:51,257</t>
 <cpl24>don't care what anyone's <PROSEGL><SN_especif+SN><genitivo>
 <cpl35>Relationship Status is on Facebook.
 <veloc_leg 59c/2.901s 20.338>
 </sub1947>

<sub1965><2L>
 <t>01:36:21,456 --> 01:36:25,085</t>
 <cpl43>If the servers are down for even a day, our <PROSEGL><SN_especif+SN>
 <cpl44>entire reputation is irreversibly destroyed.
 <veloc_leg 87c/3.629s 23.974>
 </sub1965>

<sub1972><2L>
 <t>01:36:37,772 --> 01:36:39,399</t>
 <cpl22>I am not going back to <PROSEGL><SP_prep+SN>
 <cpl30>the Caribbean Night at A-E-Pi!
 <veloc_leg 52c/1.627s 31.961>
 </sub1972>

<sub2006><2L>
 <t>01:38:06,994 --> 01:38:08,894</t>
 <cpl30>The first two are common stock <PROSEGL><SN_subst+SAdj>
 <cpl20>purchase agreements,
 <veloc_leg 50c/1.9s 26.316>
 </sub2006>

<sub2020><2L>
 <t>01:38:41,362 --> 01:38:46,299</t>
 <cpl37>You should know that Mark has already <PROSEGL><SV_verbo+SAdv>
 <cpl41>taken his percentage from 60 down to 51 .
 <veloc_leg 78c/4.937s 15.799>
 </sub2020>

<sub2090><2L>
 <t>01:42:27,455 --> 01:42:30,720</t>
 <cpl30>Well, as you know, we had some <PROSEGL><SAdj_especif+SAdj>
 <cpl32>new investors that have come in.
 <veloc_leg 62c/3.265s 18.989>
 </sub2090>

<sub2101><2L>
 <t>01:43:07,261 --> 01:43:10,992</t>
 <cpl25>What was Mr. Zuckerberg's <PROSEGL><SN_especif+SN><genitivo>
 <cpl32>ownership share diluted down to?
 <veloc_leg 57c/3.731s 15.277>
 </sub2101>

<sub2103><2L>
 <t>01:43:12,466 --> 01:43:16,095</t>

<cpl40>What was Mr. Moskowitz's ownership share <PROSEGL><SV_composto>
 <cpl16>diluted down to?
 <veloc_leg 56c/3.629s 15.431>
 </sub2103>

<sub2105><2L>
 <t>01:43:17,605 --> 01:43:20,506</t>
 <cpl38>What was Sean Parker's ownership share <PROSEGL><SV_composto>
 <cpl16>diluted down to?
 <veloc_leg 54c/2.901s 18.614>
 </sub2105>

<sub2107><2L>
 <t>01:43:21,742 --> 01:43:24,404</t>
 <cpl38>What was Peter Thiel's ownership share <PROSEGL><SV_composto>
 <cpl16>diluted down to?
 <veloc_leg 54c/2.662s 20.285>
 </sub2107>

<sub2109><2L>
 <t>01:43:25,646 --> 01:43:28,638</t>
 <cpl33>And what was your ownership share <PROSEGL><SV_composto>
 <cpl16>diluted down to?
 <veloc_leg 49c/2.992s 16.377>
 </sub2109>

<sub2122><2L>
 <t>01:44:01,048 --> 01:44:05,451</t>
 <cpl36>Along with my hoodie and my fuck-you <PROSEGL><SN_subst+SAdj>
 <cpl38>flip-flops, you pretentious douchebag!
 <veloc_leg 74c/4.403s 16.807>
 </sub2122>

<sub2259><2L>
 <t>01:52:14,841 --> 01:52:18,140</t>
 <cpl33>and start working on a settlement <PROSEGL><SN_subst+SAdj>
 <cpl28>agreement to present to you.
 <veloc_leg 61c/3.299s 18.490>
 </sub2259>

ANEXO B – Problemas de segmentação linguística nas legendas em português

Filme: *Click* (2006)

<sub2><2L>
 <t>00:01:40,934 --> 00:01:42,834</t>
 <cpl11>Achamos que <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl26>estivesse dormindo, papai.
 <veloc_leg 37c/1.9s 19.474>
 </sub2>

<sub3><2L>
 <t>00:01:43,036 --> 00:01:46,301</t>
 <cpl21>Como posso dormir com <PROSEGL><SP_prep+SN>
 <cpl31>dois ladrões de Twinkie por aí?
 <veloc_leg 52c/3.265s 15.926>
 </sub3>

<sub7><2L>
 <t>00:02:09,863 --> 00:02:11,660</t>
 <cpl17>Querido, acho que <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl21>você abriu a garagem.
 <veloc_leg 38c/1.797s 21.146>
 </sub7>

<sub19><2L>
 <t>00:02:41,127 --> 00:02:44,858</t>
 <cpl27>Será que você terá tempo de <PROSEGL><SP_prep+SV>
 <cpl34>acabar a casa na árvore algum dia?
 <veloc_leg 61c/3.731s 16.350>
 </sub19>

<sub21><2L>
 <t>00:02:49,636 --> 00:02:54,471</t>
 <cpl34>Está bem. Eu quero acabar, mas tem <PROSEGL><SV_verbo+SAdv>
 <cpl38>muita coisa rolando no trabalho agora.
 <veloc_leg 72c/4.835s 14.891>
 </sub21>

<sub25><2L>
 <t>00:03:07,587 --> 00:03:11,387</t>
 <cpl27>Isso é uma coisa que só vão <PROSEGL><SV_composto>
 <cpl28>saber daqui a 10 ou 30 anos.
 <veloc_leg 55c/3.8s 14.474>
 </sub25>

<sub51><2L>
 <t>00:04:38,144 --> 00:04:40,977</t>
 <cpl19>é que os códigos em <PROSEGL><SP_prep+SN>
 <cpl25>Manhattan nos permitem...

<veloc_leg 44c/2.833s 15.531>
</sub51>

<sub53><2L>
<t>00:04:43,383 --> 00:04:46,511</t>
<cpl27>Um idiota em um Lamborghini <PROSEGL><SN_subst+SAdj>
<cpl36>vermelho estacionou na minha vaga...
<veloc_leg 63c/3.128s 20.141>
</sub53>

<sub58><2L>
<t>00:04:58,965 --> 00:05:01,263</t>
<cpl17>Mas é um dos meus <PROSEGL><SN_especif+SN>
<cpl20>melhores arquitetos.
<veloc_leg 37c/2.298s 16.101>
</sub58>

<sub67><2L>
<t>00:05:22,422 --> 00:05:27,223</t>
<cpl28>eu disse a mim mesmo: "Vamos <PROSEGL><SV_composto>
<cpl21>redefinir elegância".
<veloc_leg 49c/4.801s 10.206>
</sub67>

<sub76><2L>
<t>00:05:53,219 --> 00:05:56,484</t>
<cpl31>E ponham uma vala no chão, para <PROSEGL><SUBORD_conj/pron_rel+oração>
<cpl33>o concurso de camisetas molhadas.
<veloc_leg 64c/3.265s 19.602>
</sub76>

<sub79><2L>
<t>00:06:00,660 --> 00:06:04,027</t>
<cpl27>Vamos tirar o átrio e fazer <PROSEGL><SV_verbo+SAdv>
<cpl22>assim, criando a vala.
<veloc_leg 49c/3.367s 14.553>
</sub79>

<sub97><2L>
<t>00:06:56,149 --> 00:07:01,212</t>
<cpl25>Só preciso de alguém para <PROSEGL><SP_prep+SV>
<cpl30>fazer uma proposta. Intrigado?
<veloc_leg 55c/5.063s 10.863>
</sub97>

<sub120><2L>
<t>00:08:27,573 --> 00:08:30,974</t>
<cpl18>Ei, o melhor é que <PROSEGL><SUBORD_conj/pron_rel+oração>
<cpl20>papai viu seu final.
<veloc_leg 38c/3.401s 11.173>

</sub120>

<sub128><2L>

<t>00:08:47,260 --> 00:08:49,922</t>

<cpl19>Ben, você progrediu <PROSEGL><SV_verbo+SAdv>

<cpl22>muito lá na raia, viu?

<veloc_leg 41c/2.662s 15.402>

</sub128>

<sub140><2L>

<t>00:09:25,898 --> 00:09:29,390</t>

<cpl31>Não está. Só está fazendo o que <PROSEGL><SUBORD_conj/pron_rel+oração>

<cpl32>Michael faz. Isso me enlouquece.

<veloc_leg 63c/3.492s 18.041>

</sub140>

<sub141><2L>

<t>00:09:29,602 --> 00:09:33,265</t>

<cpl23>Tudo o que Michael anda <PROSEGL><SV_composto>

<cpl23>fazendo não é saudável.

<veloc_leg 46c/3.663s 12.558>

</sub141>

<sub143><2L>

<t>00:09:36,609 --> 00:09:39,203</t>

<cpl18>Estamos falando de <PROSEGL><SP_prep+SN>

<cpl20>um prédio exclusivo.

<veloc_leg 38c/2.594s 14.649>

</sub143>

<sub200><2L>

<t>00:12:24,243 --> 00:12:26,643</t>

<cpl23>O prazo é terça. Se não <PROSEGL><SV_verbo+SAdv>

<cpl21>acabar, estou fora...

<veloc_leg 44c/2.4s 18.333>

</sub200>

<sub203><2L>

<t>00:12:30,516 --> 00:12:32,780</t>

<cpl22>Mas em todas as minhas <PROSEGL><SN_especif+SN>

<cpl20>escolhas, em tudo...

<veloc_leg 42c/2.264s 18.551>

</sub203>

<sub208><2L>

<t>00:12:43,596 --> 00:12:45,359</t>

<cpl19>melhor do que a que <PROSEGL><SUBORD_conj/pron_rel+oração>

<cpl15>eu sonhava ter.

<veloc_leg 34c/1.763s 19.285>

</sub208>

<sub222><2L>
 <t>00:13:45,124 --> 00:13:48,059</t>
 <cpl17>Vocês têm aqueles <PROSEGL><SN_especif+SN>
 <cpl21>controles universais?
 <veloc_leg 38c/2.935s 12.947>
 </sub222>

<sub237><2L>
 <t>00:15:11,243 --> 00:15:13,541</t>
 <cpl23>Você tem algum controle <PROSEGL><SN_núcleo>
 <cpl22>remoto universal aqui?
 <veloc_leg 45c/2.298s 19.582>
 </sub237>

<sub277><2L>
 <t>00:17:47,066 --> 00:17:50,365</t>
 <cpl20>Sorte sua. Ainda não <PROSEGL><SV_verbo+SAdv>
 <cpl24>entrou no nosso sistema.
 <veloc_leg 44c/3.299s 13.337>
 </sub277>

<sub299><2L>
 <t>00:19:22,694 --> 00:19:25,322</t>
 <cpl21>Obrigado por entender <PROSEGL><SV_verbo+SAdv>
 <cpl14>na boa também.
 <veloc_leg 35c/2.628s 13.318>
 </sub299>

<sub339><2L>
 <t>00:22:53,138 --> 00:22:56,835</t>
 <cpl28>Está com ciúmes porque Donna <PROSEGL><COORD_conj+oração>
 <cpl25>e eu vamos à ginástica...
 <veloc_leg 53c/3.697s 14.336>
 </sub339>

<sub345><2L>
 <t>00:23:14,960 --> 00:23:19,863</t>
 <cpl25>Um marido a quem eu possa <PROSEGL><SV_composto>
 <cpl18>dar suco e beijar.
 <veloc_leg 43c/4.903s 8.770>
 </sub345>

<sub346><2L>
 <t>00:23:20,065 --> 00:23:23,831</t>
 <cpl26>Você já traiu três maridos <PROSEGL><SN_subst+SAdj>
 <cpl27>diferentes com seus irmãos.
 <veloc_leg 53c/3.766s 14.073>
 </sub346>

<sub350><2L>
 <t>00:23:33,145 --> 00:23:36,046</t>
 <cpl20>Não devia ter ido ao <PROSEGL><SP_prep+SN>
 <cpl28>programa do Montel Williams.
 <veloc_leg 48c/2.901s 16.546>
 </sub350>

<sub355><2L>
 <t>00:23:44,990 --> 00:23:47,458</t>
 <cpl13>Todos eles me <PROSEGL><SV_(verbo)+oblíquo+SV>
 <cpl27>abandonaram emocionalmente.
 <veloc_leg 40c/2.468s 16.207>
 </sub355>

<sub366><2L>
 <t>00:24:25,597 --> 00:24:30,534</t>
 <cpl28>Foi um lance na arquibancada <PROSEGL><SAdj_especific+SAdj>
 <cpl27>direita para Hideki Matsui.
 <veloc_leg 55c/4.937s 11.140>
 </sub366>

<sub372><2L>
 <t>00:24:49,888 --> 00:24:52,356</t>
 <cpl16>Ele late em mais <PROSEGL><SAdv_adv+adv>
 <cpl16>de seis línguas.
 <veloc_leg 32c/2.468s 12.966>
 </sub372>

<sub375><2L>
 <t>00:24:57,429 --> 00:25:00,296</t>
 <cpl19>Ei, pai, viu o novo <PROSEGL><SN_subst+SAdj>
 <cpl23>cachorro robô do Kevin?
 <veloc_leg 42c/2.867s 14.649>
 </sub375>

<sub377><2L>
 <t>00:25:02,434 --> 00:25:06,029</t>
 <cpl23>"Muito legal"? Ele vale <PROSEGL><SV_verbo+SAdv>
 <cpl22>mais do que seu carro.
 <veloc_leg 45c/3.595s 12.517>
 </sub377>

<sub379><2L>
 <t>00:25:20,285 --> 00:25:24,654</t>
 <cpl22>Está bem. Isso é algum <PROSEGL><SN_especific+SN>
 <cpl36>reality show? Você é o apresentador?
 <veloc_leg 58c/4.369s 13.275>
 </sub379>

<sub386><2L>

<t>00:25:45,810 --> 00:25:51,248</t>
 <cpl23>Você queria um controle <PROSEGL><SN_núcleo>
 <cpl25>remoto para seu universo.
 <veloc_leg 48c/5.438s 8.827>
 </sub386>

<sub435><2L>
 <t>00:28:20,365 --> 00:28:22,925</t>
 <cpl16>Acho que estamos <PROSEGL><SV_verbo+SAdv>
 <cpl20>depois da concepção.
 <veloc_leg 36c/2.56s 14.063>
 </sub435>

<sub437><2L>
 <t>00:28:25,670 --> 00:28:28,764</t>
 <cpl22>Não! Estamos na vagina <PROSEGL><SAdv_adv+adv>
 <cpl13>da minha mãe?
 <veloc_leg 35c/3.094s 11.312>
 </sub437>

<sub458><2L>
 <t>00:29:24,529 --> 00:29:26,156</t>
 <cpl24>Como se já não estivesse <PROSEGL><SV_verbo+SAdv>
 <cpl10>assustado!
 <veloc_leg 34c/1.627s 20.897>
 </sub458>

<sub459><2L>
 <t>00:29:26,364 --> 00:29:28,195</t>
 <cpl23>É simples. Pense em uma <PROSEGL><SN_especif+SN>
 <cpl19>época e um lugar...
 <veloc_leg 42c/1.831s 22.938>
 </sub459>

<sub465><2L>
 <t>00:29:43,948 --> 00:29:46,178</t>
 <cpl18>Sou eu, ali, com a <PROSEGL><SN_especif+SN>
 <cpl15>camisa do Jets.
 <veloc_leg 33c/2.23s 14.798>
 </sub465>

<sub469><2L>
 <t>00:29:55,360 --> 00:29:59,660</t>
 <cpl28>Michael, o jantar está quase <PROSEGL><SAdv_adv+adv>
 <cpl28>pronto. Convide seus amigos.
 <veloc_leg 56c/4.3s 13.023>
 </sub469>

<sub481><2L>
 <t>00:30:45,076 --> 00:30:48,068</t>

<cpl20>Tenho tanta fome que <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl13>comeria isso.
 <veloc_leg 33c/2.992s 11.029>
 </sub481>

<sub488><2L>
 <t>00:31:17,342 --> 00:31:20,937</t>
 <cpl21>Este é um dispositivo <PROSEGL><SN_subst+SAdj>
 <cpl26>poderoso. Use com cuidado.
 <veloc_leg 47c/3.595s 13.074>
 </sub488>

<sub495><2L>
 <t>00:31:42,934 --> 00:31:47,769</t>
 <cpl27>Não, Linda, acho que os pés <PROSEGL><SN_subst+SP>
 <cpl26>de Stacy são mais bonitos.
 <veloc_leg 53c/4.835s 10.962>
 </sub495>

<sub497><2L>
 <t>00:31:49,707 --> 00:31:52,767</t>
 <cpl24>Michael. Estou indo para <PROSEGL><SP_prep+SN>
 <cpl23>Hamptons daqui a pouco.
 <veloc_leg 47c/3.06s 15.359>
 </sub497>

<sub507><2L>
 <t>00:32:22,473 --> 00:32:24,964</t>
 <cpl20>Janine, uma amiga da <PROSEGL><SP_prep+SN>
 <cpl22>minha mulher, é linda.
 <veloc_leg 42c/2.491s 16.861>
 </sub507>

<sub517><2L>
 <t>00:33:03,881 --> 00:33:07,715</t>
 <cpl21>Está derrubando todas <PROSEGL><SN_especif+SN>
 <cpl34>as bolas, isso sim. Não pegou uma.
 <veloc_leg 55c/3.834s 14.345>
 </sub517>

<sub523><2L>
 <t>00:33:35,713 --> 00:33:38,147</t>
 <cpl24>Mas viu a bolsa dela, da <PROSEGL><SP_prep+SN>
 <cpl22>Kate Spade? Era linda.
 <veloc_leg 46c/2.434s 18.899>
 </sub523>

<sub553><2L>
 <t>00:35:19,117 --> 00:35:23,247</t>
 <cpl28>Meu pênis cresceu, depois de <PROSEGL><SP_prep+SN>

<cpl29>velho. Só para vocês saberem.
 <veloc_leg 57c/4.13s 13.801>
 </sub553>

<sub556><2L>
 <t>00:35:31,262 --> 00:35:34,322</t>
 <cpl13>Venha cá, vou <PROSEGL><SV_composto>
 <cpl21>refrescar seu hálito.
 <veloc_leg 34c/3.06s 11.111>
 </sub556>

<sub562><2L>
 <t>00:35:47,945 --> 00:35:51,676</t>
 <cpl26>Que bom. Pode ficar com os <PROSEGL><SN_especif+SN>
 <cpl33>meninos por 10 minutos para mim?
 <veloc_leg 59c/3.731s 15.813>
 </sub562>

<sub605><2L>
 <t>00:38:15,192 --> 00:38:18,684</t>
 <cpl29>Você disse que almoçaríamos e <PROSEGL><COORD_conj+oração>
 <cpl35>compraríamos o presente do seu pai.
 <veloc_leg 64c/3.492s 18.328>
 </sub605>

<sub617><2L>
 <t>00:38:57,935 --> 00:39:01,336</t>
 <cpl19>É simples. Deve ter <PROSEGL><SV_composto>
 <cpl27>ligado o piloto automático.
 <veloc_leg 46c/3.401s 13.525>
 </sub617>

<sub624><2L>
 <t>00:39:18,856 --> 00:39:21,086</t>
 <cpl19>O controle tira sua <PROSEGL><SN_especif+SN>
 <cpl13>mente dali...
 <veloc_leg 32c/2.23s 14.350>
 </sub624>

<sub625><2L>
 <t>00:39:21,292 --> 00:39:23,886</t>
 <cpl22>mas seu corpo continua <PROSEGL><SV_composto>
 <cpl25>ouvindo as coisas chatas.
 <veloc_leg 47c/2.594s 18.119>
 </sub625>

<sub642><2L>
 <t>00:40:10,274 --> 00:40:13,437</t>
 <cpl23>Quer ver o garanhão que <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl13>eu era? Quer?

<veloc_leg 36c/3.163s 11.382>
</sub642>

<sub644><2L>
<t>00:40:16,113 --> 00:40:18,707</t>
<cpl20>que eu tive antes de <PROSEGL><SP_prep+SN>
<cpl17>conhecer a Donna.
<veloc_leg 37c/2.594s 14.264>
</sub644>

<sub671><2L>
<t>00:42:09,026 --> 00:42:12,086</t>
<cpl23>Assistam aos desenhos e <PROSEGL><COORD_conj+oração>
<cpl31>divirtam-se com a mamãe, certo?
<veloc_leg 54c/3.06s 17.647>
</sub671>

<sub676><2L>
<t>00:42:27,845 --> 00:42:31,212</t>
<cpl26>Bom, logo vou ser sócio da <PROSEGL><SP_prep+SN>
<cpl37>empresa e farei muitos prédios novos.
<veloc_leg 63c/3.367s 18.711>
</sub676>

<sub677><2L>
<t>00:42:31,415 --> 00:42:34,077</t>
<cpl23>Se vocês tiverem alguma <PROSEGL><SN_especif+SN>
<cpl31>idéia legal, desenhem para mim.
<veloc_leg 54c/2.662s 20.285>
</sub677>

<sub693><2L>
<t>00:44:10,147 --> 00:44:14,379</t>
<cpl18>Certo. Vamos ficar <PROSEGL><SV_composto>
<cpl33>bronzeados como você sempre quis.
<veloc_leg 51c/4.232s 12.051>
</sub693>

<sub713><2L>
<t>00:45:18,782 --> 00:45:22,582</t>
<cpl23>Vou lhes dar um exemplo <PROSEGL><SN_subst+SAdj>
<cpl38>prático, que tenho certeza que eu fiz.
<veloc_leg 61c/3.8s 16.053>
</sub713>

<sub718><2L>
<t>00:45:44,842 --> 00:45:47,606</t>
<cpl18>Acontece quando um <PROSEGL><SN_especif+SN>
<cpl28>cara se aproxima de outro...
<veloc_leg 46c/2.764s 16.643>

</sub718>

<sub726><2L>

<t>00:46:32,122 --> 00:46:34,181</t>

<cpl16>Você sempre pode <PROSEGL><SV_composto>

<cpl18>contar com Matsui.

<veloc_leg 34c/2.059s 16.513>

</sub726>

<sub748><2L>

<t>00:47:45,596 --> 00:47:49,464</t>

<cpl25>Antes de pedirmos, estava <PROSEGL><SV_composto>

<cpl33>olhando a proposta. Querem saber?

<veloc_leg 58c/3.868s 14.995>

</sub748>

<sub770><2L>

<t>00:49:09,947 --> 00:49:12,347</t>

<cpl28>e terão o melhor, daqui para <PROSEGL><SAdv_adv+adv>

<cpl17>frente, está bem?

<veloc_leg 45c/2.4s 18.750>

</sub770>

<sub775><2L>

<t>00:49:24,795 --> 00:49:28,196</t>

<cpl23>Estas são as bicicletas <PROSEGL><SN_subst+SAdj>

<cpl22>mais legais da cidade.

<veloc_leg 45c/3.401s 13.231>

</sub775>

<sub776><2L>

<t>00:49:28,398 --> 00:49:30,832</t>

<cpl17>Kevin O'Doyle vai <PROSEGL><SV_composto>

<cpl16>sujar as calças.

<veloc_leg 33c/2.434s 13.558>

</sub776>

<sub778><2L>

<t>00:49:34,604 --> 00:49:38,472</t>

<cpl24>Não é uma bicicleta, nem <PROSEGL><COORD_conj+oração>

<cpl36>um chapéu de caubói, mas vai gostar.

<veloc_leg 60c/3.868s 15.512>

</sub778>

<sub786><2L>

<t>00:50:00,564 --> 00:50:03,795</t>

<cpl19>Eu sabia! Sabia que <PROSEGL><SUBORD_conj/pron_rel+oração>

<cpl24>enlouqueceria! Amo você!

<veloc_leg 43c/3.231s 13.309>

</sub786>

<sub790><2L>
 <t>00:50:12,509 --> 00:50:16,946</t>
 <cpl18>Amor! A que estava <PROSEGL><SV_composto>
 <cpl28>tocando quando nos beijamos.
 <veloc_leg 46c/4.437s 10.367>
 </sub790>

<sub835><2L>
 <t>00:53:01,011 --> 00:53:03,809</t>
 <cpl29>Eu disse: "Consiga a conta do <PROSEGL><SP_prep+SN>
 <cpl28>Watsuhita e será promovido".
 <veloc_leg 57c/2.798s 20.372>
 </sub835>

<sub842><2L>
 <t>00:53:31,007 --> 00:53:35,307</t>
 <cpl28>Esqueci de dizer que saí com <PROSEGL><SP_prep+SN>
 <cpl34>sua amiga Janine no fim de semana.
 <veloc_leg 62c/4.3s 14.419>
 </sub842>

<sub852><2L>
 <t>00:54:22,993 --> 00:54:24,961</t>
 <cpl18>Você já pensou nas <PROSEGL><SP_prep+SN>
 <cpl16>conseqüências...
 <veloc_leg 34c/1.968s 17.276>
 </sub852>

<sub854><2L>
 <t>00:54:27,998 --> 00:54:30,364</t>
 <cpl20>Sim. Acha que eu não <PROSEGL><SV_verbo+SAdv>
 <cpl16>devo fazer isso?
 <veloc_leg 36c/2.366s 15.216>
 </sub854>

<sub862><2L>
 <t>00:54:54,224 --> 00:54:57,853</t>
 <cpl22>Ele vive atrás do pote <PROSEGL><SN_subst+SAdj>
 <cpl29>de ouro, no fim do arco-íris.
 <veloc_leg 51c/3.629s 14.053>
 </sub862>

<sub870><2L>
 <t>00:55:41,971 --> 00:55:44,496</t>
 <cpl20>Está bem. Querem que <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl17>eu veja? Vou ver.
 <veloc_leg 37c/2.525s 14.653>
 </sub870>

<sub880><2L>
 <t>00:56:14,037 --> 00:56:17,336</t>
 <cpl28>É puxar o saco do chefe, não <PROSEGL><SV_verbo+SAdv>
 <cpl39>ousar, fazê-lo ganhar muito dinheiro...
 <veloc_leg 67c/3.299s 20.309>
 </sub880>

<sub906><2L>
 <t>00:57:43,893 --> 00:57:46,157</t>
 <cpl23>Infelizmente, houve uma <PROSEGL><SN_especif+SN>
 <cpl26>confusão no trabalho dele.
 <veloc_leg 49c/2.264s 21.643>
 </sub906>

<sub917><2L>
 <t>00:58:16,092 --> 00:58:19,084</t>
 <cpl17>o maior sócio que <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl25>eu poderia sonhar em ter.
 <veloc_leg 42c/2.992s 14.037>
 </sub917>

<sub932><2L>
 <t>00:59:00,837 --> 00:59:04,898</t>
 <cpl32>O incidente do fim de semana que <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl27>ela passou com meu irmão...
 <veloc_leg 59c/4.061s 14.528>
 </sub932>

<sub934><2L>
 <t>00:59:07,110 --> 00:59:10,045</t>
 <cpl21>Mas o Dr. Bergman nos <PROSEGL><SV_(verbo)+oblíquo+SV>
 <cpl22>ajudou, não foi, amor?
 <veloc_leg 43c/2.935s 14.651>
 </sub934>

<sub946><2L>
 <t>00:59:42,111 --> 00:59:45,308</t>
 <cpl19>E agora ela prefere <PROSEGL><SV_composto>
 <cpl28>ser chamada de Alan. Lembra?
 <veloc_leg 47c/3.197s 14.701>
 </sub946>

<sub968><2L>
 <t>01:00:43,806 --> 01:00:45,865</t>
 <cpl21>Que tal assistirmos a <PROSEGL><SP_prep+SN>
 <cpl19>Dragon Tales juntos?
 <veloc_leg 40c/2.059s 19.427>
 </sub968>

<sub976><2L>

<t>01:01:11,100 --> 01:01:13,796</t>
 <cpl27>Não grite com Peanut porque <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl26>descobriu seu esconderijo.
 <veloc_leg 53c/2.696s 19.659>
 </sub976>

<sub986><2L>
 <t>01:01:46,335 --> 01:01:49,031</t>
 <cpl26>ele tentou trepar na minha <PROSEGL><SN_especif+SN>
 <cpl23>perna, e eu não deixei.
 <veloc_leg 49c/2.696s 18.175>
 </sub986>

<sub997><2L>
 <t>01:02:21,671 --> 01:02:25,607</t>
 <cpl28>Meu Deus, o pato não demorou <PROSEGL><SV_verbo+SAdv>
 <cpl30>nada para esquecer o Sundance.
 <veloc_leg 58c/3.936s 14.736>
 </sub997>

<sub1016><2L>
 <t>01:03:39,582 --> 01:03:42,608</t>
 <cpl27>Avancei o sexo uma vez, não <PROSEGL><SV_verbo+SAdv>
 <cpl33>quero que seja assim para sempre.
 <veloc_leg 60c/3.026s 19.828>
 </sub1016>

<sub1039><2L>
 <t>01:05:02,331 --> 01:05:06,529</t>
 <cpl13>Eu avisei que <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl23>não poderia devolvê-lo.
 <veloc_leg 36c/4.198s 8.576>
 </sub1039>

<sub1045><2L>
 <t>01:05:27,924 --> 01:05:30,722</t>
 <cpl21>E se eu avançar, quem <PROSEGL><COORD_conj+oração>
 <cpl22>sabe por quanto tempo?
 <veloc_leg 43c/2.798s 15.368>
 </sub1045>

<sub1049><2L>
 <t>01:05:40,336 --> 01:05:41,963</t>
 <cpl19>Quero que saiba que <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl22>estou com um problema.
 <veloc_leg 41c/1.627s 25.200>
 </sub1049>

<sub1051><2L>
 <t>01:05:45,174 --> 01:05:47,267</t>

<cpl22>Vou trabalhar e tentar <PROSEGL><SV_composto>
 <cpl14>resolver isso.
 <veloc_leg 36c/2.093s 17.200>
 </sub1051>

<sub1053><2L>
 <t>01:05:49,312 --> 01:05:51,371</t>
 <cpl20>Querida, não podemos <PROSEGL><SV_composto>
 <cpl15>brigar. Acabou.
 <veloc_leg 35c/2.059s 16.999>
 </sub1053>

<sub1058><2L>
 <t>01:06:02,591 --> 01:06:05,890</t>
 <cpl26>Quero ficar com você minha <PROSEGL><SN_especif+SN>
 <cpl31>vida toda, por isso vou embora.
 <veloc_leg 57c/3.299s 17.278>
 </sub1058>

<sub1070><2L>
 <t>01:07:21,537 --> 01:07:24,995</t>
 <cpl26>John, acho que não devemos <PROSEGL><SV_composto>
 <cpl17>gastar energia...
 <veloc_leg 43c/3.458s 12.435>
 </sub1070>

<sub1071><2L>
 <t>01:07:25,207 --> 01:07:28,472</t>
 <cpl25>puxando o saco e querendo <PROSEGL><SV_composto>
 <cpl30>saber quem tem a melhor roupa.
 <veloc_leg 55c/3.265s 16.845>
 </sub1071>

<sub1074><2L>
 <t>01:07:36,719 --> 01:07:38,983</t>
 <cpl14>Esta é a maior <PROSEGL><SN_subst+SAdj>
 <cpl22>maluquice que já ouvi.
 <veloc_leg 36c/2.264s 15.901>
 </sub1074>

<sub1080><2L>
 <t>01:07:57,873 --> 01:08:02,867</t>
 <cpl29>tem uma coisa, ou melhor, uma <PROSEGL><SN_especif+SN>
 <cpl27>pessoa de quem quero falar.
 <veloc_leg 56c/4.994s 11.213>
 </sub1080>

<sub1082><2L>
 <t>01:08:06,515 --> 01:08:09,143</t>
 <cpl30>Michael, Janine e eu decidimos <PROSEGL><SV_composto>

<cpl19>sair da correria...
 <veloc_leg 49c/2.628s 18.645>
 </sub1082>

<sub1092><2L>
 <t>01:08:33,075 --> 01:08:35,373</t>
 <cpl23>se um dia você acabasse <PROSEGL><SV_composto>
 <cpl18>virando nosso CEO.
 <veloc_leg 41c/2.298s 17.842>
 </sub1092>

<sub1124><2L>
 <t>01:11:04,927 --> 01:11:08,693</t>
 <cpl28>Veja só quem fala! O capitão <PROSEGL><SN_núcleo>
 <cpl25>Twinkie, do navio Gordão!
 <veloc_leg 53c/3.766s 14.073>
 </sub1124>

<sub1135><2L>
 <t>01:11:35,224 --> 01:11:38,387</t>
 <cpl22>Meu namorado, o garoto <PROSEGL><SN_subst+SAdj>
 <cpl21>mais lindo da escola.
 <veloc_leg 43c/3.163s 13.595>
 </sub1135>

<sub1147><2L>
 <t>01:12:17,533 --> 01:12:19,296</t>
 <cpl16>Donna e eu íamos <PROSEGL><SV_composto>
 <cpl23>tomar chocolate quente.
 <veloc_leg 39c/1.763s 22.121>
 </sub1147>

<sub1152><2L>
 <t>01:12:29,445 --> 01:12:32,209</t>
 <cpl23>Mas não devia mexer com <PROSEGL><SP_prep+SN>
 <cpl18>a mulher de outro!
 <veloc_leg 41c/2.764s 14.834>
 </sub1152>

<sub1164><2L>
 <t>01:13:06,115 --> 01:13:08,515</t>
 <cpl20>Não. Não, não queria <PROSEGL><SV_composto>
 <cpl17>perder tudo isso.
 <veloc_leg 37c/2.4s 15.417>
 </sub1164>

<sub1187><2L>
 <t>01:14:45,547 --> 01:14:47,640</t>
 <cpl23>É, se ele não o tivesse <PROSEGL><SV_composto>
 <cpl12>derrubado...

<veloc_leg 35c/2.093s 16.722>
</sub1187>

<sub1191><2L>
<t>01:14:55,123 --> 01:14:57,353</t>
<cpl20>Foi a única pessoa a <PROSEGL><SP_prep+SV>
<cpl26>engordar na quimioterapia.
<veloc_leg 46c/2.23s 20.628>
</sub1191>

<sub1194><2L>
<t>01:15:01,230 --> 01:15:03,221</t>
<cpl21>Achei que já tivessem <PROSEGL><SV_composto>
<cpl21>parado com a morfina.
<veloc_leg 42c/1.991s 21.095>
</sub1194>

<sub1210><2L>
<t>01:15:52,881 --> 01:15:55,509</t>
<cpl15>O que houve com <PROSEGL><SP_prep+SN>
<cpl23>"para sempre", querida?
<veloc_leg 38c/2.628s 14.460>
</sub1210>

<sub1212><2L>
<t>01:15:58,053 --> 01:16:01,386</t>
<cpl23>Michael, estou tentando <PROSEGL><SV_composto>
<cpl38>ter uma relação civilizada com você...
<veloc_leg 61c/3.333s 18.302>
</sub1212>

<sub1219><2L>
<t>01:16:17,873 --> 01:16:20,671</t>
<cpl20>Por favor, aprenda a <PROSEGL><SP_prep+SV>
<cpl25>aceitar isso. Preciso ir.
<veloc_leg 45c/2.798s 16.083>
</sub1219>

<sub1223><2L>
<t>01:16:32,454 --> 01:16:35,082</t>
<cpl20>Construindo um mundo <PROSEGL><SN_subst+SAdj>
<cpl16>melhor para nós.
<veloc_leg 36c/2.628s 13.699>
</sub1223>

<sub1230><2L>
<t>01:16:58,113 --> 01:17:02,550</t>
<cpl23>Sei, minha bunda. Vamos <PROSEGL><SV_composto>
<cpl24>conversar na minha sala?
<veloc_leg 47c/4.437s 10.593>

</sub1230>

<sub1236><2L>

<t>01:17:24,539 --> 01:17:28,031</t>

<cpl34>Não. Tenho malhado com Bill, cinco <PROSEGL><SAdv_adv+adv>

<cpl30>vezes por semana, como sempre.

<veloc_leg 64c/3.492s 18.328>

</sub1236>

<sub1245><2L>

<t>01:18:02,077 --> 01:18:05,945</t>

<cpl20>Temos de passar mais <PROSEGL><SAdv_adv+adv>

<cpl24>tempo juntos. Todos nós.

<veloc_leg 44c/3.868s 11.375>

</sub1245>

<sub1250><2L>

<t>01:18:18,126 --> 01:18:20,560</t>

<cpl18>Sabia que você não <PROSEGL><SV_verbo+SAdv>

<cpl18>estava legal hoje.

<veloc_leg 36c/2.434s 14.790>

</sub1250>

<sub1277><2L>

<t>01:20:09,571 --> 01:20:11,903</t>

<cpl17>Pode me levar até <PROSEGL><SN_subst+SP>

<cpl30>a última vez em que nos vimos?

<veloc_leg 47c/2.332s 20.154>

</sub1277>

<sub1293><2L>

<t>01:21:02,357 --> 01:21:05,417</t>

<cpl23>Você, eu e Ben podíamos <PROSEGL><SV_composto>

<cpl21>sair à noite, só nós.

<veloc_leg 44c/3.06s 14.379>

</sub1293>

<sub1301><2L>

<t>01:21:27,782 --> 01:21:32,116</t>

<cpl21>Se for comigo, eu lhe <PROSEGL><SV_(verbo)+oblíquo+SV>

<cpl25>mostro o truque da moeda.

<veloc_leg 46c/4.334s 10.614>

</sub1301>

<sub1305><2L>

<t>01:21:38,460 --> 01:21:40,928</t>

<cpl22>Eu sei como faz aquele <PROSEGL><SN_especif+SN>

<cpl28>truque idiota. Sempre soube.

<veloc_leg 50c/2.468s 20.259>

</sub1305>

<sub1352><2L>
 <t>01:25:59,521 --> 01:26:03,184</t>
 <cpl26>Michael, você não me beija <PROSEGL><SV_verbo+SAdv>
 <cpl27>assim desde que era garoto.
 <veloc_leg 53c/3.663s 14.469>
 </sub1352>

<sub1355><2L>
 <t>01:26:10,765 --> 01:26:14,565</t>
 <cpl16>Seu pai adoraria <PROSEGL><SV_composto>
 <cpl20>estar aqui, Michael.
 <veloc_leg 36c/3.8s 9.474>
 </sub1355>

<sub1400><2L>
 <t>01:29:35,336 --> 01:29:39,466</t>
 <cpl14>Por que me fez <PROSEGL><SV_composto>
 <cpl26>desperdiçar uma vida toda?
 <veloc_leg 40c/4.13s 9.685>
 </sub1400>

<sub1410><2L>
 <t>01:30:06,567 --> 01:30:09,502</t>
 <cpl17>Melhor, agora que <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl17>vocês estão aqui.
 <veloc_leg 34c/2.935s 11.584>
 </sub1410>

<sub1425><2L>
 <t>01:30:55,716 --> 01:30:58,651</t>
 <cpl24>ela entende. Sabe que os <PROSEGL><SN_especif+SN>
 <cpl30>negócios são mais importantes.
 <veloc_leg 54c/2.935s 18.399>
 </sub1425>

<sub1426><2L>
 <t>01:30:58,853 --> 01:31:01,617</t>
 <cpl24>Desculpem, mas vocês vão <PROSEGL><SV_composto>
 <cpl36>ter que sair. Ele precisa descansar.
 <veloc_leg 60c/2.764s 21.708>
 </sub1426>

<sub1431><2L>
 <t>01:31:25,746 --> 01:31:28,271</t>
 <cpl20>Amigo, você não pode <PROSEGL><SV_composto>
 <cpl11>sair daqui.
 <veloc_leg 31c/2.525s 12.277>
 </sub1431>

<sub1439><2L>
 <t>01:32:01,048 --> 01:32:04,040</t>
 <cpl25>Michael, isso não precisa <PROSEGL><SV_composto>
 <cpl13>acabar agora.
 <veloc_leg 38c/2.992s 12.701>
 </sub1439>

<sub1473><2L>
 <t>01:35:41,535 --> 01:35:43,503</t>
 <cpl13>Bem, você não <PROSEGL><SV_verbo+SAdv>
 <cpl17>é lá muito jovem.
 <veloc_leg 30c/1.968s 15.244>
 </sub1473>

<sub1475><2L>
 <t>01:35:45,740 --> 01:35:49,039</t>
 <cpl26>Não, não, não. A vida está <PROSEGL><SV_verbo+SAdv>
 <cpl17>apenas começando.
 <veloc_leg 43c/3.299s 13.034>
 </sub1475>

<sub1488><2L>
 <t>01:36:36,690 --> 01:36:38,419</t>
 <cpl15>Um mágico nunca <PROSEGL><SV_verbo+SAdv>
 <cpl21>revela seus segredos.
 <veloc_leg 36c/1.729s 20.821>
 </sub1488>

<sub1489><2L>
 <t>01:36:38,626 --> 01:36:41,094</t>
 <cpl15>Mas se você nos <PROSEGL><SV_(verbo)+oblíquo+SV>
 <cpl28>convidar mais para jantar...
 <veloc_leg 43c/2.468s 17.423>
 </sub1489>

<sub1518><2L>
 <t>01:38:25,199 --> 01:38:27,929</t>
 <cpl29>Um dia, você vai ser a garota <PROSEGL><SN_subst+SAdj>
 <cpl20>mais linda do mundo.
 <veloc_leg 49c/2.73s 17.949>
 </sub1518>

<sub1520><2L>
 <t>01:38:29,737 --> 01:38:31,637</t>
 <cpl19>Então, amanhã vamos <PROSEGL><SV_composto>
 <cpl17>aprender cálculo.
 <veloc_leg 36c/1.9s 18.947>
 </sub1520>

<sub1527><2L>

<t>01:38:45,252 --> 01:38:49,655</t>
 <cpl28>O tempo passa voando para um <PROSEGL><SN_especif+SN>
 <cpl38>cachorro. Divirta-se com o lance real.
 <veloc_leg 66c/4.403s 14.990>
 </sub1527>

<sub1539><2L>
 <t>01:39:50,584 --> 01:39:53,052</t>
 <cpl13>Vou pôr minha <PROSEGL><SN_especif+SN>
 <cpl23>fantasia de Pocahontas.
 <veloc_leg 36c/2.468s 14.587>
 </sub1539>

<sub1548><2L>
 <t>01:41:04,758 --> 01:41:07,283</t>
 <cpl19>Quem quer fazer uma <PROSEGL><SN_especif+SN>
 <cpl22>briga de travesseiros?
 <veloc_leg 41c/2.525s 16.238>
 </sub1548>

Filme: *Exodus* (2015)

<sub9><2L>
 <t>00:02:12,873 --> 00:02:17,310</t>
 <cpl40>É verdade que o exército hitita está com <PROSEGL><SP_prep+SN>
 <cpl40>16.000 soldados nos arredores de Kadesh.
 <veloc_leg 80c/4.437s 18.030>
 </sub9>

<sub11><2L>
 <t>00:02:19,680 --> 00:02:22,706</t>
 <cpl16>Os hititas estão <PROSEGL><SV_composto>
 <cpl28>tentando cruzar a fronteira.
 <veloc_leg 44c/3.026s 14.541>
 </sub11>

<sub106><2L>
 <t>00:13:44,264 --> 00:13:46,698</t>
 <cpl17>travando uma luta <PROSEGL><SN_subst+SAdj>
 <cpl22>desesperada pela vida.
 <veloc_leg 39c/2.434s 16.023>
 </sub106>

<sub195><2L>
 <t>00:19:50,063 --> 00:19:52,657</t>
 <cpl15>Suas tribos são <PROSEGL><SV_verbo+SAdv>
 <cpl29>mais ferozes que os egípcios.
 <veloc_leg 44c/2.594s 16.962>

</sub195>

<sub316><2L>

<t>00:28:44,096 --> 00:28:47,623</t>

<cpl17>É uma das ironias <PROSEGL><SN_subst+SAdj>

<cpl25>mais deprimentes da vida,

<veloc_leg 42c/3.527s 11.908>

</sub316>

<sub382><2L>

<t>00:33:28,881 --> 00:33:31,748</t>

<cpl15>como eu poderia <PROSEGL><SV_composto>

<cpl25>acreditar nessa história.

<veloc_leg 40c/2.867s 13.952>

</sub382>

<sub430><2L>

<t>00:35:40,779 --> 00:35:43,680</t>

<cpl21>Ou talvez você queira <PROSEGL><SV_composto>

<cpl21>responder a pergunta.

<veloc_leg 42c/2.901s 14.478>

</sub430>

<sub536><2L>

<t>00:50:02,540 --> 00:50:05,475</t>

<cpl21>e confio no que ainda <PROSEGL><SAdv_adv+adv>

<cpl8>não sei.

<veloc_leg 29c/2.935s 9.881>

</sub536>

<sub587><2L>

<t>00:53:52,903 --> 00:53:56,862</t>

<cpl14>É ruim crescer <PROSEGL><SV_composto>

<cpl24>acreditando em si mesmo?

<veloc_leg 38c/3.959s 9.598>

</sub587>

<sub609><2L>

<t>00:56:55,986 --> 00:56:59,444</t>

<cpl22>Você deveria ver o que <PROSEGL><SUBORD_conj/pron_rel+oração>

<cpl36>está acontecendo com seu povo agora.

<veloc_leg 58c/3.458s 16.773>

</sub609>

<sub637><2L>

<t>00:58:38,188 --> 00:58:41,180</t>

<cpl10>Eu não fui <PROSEGL><SV_verbo+SAdv>

<cpl31>completamente sincero com você.

<veloc_leg 41c/2.992s 13.703>

</sub637>

<sub689><2L>
 <t>01:04:24,735 --> 01:04:27,795</t>
 <cpl12>É um caminho <PROSEGL><SN_subst+SAdj>
 <cpl20>estreito e perigoso.
 <veloc_leg 32c/3.06s 10.458>
 </sub689>

<sub711><2L>
 <t>01:08:39,656 --> 01:08:42,147</t>
 <cpl16>Foi por isso que <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl40>só mandou dois assassinos para me matar?
 <veloc_leg 56c/2.491s 22.481>
 </sub711>

<sub744><2L>
 <t>01:10:07,344 --> 01:10:11,906</t>
 <cpl15>Eu não esperava <PROSEGL><SV_composto>
 <cpl21>ouvir um simples sim,
 <veloc_leg 36c/4.562s 7.891>
 </sub744>

<sub745><2L>
 <t>01:10:12,115 --> 01:10:15,710</t>
 <cpl13>mas não quero <PROSEGL><SV_composto>
 <cpl21>ouvir um simples não.
 <veloc_leg 34c/3.595s 9.458>
 </sub745>

<sub937><2L>
 <t>01:33:43,784 --> 01:33:47,276</t>
 <cpl29>mais de 10 milhões de cúbitos <PROSEGL><SN_núcleo>
 <cpl31>quadrados de colheita perdidos.
 <veloc_leg 60c/3.492s 17.182>
 </sub937>

<sub1094><2L>
 <t>01:53:50,723 --> 01:53:54,124</t>
 <cpl23>Por que nós não podemos <PROSEGL><SV_composto>
 <cpl21>fazer aquele caminho?
 <veloc_leg 44c/3.401s 12.937>
 </sub1094>

Filme: *Flightplan* (2004)

<sub7><2L>
 <t>05:44:17,844 --> 05:44:21,940</t>
 <cpl20>avid? Podemos sentar <PROSEGL><SV_verbo+SAdv>
 <cpl19>um pouco no jardim?

<veloc_leg 39c/4.096s 9.521>
</sub7>

<sub23><2L>
<t>05:47:58,265 --> 05:48:01,723</t>
<cpl32>- Vai ficar tudo bem, doçura. Eu <PROSEGL>
<cpl26>prometo. - Estou com medo.
<veloc_leg 58c/3.458s 16.773>
</sub23>

<sub54><2L>
<t>05:51:48,328 --> 05:51:51,388</t>
<cpl32>Pode ficar com ele até entrarmos <PROSEGL><SV_verbo+SAdv>
<cpl20>no avião, combinado?
<veloc_leg 52c/3.06s 16.993>
</sub54>

<sub71><2L>
<t>05:53:30,396 --> 05:53:32,591</t>
<cpl36>Não, não, não. Nada de laptop. Agora <PROSEGL><SAdv_adv+adv>
<cpl4>não.
<veloc_leg 40c/2.195s 18.223>
</sub71>

<sub83><2L>
<t>05:54:35,395 --> 05:54:38,023</t>
<cpl34>Pessoas a dois aviões daqui também <PROSEGL><SV_verbo+SAdv>
<cpl8>ouviram.
<veloc_leg 42c/2.628s 15.982>
</sub83>

<sub93><2L>
<t>05:55:31,584 --> 05:55:35,884</t>
<cpl37>Precisam tirar o gelo, senão o ar não <PROSEGL><SV_verbo+SAdv>
<cpl25>passa direito pelas asas.
<veloc_leg 62c/4.3s 14.419>
</sub93>

<sub109><2L>
<t>05:59:45,738 --> 05:59:49,606</t>
<cpl37>É claro Que a 36 mil pés, não dá para <PROSEGL><SP_prep+SV>
<cpl22>sair do cinema, não é?
<veloc_leg 59c/3.868s 15.253>
</sub109>

<sub121><2L>
<t>06:01:16,896 --> 06:01:20,593</t>
<cpl27>E deve estar com um ursinho <PROSEGL><SN_subst+SAdj>
<cpl26>de pelúcia de um braço só.
<veloc_leg 53c/3.697s 14.336>

</sub121>

<sub137><2L>

<t>06:03:33,666 --> 06:03:37,762</t>

<cpl32>Parece Que a aeronave é grande o <PROSEGL><SAdv_adv+adv>

<cpl36>bastante para se perder uma criança.

<veloc_leg 68c/4.096s 16.602>

</sub137>

<sub139><2L>

<t>06:03:42,174 --> 06:03:45,109</t>

<cpl31>Se ela estiver perto de você ou <PROSEGL><COORD_conj+oração>

<cpl28>brincando com seus filhos...

<veloc_leg 59c/2.935s 20.102>

</sub139>

<sub146><2L>

<t>06:04:26,085 --> 06:04:28,315</t>

<cpl24>Estou certa Que ela está <PROSEGL><SV_verbo+SAdv>

<cpl18>perfeitamente bem.

<veloc_leg 42c/2.23s 18.834>

</sub146>

<sub201><2L>

<t>06:08:28,494 --> 06:08:30,962</t>

<cpl37>Fomos os dois primeiros passageiros a <PROSEGL><SP_prep+SV>

<cpl16>entrar no avião.

<veloc_leg 53c/2.468s 21.475>

</sub201>

<sub206><2L>

<t>06:08:45,410 --> 06:08:50,245</t>

<cpl40>A lista diz Que o 26-A está livre. Fiz a <PROSEGL><SN_especif+SN>

<cpl21>contagem. VerifiQuei.

<veloc_leg 61c/4.835s 12.616>

</sub206>

<sub260><2L>

<t>06:12:57,796 --> 06:13:01,527</t>

<cpl34>Isso faz sentido? Podia pensar num <PROSEGL><SP_prep+SN>

<cpl26>motivo para eu fazer isso?

<veloc_leg 60c/3.731s 16.081>

</sub260>

<sub265><2L>

<t>06:13:16,748 --> 06:13:20,878</t>

<cpl38>Eu digo Que está incorreta. Então deve <PROSEGL><SV_composto>

<cpl32>checar o avião de cabo a rabo...

<veloc_leg 70c/4.13s 16.949>

</sub265>

<sub300><2L>
 <t>06:16:38,016 --> 06:16:41,281</t>
 <cpl38>E mulheres com crianças imaginárias se <PROSEGL><SV_(verbo)+oblíquo+SV>
 <cpl11>Qualificam.
 <veloc_leg 49c/3.265s 15.008>
 </sub300>

<sub333><2L>
 <t>06:19:20,545 --> 06:19:24,174</t>
 <cpl31>- O botão é muito alto para uma <PROSEGL><SN_especif+SN>
 <cpl34>criança... - Bem como o bagageiro,
 <veloc_leg 65c/3.629s 17.911>
 </sub333>

<sub345><2L>
 <t>06:19:58,917 --> 06:20:03,684</t>
 <cpl32>Sr. Carson, sou responsável pela <PROSEGL><SP_prep+SN>
 <cpl31>segurança de todos neste avião.
 <veloc_leg 63c/4.767s 13.216>
 </sub345>

<sub346><2L>
 <t>06:20:03,755 --> 06:20:07,782</t>
 <cpl39>Mesmo os delirantes. E não deixarei Que
 <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl34>nenhum passageiro saia do lugar...
 <veloc_leg 73c/4.027s 18.128>
 </sub346>

<sub356><2L>
 <t>06:20:39,791 --> 06:20:42,487</t>
 <cpl31>Que não estão recebendo nenhuma <PROSEGL><SN_especif+SN>
 <cpl21>atenção no momento...
 <veloc_leg 52c/2.696s 19.288>
 </sub356>

<sub359><2L>
 <t>06:20:49,033 --> 06:20:52,400</t>
 <cpl35>Se acha Que devemos fazer mais para <PROSEGL><SP_prep+SV>
 <cpl31>atender às suas necessidades...
 <veloc_leg 66c/3.367s 19.602>
 </sub359>

<sub360><2L>
 <t>06:20:52,470 --> 06:20:56,497</t>
 <cpl35> sugiro reclamar com o Atendimento a <PROSEGL><SP_prep+SN>
 <cpl26>Clientes Quando pousarmos.
 <veloc_leg 61c/4.027s 15.148>
 </sub360>

<sub372><2L>
 <t>06:21:41,886 --> 06:21:45,185</t>
 <cpl36>Eu vi os dois olhando para dentro do <PROSEGL><SP_prep+SN>
 <cpl22>Quarto da minha filha!
 <veloc_leg 58c/3.299s 17.581>
 </sub372>

<sub386><2L>
 <t>06:22:34,272 --> 06:22:37,833</t>
 <cpl33>mas eu estive no Hilton ontem, em <PROSEGL><SP_prep+SN>
 <cpl23>Tiergarten, a negócios.
 <veloc_leg 56c/3.561s 15.726>
 </sub386>

<sub422><2L>
 <t>06:25:16,367 --> 06:25:19,336</t>
 <cpl39>Não me importa se não sou politicamente <PROSEGL><SAdj_especif+SAdj>
 <cpl8>correta!
 <veloc_leg 47c/2.969s 15.830>
 </sub422>

<sub423><2L>
 <t>06:25:19,403 --> 06:25:23,669</t>
 <cpl39>Ótimo! Vamos prender os dois! Assim Que
 <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl35>você responder a mais uma pergunta.
 <veloc_leg 74c/4.266s 17.346>
 </sub423>

<sub428><2L>
 <t>06:25:44,662 --> 06:25:47,597</t>
 <cpl33>Ainda não sei, mas acho Que estão <PROSEGL><SV_composto>
 <cpl21>seQuestrando o avião.
 <veloc_leg 54c/2.935s 18.399>
 </sub428>

<sub430><2L>
 <t>06:25:49,600 --> 06:25:53,161</t>
 <cpl28>A morte de seu marido começa <PROSEGL><SV_composto>
 <cpl30>a fazer mais sentido para mim.
 <veloc_leg 58c/3.561s 16.288>
 </sub430>

<sub449><2L>
 <t>06:27:34,572 --> 06:27:38,838</t>
 <cpl39>Srta. Pratt, de acordo com o diretor do <PROSEGL><SP_prep+SN>
 <cpl13>necrotério...
 <veloc_leg 52c/4.266s 12.189>
 </sub449>

<sub450><2L>
 <t>06:27:38,909 --> 06:27:42,709</t>
 <cpl33>sua filha Julia, morreu de lesões <PROSEGL><SN_subst+SAdj>
 <cpl11>internas...
 <veloc_leg 44c/3.8s 11.579>
 </sub450>

<sub464><2L>
 <t>06:28:37,902 --> 06:28:42,703</t>
 <cpl33>independentemente da seriedade da <PROSEGL><SP_prep+SN>
 <cpl30>situação, ponha isso em risco.
 <veloc_leg 63c/4.801s 13.122>
 </sub464>

<sub477><2L>
 <t>06:30:06,967 --> 06:30:10,767</t>
 <cpl26>O capitão achou Que talvez <PROSEGL><SV_verbo+SAdv>
 <cpl20>Queira falar comigo.
 <veloc_leg 46c/3.8s 12.105>
 </sub477>

<sub516><2L>
 <t>06:33:16,991 --> 06:33:21,951</t>
 <cpl38>Você pode imaginar... seu marido e sua <PROSEGL><SN_especif+SN>
 <cpl16>filha numa ilha,
 <veloc_leg 54c/4.96s 10.887>
 </sub516>

<sub518><2L>
 <t>06:33:25,266 --> 06:33:30,898</t>
 <cpl35>Eles sempre estarão lá. E você pode <PROSEGL><SV_composto>
 <cpl35>nadar até eles sempre Que precisar.
 <veloc_leg 70c/5.632s 12.429>
 </sub518>

<sub569><2L>
 <t>06:41:31,685 --> 06:41:36,054</t>
 <cpl34>É o Que pilotos fazem Quando há um <PROSEGL><SN_especif+SN>
 <cpl33>passageiro descontrolado a bordo.
 <veloc_leg 67c/4.369s 15.335>
 </sub569>

<sub574><2L>
 <t>06:41:50,370 --> 06:41:55,137</t>
 <cpl35>Senhoras e senhores, como devem ter <PROSEGL><SV_composto>
 <cpl33>notado, começamos a descer agora.
 <veloc_leg 68c/4.767s 14.265>
 </sub574>

<sub587><2L>
 <t>06:42:40,620 --> 06:42:43,589</t>
 <cpl35>Por favor. Olhe, não há mais tantos <PROSEGL><SAdv_adv+adv>
 <cpl22>lugares para procurar.
 <veloc_leg 57c/2.969s 19.198>
 </sub587>

<sub616><2L>
 <t>06:48:30,236 --> 06:48:34,070</t>
 <cpl33>Querem o avião vazio. Primeiro os <PROSEGL><SN_especif+SN>
 <cpl33>passageiros, depois a tripulação.
 <veloc_leg 66c/3.834s 17.214>
 </sub616>

<sub619><2L>
 <t>06:48:39,946 --> 06:48:44,144</t>
 <cpl37>Ela está perturbada. Usou o caixão do <PROSEGL><SN_especif+SN>
 <cpl33>marido para trazer os explosivos.
 <veloc_leg 70c/4.198s 16.675>
 </sub619>

<sub627><2L>
 <t>06:49:11,077 --> 06:49:15,275</t>
 <cpl34>Nunca soube de um seQuestrador Que
 <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl33>deixasse as pessoas saírem antes.
 <veloc_leg 67c/4.198s 15.960>
 </sub627>

<sub628><2L>
 <t>06:49:15,348 --> 06:49:19,944</t>
 <cpl31>Minha recomendação é Que peça à <PROSEGL><SP_prep+SN>
 <cpl35>empresa para fazer a transferência.
 <veloc_leg 66c/4.596s 14.360>
 </sub628>

<sub636><2L>
 <t>06:49:56,089 --> 06:50:00,219</t>
 <cpl35>Que ela está nos instrumentos e Que
 <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl36>estou envolvida. Quando você saiu...
 <veloc_leg 71c/4.13s 17.191>
 </sub636>

<sub640><2L>
 <t>06:50:10,236 --> 06:50:13,831</t>
 <cpl34>Meio avião acha Que há terroristas <PROSEGL><SN_subst+SAdj>
 <cpl15>árabes a bordo.
 <veloc_leg 49c/3.595s 13.630>
 </sub640>

<sub664><2L>
 <t>06:51:51,804 --> 06:51:55,331</t>
 <cpl35>Já os informei Que temos uma pessoa <PROSEGL><SN_subst+SAdj>
 <cpl21>desaparecida a bordo.
 <veloc_leg 56c/3.527s 15.878>
 </sub664>

<sub675><2L>
 <t>06:53:42,915 --> 06:53:45,850</t>
 <cpl28>O suspeito será trazido pelo <PROSEGL><SP_prep+SN>
 <cpl16>policial do vôo.
 <veloc_leg 44c/2.935s 14.991>
 </sub675>

<sub697><2L>
 <t>06:55:48,808 --> 06:55:52,710</t>
 <cpl29>Srta. Pratt, seu dinheiro foi <PROSEGL><SV_composto>
 <cpl23>transferido como pediu.
 <veloc_leg 52c/3.902s 13.326>
 </sub697>

<sub705><2L>
 <t>06:56:54,240 --> 06:56:57,232</t>
 <cpl24>Você sai do avião Quando <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl25>eu disser para você sair.
 <veloc_leg 49c/2.992s 16.377>
 </sub705>

<sub730><2L>
 <t>07:00:57,883 --> 07:01:02,217</t>
 <cpl31>Precisávamos de um seQuestrador <PROSEGL><SN_subst+SAdj>
 <cpl36>convincente. Que conhecesse o avião.
 <veloc_leg 67c/4.334s 15.459>
 </sub730>

<sub756><2L>
 <t>07:07:39,484 --> 07:07:43,682</t>
 <cpl37>Central, aQui é FBI Zona 1 . Todas as <PROSEGL><SN_especif+SN>
 <cpl33>unidades em posição no perímetro.
 <veloc_leg 70c/4.198s 16.675>
 </sub756>

Filme: *Madagascar* (2005)

<sub2><2L>
 <t>00:01:30,249 --> 00:01:34,015</t>
 <cpl30>Alex, não me interrompa quando <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl24>estou sonhando acordado.

<veloc_leg 54c/3.766s 14.339>
</sub2>

<sub18><2L>
<t>00:02:20,065 --> 00:02:23,125</t>
<cpl20>Ainda não chegou nas <PROSEGL><SP_prep+SN>
<cpl27>prateleiras. Dê uma olhada.
<veloc_leg 47c/3.06s 15.359>
</sub18>

<sub28><2L>
<t>00:02:50,729 --> 00:02:55,564</t>
<cpl31>"Paro aqui, trote ali, como uma <PROSEGL><SN_especif+SN>
<cpl33>graminha, ando de volta para cá".
<veloc_leg 64c/4.835s 13.237>
</sub28>

<sub75><2L>
<t>00:06:07,959 --> 00:06:11,019</t>
<cpl24>Quero que fique bonito e <PROSEGL><COORD_conj+oração>
<cpl15>fofo, pracinha.
<veloc_leg 39c/3.06s 12.745>
</sub75>

<sub85><2L>
<t>00:06:42,127 --> 00:06:45,654</t>
<cpl33>Não esqueçam: não castrem os seus <PROSEGL><SN_especif+SN>
<cpl21>animais de estimação!
<veloc_leg 54c/3.527s 15.310>
</sub85>

<sub99><2L>
<t>00:07:34,946 --> 00:07:38,245</t>
<cpl19>Isso é fruto de uma <PROSEGL><SN_especif+SN>
<cpl21>estranha conspiração.
<veloc_leg 40c/3.299s 12.125>
</sub99>

<sub224><2L>
<t>00:14:54,652 --> 00:14:57,314</t>
<cpl21>Voltaremos de manhã e <PROSEGL><COORD_conj+oração>
<cpl21>ninguêm nunca saberá.
<veloc_leg 42c/2.662s 15.778>
</sub224>

<sub230><2L>
<t>00:15:17,342 --> 00:15:20,903</t>
<cpl35>Eu também. Tenho que poupar a minha <PROSEGL><SN_especif+SN>
<cpl16>voz para amanhã.
<veloc_leg 51c/3.561s 14.322>

</sub230>

<sub234><2L>

<t>00:15:40,565 --> 00:15:43,363</t>

<cpl28>Esqueceram de desligar o som <PROSEGL><SN_subst+SAdj>

<cpl17>ambiente de novo.

<veloc_leg 45c/2.798s 16.083>

</sub234>

<sub254><2L>

<t>00:17:00,912 --> 00:17:03,574</t>

<cpl24>Há quanto tempo ele está <PROSEGL><SV_composto>

<cpl18>trabalhando nisso?

<veloc_leg 42c/2.662s 15.778>

</sub254>

<sub268><2L>

<t>00:17:44,355 --> 00:17:48,052</t>

<cpl31>Iráo transferir o Marty. Não se <PROSEGL><SV_(verbo)+oblíquo+SV>

<cpl28>morde a mão que te alimenta.

<veloc_leg 59c/3.697s 15.959>

</sub268>

<sub275><2L>

<t>00:18:26,330 --> 00:18:28,798</t>

<cpl33>Um de nós tem que ficar aqui caso <PROSEGL><COORD_conj+oração>

<cpl10>ele volte.

<veloc_leg 43c/2.468s 17.423>

</sub275>

<sub290><2L>

<t>00:19:16,814 --> 00:19:20,341</t>

<cpl29>É uma missão urgente para que <PROSEGL><SUBORD_conj/pron_rel+oração>

<cpl28>Marty não jogue a vida fora.

<veloc_leg 57c/3.527s 16.161>

</sub290>

<sub296><2L>

<t>00:19:37,401 --> 00:19:41,201</t>

<cpl35>E ele tem riscas pretas e brancas e <PROSEGL><COORD_conj+oração>

<cpl30>tudo mais e depois cantamos...

<veloc_leg 65c/3.8s 17.105>

</sub296>

<sub309><2L>

<t>00:20:27,051 --> 00:20:30,509</t>

<cpl34>Falaram "Grand Central Station" ou <PROSEGL><COORD_conj+oração>

<cpl20>foi só "enganation"?

<veloc_leg 54c/3.458s 15.616>

</sub309>

<sub361><2L>
 <t>00:23:20,224 --> 00:23:24,354</t>
 <cpl32>Vamos levar o meu amiguinho para <PROSEGL><SP_prep+SN>
 <cpl32>casa e esquecer o que aconteceu.
 <veloc_leg 64c/4.13s 15.496>
 </sub361>

<sub372><2L>
 <t>00:24:30,536 --> 00:24:33,130</t>
 <cpl31>Serão devolvidos ao seu habitat <PROSEGL><SN_subst+SAdj>
 <cpl10>natural...
 <veloc_leg 41c/2.594s 15.806>
 </sub372>

<sub401><2L>
 <t>00:26:03,462 --> 00:26:06,056</t>
 <cpl26>Tenho consulta marcada com <PROSEGL><SP_prep+SN>
 <cpl23>o dr. Goldberg às 5:00.
 <veloc_leg 49c/2.594s 18.890>
 </sub401>

<sub403><2L>
 <t>00:26:08,867 --> 00:26:11,461</t>
 <cpl31>Nenhum outro zôo bancaria o meu <PROSEGL><SN_especif+SN>
 <cpl13>plano médico.
 <veloc_leg 44c/2.594s 16.962>
 </sub403>

<sub411><2L>
 <t>00:26:32,591 --> 00:26:37,119</t>
 <cpl31>"Não sei quem sou. Tenho que me <PROSEGL><SV_(verbo)+oblíquo+SV>
 <cpl19>achar na natureza".
 <veloc_leg 50c/4.528s 11.042>
 </sub411>

<sub421><2L>
 <t>00:27:03,221 --> 00:27:07,123</t>
 <cpl31>Enviar para RESERVA SELVAGEM DO <PROSEGL><SP_prep+SN>
 <cpl14>QUÊNIA, ÀFRICA
 <veloc_leg 45c/3.902s 11.533>
 </sub421>

<sub437><2L>
 <t>00:28:24,736 --> 00:28:28,968</t>
 <cpl23>Se o amavam é só porque <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl24>não o conheciam direito.
 <veloc_leg 47c/4.232s 11.106>
 </sub437>

<sub528><2L>
 <t>00:34:23,595 --> 00:34:26,655</t>
 <cpl38>Estou dizendo, pode ser o Zoológico de <PROSEGL><SP_prep+SN>
 <cpl10>San Diego.
 <veloc_leg 48c/3.06s 15.686>
 </sub528>

<sub536><2L>
 <t>00:34:44,949 --> 00:34:48,043</t>
 <cpl28>Depois enterro você, clono e <PROSEGL><COORD_conj+oração>
 <cpl21>mato todos os clones.
 <veloc_leg 49c/3.094s 15.837>
 </sub536>

<sub541><2L>
 <t>00:35:02,000 --> 00:35:05,959</t>
 <cpl30>Agora terei que competir com a <PROSEGL><SN_especif+SN>
 <cpl33>Shamu e seu sorrisinho malicioso.
 <veloc_leg 63c/3.959s 15.913>
 </sub541>

<sub550><2L>
 <t>00:35:26,958 --> 00:35:30,086</t>
 <cpl30>Você tem que ser um pouco mais <PROSEGL><SAdj_especif+SAdj>
 <cpl11>compreen...
 <veloc_leg 41c/3.128s 13.107>
 </sub550>

<sub556><2L>
 <t>00:35:47,078 --> 00:35:52,277</t>
 <cpl31>Primeiro dizem: "Temos um ótimo <PROSEGL><SN_subst+SAdj>
 <cpl31>espaço para soltar os animais".
 <veloc_leg 62c/5.199s 11.925>
 </sub556>

<sub590><2L>
 <t>00:37:53,137 --> 00:37:56,470</t>
 <cpl34>Não sei, mas ele está perdendo uma <PROSEGL><SN_especif+SN>
 <cpl11>superfesta.
 <veloc_leg 45c/3.333s 13.501>
 </sub590>

<sub597><2L>
 <t>00:38:18,396 --> 00:38:21,593</t>
 <cpl36>Muito obrigado, galera. Obrigado por <PROSEGL><SP_prep+SV>
 <cpl27>me esperarem. Gostei muito!
 <veloc_leg 63c/3.197s 19.706>
 </sub597>

<sub598><2L>
 <t>00:38:24,569 --> 00:38:28,505</t>

<cpl31>Chegamos de Nova York e estamos <PROSEGL><SV_composto>
 <cpl24>procurando o supervisor.
 <veloc_leg 55c/3.936s 13.974>
 </sub598>

<sub614><2L>
 <t>00:39:21,526 --> 00:39:25,986</t>
 <cpl30>Vieram nos matar, pegar nossas <PROSEGL><SN_especif+SN>
 <cpl35>mulheres e nossos metais preciosos.
 <veloc_leg 65c/4.46s 14.574>
 </sub614>

<sub638><2L>
 <t>00:40:50,281 --> 00:40:52,909</t>
 <cpl20>Aquele leão grande e <PROSEGL><COORD_conj+oração>
 <cpl18>mau assustou você?
 <veloc_leg 38c/2.628s 14.460>
 </sub638>

<sub645><2L>
 <t>00:41:11,969 --> 00:41:14,870</t>
 <cpl26>Tem algo naquele do cabelo <PROSEGL><SN_subst+SAdj>
 <cpl9>esquisito
 <veloc_leg 35c/2.901s 12.065>
 </sub645>

<sub649><2L>
 <t>00:41:34,425 --> 00:41:38,020</t>
 <cpl31>Sua Excelência, o nosso ilustre <PROSEGL><SN_subst+SAdj>
 <cpl18>Rei Julien XIII...
 <veloc_leg 49c/3.595s 13.630>
 </sub649>

<sub650><2L>
 <t>00:41:38,095 --> 00:41:41,724</t>
 <cpl28>...autoproclamado senhor dos <PROSEGL><SP_prep+SN>
 <cpl31>lêmures, etc. "Hurra", pessoal!
 <veloc_leg 59c/3.629s 16.258>
 </sub650>

<sub657><2L>
 <t>00:42:11,562 --> 00:42:15,896</t>
 <cpl31>Somos muito gratos por mandarem <PROSEGL><SV_verbo+SAdv>
 <cpl17>embora as foosas.
 <veloc_leg 48c/4.334s 11.075>
 </sub657>

<sub661><2L>
 <t>00:42:24,175 --> 00:42:27,167</t>
 <cpl34>Parece bom. Só queremos saber onde <PROSEGL><SV_verbo+SAdv>

<cpl18>estão os homens...
 <veloc_leg 52c/2.992s 17.380>
 </sub661>

<sub671><2L>
 <t>00:42:56,240 --> 00:42:59,334</t>
 <cpl30>Só temos que perguntar a esses <PROSEGL><SN_especif+SN>
 <cpl29>"Bozos" onde estão os homens.
 <veloc_leg 59c/3.094s 19.069>
 </sub671>

<sub684><2L>
 <t>00:43:38,516 --> 00:43:43,010</t>
 <cpl38>...natureza onde se mora numa cabana e <PROSEGL><COORD_conj+oração>
 <cpl23>se limpa com uma folha?
 <veloc_leg 61c/4.494s 13.574>
 </sub684>

<sub720><2L>
 <t>00:45:39,870 --> 00:45:43,772</t>
 <cpl33>Abusou do desejo de aniversário e <PROSEGL><COORD_conj+oração>
 <cpl27>trouxe azar para todos nós.
 <veloc_leg 60c/3.902s 15.377>
 </sub720>

<sub728><2L>
 <t>00:46:04,562 --> 00:46:09,499</t>
 <cpl27>Este é o seu lado da ilha e <PROSEGL><COORD_conj+oração>
 <cpl28>este é o nosso lado da ilha.
 <veloc_leg 55c/4.937s 11.140>
 </sub728>

<sub729><2L>
 <t>00:46:09,567 --> 00:46:13,799</t>
 <cpl34>Este é o lado ruim, onde você pode <PROSEGL><SV_composto>
 <cpl31>pular feito um cavalo mágico...
 <veloc_leg 65c/4.232s 15.359>
 </sub729>

<sub734><2L>
 <t>00:46:24,615 --> 00:46:28,949</t>
 <cpl32>Tudo bem. Vocês têm o seu lado e <PROSEGL><COORD_conj+oração>
 <cpl15>eu terei o meu.
 <veloc_leg 47c/4.334s 10.844>
 </sub734>

<sub769><2L>
 <t>00:48:31,742 --> 00:48:34,006</t>
 <cpl32>Não se preocupe, os gatos sempre <PROSEGL><SV_verbo+SAdv>
 <cpl10>caem de...

<veloc_leg 42c/2.264s 18.551>
</sub769>

<sub785><2L>
<t>00:49:23,394 --> 00:49:28,093</t>
<cpl35>...estamos curiosos sobre os nossos <PROSEGL><SN_especif+SN>
<cpl30>convidados, os gigantes de NY.
<veloc_leg 65c/4.699s 13.833>
</sub785>

<sub787><2L>
<t>00:49:30,134 --> 00:49:34,093</t>
<cpl32>Gostei deles. Gostei deles antes <PROSEGL><SAdv_adv+adv>
<cpl21>mesmo de os conhecer.
<veloc_leg 53c/3.959s 13.387>
</sub787>

<sub797><2L>
<t>00:50:02,866 --> 00:50:07,030</t>
<cpl26>Esperem aí, elas não estão <PROSEGL><SV_(verbo)+oblíquo+SV>
<cpl28>nos atacando neste instante.
<veloc_leg 54c/4.164s 12.968>
</sub797>

<sub808><2L>
<t>00:50:42,840 --> 00:50:46,105</t>
<cpl21>Talvez tivéssemos que <PROSEGL><SUBORD_conj/pron_rel+oração>
<cpl15>temê-lo também.
<veloc_leg 36c/3.265s 11.026>
</sub808>

<sub812><2L>
<t>00:50:56,587 --> 00:51:00,318</t>
<cpl29>Seu comentário malicioso será <PROSEGL><SV_composto>
<cpl22>excluído do relatório.
<veloc_leg 51c/3.731s 13.669>
</sub812>

<sub816><2L>
<t>00:51:10,701 --> 00:51:16,367</t>
<cpl30>...queremos ter certeza de que <PROSEGL><SUBORD_conj/pron_rel+oração>
<cpl19>acordem no paraíso.
<veloc_leg 49c/5.666s 8.648>
</sub816>

<sub822><2L>
<t>00:51:53,944 --> 00:51:58,108</t>
<cpl33>Andei pensando. Se você desse uma <PROSEGL><SN_especif+SN>
<cpl22>chance a este lugar...
<veloc_leg 55c/4.164s 13.208>

</sub822>

<sub872><2L>

<t>00:55:52,849 --> 00:55:57,343</t>

<cpl23>27, 28, 29. 30 pretas e <PROSEGL><COORD_conj+oração>

<cpl14>só 29 brancas.

<veloc_leg 37c/4.494s 8.233>

</sub872>

<sub882><2L>

<t>00:56:34,958 --> 00:56:39,122</t>

<cpl32>Logo colocaremos o meu excelente <PROSEGL><SN_subst+SAdj>

<cpl14>plano em ação.

<veloc_leg 46c/4.164s 11.047>

</sub882>

<sub884><2L>

<t>00:56:41,998 --> 00:56:45,661</t>

<cpl32>...atê que estejam profundamente <PROSEGL><SAdj_especif+SAdj>

<cpl12>adormecidos.

<veloc_leg 44c/3.663s 12.012>

</sub884>

<sub908><2L>

<t>00:58:28,705 --> 00:58:32,038</t>

<cpl24>Vamos bombear o sangue e <PROSEGL><COORD_conj+oração>

<cpl29>mandar ar fresco aos pulmões.

<veloc_leg 53c/3.333s 15.902>

</sub908>

<sub936><2L>

<t>00:59:54,858 --> 01:00:00,194</t>

<cpl35>Me sinto um pastrami de 1 km em pão <PROSEGL><SN_subst+SAdj>

<cpl32>de centeio saindo da lanchonete!

<veloc_leg 67c/5.336s 12.556>

</sub936>

<sub976><2L>

<t>01:02:52,135 --> 01:02:56,595</t>

<cpl31>Por favor, Maurice! O que é uma <PROSEGL><SN_especif+SN>

<cpl29>simples mordida entre amigos?

<veloc_leg 60c/4.46s 13.453>

</sub976>

<sub1027><2L>

<t>01:07:38,988 --> 01:07:42,981</t>

<cpl23>O navio voltou para nos <PROSEGL><SV_(verbo)+oblíquo+SV>

<cpl27>pegar. Temos que sinalizar.

<veloc_leg 50c/3.993s 12.522>

</sub1027>

<sub1054><2L>
 <t>01:09:13,416 --> 01:09:17,853</t>
 <cpl33>Onde está aquele leão psicótico e <PROSEGL><COORD_conj+oração>
 <cpl28>o nosso amigo monocromático?
 <veloc_leg 61c/4.437s 13.748>
 </sub1054>

<sub1060><2L>
 <t>01:09:33,736 --> 01:09:37,263</t>
 <cpl30>Temos que ganhar os corações e <PROSEGL><COORD_conj+oração>
 <cpl22>as mentes dos nativos.
 <veloc_leg 52c/3.527s 14.743>
 </sub1060>

<sub1075><2L>
 <t>01:10:32,528 --> 01:10:36,589</t>
 <cpl32>E podemos voltar à civilização e <PROSEGL><COORD_conj+oração>
 <cpl30>tudo voltará a ser como antes.
 <veloc_leg 62c/4.061s 15.267>
 </sub1075>

<sub1135><2L>
 <t>01:15:43,706 --> 01:15:47,767</t>
 <cpl35>Nunca mais quero ver vocês na minha <PROSEGL><SN_especif+SN>
 <cpl13>área de novo.
 <veloc_leg 48c/4.061s 11.820>
 </sub1135>

<sub1165><2L>
 <t>01:17:18,267 --> 01:17:20,701</t>
 <cpl29>que tem o coração maior que o <PROSEGL><SN_especif+SN>
 <cpl9>estômago.
 <veloc_leg 38c/2.434s 15.612>
 </sub1165>

<sub1176><2L>
 <t>01:17:52,835 --> 01:17:56,566</t>
 <cpl29>Mas antes de partir tenho uma <PROSEGL><SN_especif+SN>
 <cpl20>comunicação a fazer.
 <veloc_leg 49c/3.731s 13.133>
 </sub1176>

<sub1182><2L>
 <t>01:18:11,253 --> 01:18:14,086</t>
 <cpl26>lhe darei este presente de <PROSEGL><SP_prep+SN>
 <cpl10>despedida.
 <veloc_leg 36c/2.833s 12.707>
 </sub1182>

<sub1201><2L>
 <t>01:18:58,667 --> 01:19:01,363</t>
 <cpl27>Podíamos voltar aqui de vez <PROSEGL><SAdj_especif+SAdj>
 <cpl10>em quando.
 <veloc_leg 37c/2.696s 13.724>
 </sub1201>

Filme: *Social network* (2010)

<sub75><2L>
 <t>02:03:50,001 --> 02:03:53,437</t>
 <cpl21>Sinceramente... devia <PROSEGL><SV_(verbo)+oblíquo+SV>
 <cpl15>me apoiar mais.
 <veloc_leg 36c/3.436s 10.477>
 </sub75>

<sub84><2L>
 <t>02:04:17,562 --> 02:04:20,998</t>
 <cpl22>Vou conhecer gente que <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl21>eu nunca conheceria?!
 <veloc_leg 43c/3.436s 12.515>
 </sub84>

<sub90><2L>
 <t>02:04:36,915 --> 02:04:39,406</t>
 <cpl21>Você, por acaso, é de <PROSEGL><SP_prep+SN>
 <cpl23>Long Island, Wimbledon?
 <veloc_leg 44c/2.491s 17.664>
 </sub90>

<sub110><2L>
 <t>02:05:26,932 --> 02:05:31,301</t>
 <cpl27>Você ainda vai ser um gênio <PROSEGL><SN_subst+SAdj>
 <cpl29>bem-sucedido da computação...
 <veloc_leg 56c/4.369s 12.818>
 </sub110>

<sub316><2L>
 <t>02:19:54,031 --> 02:19:57,660</t>
 <cpl24>Eu-- Já me desculpei com <PROSEGL><SP_prep+SN>
 <cpl24>o The Crimson, à ABHW...
 <veloc_leg 48c/3.629s 13.227>
 </sub316>

<sub348><2L>
 <t>02:21:14,979 --> 02:21:18,938</t>
 <cpl23>Temos um computador com <PROSEGL><SP_prep+SN>
 <cpl33>um endereço virtual de 16 bits...

<veloc_leg 56c/3.959s 14.145>
</sub348>

<sub358><2L>
<t>02:21:44,975 --> 02:21:48,467</t>
<cpl32>Sr. Zuckerberg, gente ainda mais <PROSEGL><SAdj_especif+SAdj>
<cpl30>genial não passou neste curso.
<veloc_leg 62c/3.492s 17.755>
</sub358>

<sub359><2L>
<t>02:21:48,646 --> 02:21:53,140</t>
<cpl36>1 bit válido, 1 bit modificado, 1 de <PROSEGL><SP_prep+SV>
<cpl33>referência e 5 bits de permissão.
<veloc_leg 69c/4.494s 15.354>
</sub359>

<sub394><2L>
<t>02:23:19,870 --> 02:23:22,930</t>
<cpl24>Um aplicativo de MP3 que <PROSEGL><SUBORD_conj/pron_rel+oração>
<cpl28>reconhece seu gosto musical.
<veloc_leg 52c/3.06s 16.993>
</sub394>

<sub411><2L>
<t>02:24:03,347 --> 02:24:06,111</t>
<cpl25>Divya e o meu irmão sabem <PROSEGL><SV_composto>
<cpl15>ser grosseiros.
<veloc_leg 40c/2.764s 14.472>
</sub411>

<sub503><2L>
<t>02:28:14,198 --> 02:28:17,429</t>
<cpl21>Num mundo regido pela <PROSEGL><SP_prep+SN>
<cpl24>estratificação social...
<veloc_leg 45c/3.231s 13.928>
</sub503>

<sub536><2L>
<t>02:29:44,087 --> 02:29:46,612</t>
<cpl30>ele mencionou Tyler ou Cameron <PROSEGL><SN_núcleo>
<cpl13>Winklevoss...
<veloc_leg 43c/2.525s 17.030>
</sub536>

<sub537><2L>
<t>02:29:46,790 --> 02:29:48,951</t>
<cpl17>Divya Narendra ou <PROSEGL><COORD_conj+oração>
<cpl18>HarvardConnection?
<veloc_leg 35c/2.161s 16.196>

</sub537>

<sub539><2L>

<t>02:29:51,862 --> 02:29:55,491</t>

<cpl28>mas pelo que tinham, viu que <PROSEGL><SUBORD_conj/pron_rel+oração>

<cpl22>não valeria seu tempo.

<veloc_leg 50c/3.629s 13.778>

</sub539>

<sub544><2L>

<t>02:30:06,877 --> 02:30:09,869</t>

<cpl29>por acharem que ele precisava <PROSEGL><SV_composto>

<cpl22>reabilita sua imagem.

<veloc_leg 51c/2.992s 17.045>

</sub544>

<sub555><2L>

<t>02:30:37,708 --> 02:30:40,973</t>

<cpl31>ele convencia os autores de que <PROSEGL><SUBORD_conj/pron_rel+oração>

<cpl30>construía o HarvardConnection?

<veloc_leg 61c/3.265s 18.683>

</sub555>

<sub560><2L>

<t>02:30:50,153 --> 02:30:53,145</t>

<cpl20>"Li a sua info sobre <PROSEGL><SP_prep+SN>

<cpl24>a HarvardConnection... "

<veloc_leg 44c/2.992s 14.706>

</sub560>

<sub562><2L>

<t>02:30:56,159 --> 02:31:00,755</t>

<cpl35>"Falamos quando eu tiver uma versão <PROSEGL><SN_subst+SAdj>

<cpl24>básica, amanhã à noite."

<veloc_leg 59c/4.596s 12.837>

</sub562>

<sub571><2L>

<t>02:31:24,921 --> 02:31:28,186</t>

<cpl27>Quem mais acha que ele está <PROSEGL><SV_(verbo)+oblíquo+SV>

<cpl13>nos evitando?

<veloc_leg 40c/3.265s 12.251>

</sub571>

<sub611><2L>

<t>02:33:51,768 --> 02:33:55,864</t>

<cpl28>39 dias após a 1º reunião de <PROSEGL><SP_prep+SV>

<cpl31>Zuckerberg com meus clientes...

<veloc_leg 59c/4.096s 14.404>

</sub611>

<sub619><2L>
 <t>02:34:24,267 --> 02:34:28,431</t>
 <cpl25>"Cameron, ainda duvido da <PROSEGL><SP_prep+SV>
 <cpl31>funcionalidade do site para..."
 <veloc_leg 56c/4.164s 13.449>
 </sub619>

<sub637><2L>
 <t>02:35:10,814 --> 02:35:13,783</t>
 <cpl24>Vai ser um prazer te ver <PROSEGL><SV_composto>
 <cpl23>preencher nosso cheque.
 <veloc_leg 47c/2.969s 15.830>
 </sub637>

<sub642><2L>
 <t>02:35:29,099 --> 02:35:32,830</t>
 <cpl33>Uma gata da tua turma de História <PROSEGL><SN_subst+SP>
 <cpl27>da Arte, Stephanie Attis...
 <veloc_leg 60c/3.731s 16.081>
 </sub642>

<sub700><2L>
 <t>02:39:06,716 --> 02:39:09,514</t>
 <cpl24>Não sei. Podia ser de um <PROSEGL><SN_especif+SN>
 <cpl22>GatosSósiadoHitler...
 <veloc_leg 46c/2.798s 16.440>
 </sub700>

<sub731><2L>
 <t>02:41:00,530 --> 02:41:04,523</t>
 <cpl33>Nem traficante passaria drogas de <PROSEGL><SAdv_adv+adv>
 <cpl30>graça pra 650 pessoas num dia.
 <veloc_leg 63c/3.993s 15.778>
 </sub731>

<sub732><2L>
 <t>02:41:04,701 --> 02:41:08,102</t>
 <cpl18>E ele não tem mais <PROSEGL><SAdv_adv+adv>
 <cpl16>do que 3 amigos.
 <veloc_leg 34c/3.401s 9.997>
 </sub732>

<sub750><2L>
 <t>02:41:52,315 --> 02:41:55,250</t>
 <cpl19>Falou em mandar uma <PROSEGL><SN_especif+SN>
 <cpl26>notificação extrajudicial.
 <veloc_leg 45c/2.935s 15.332>
 </sub750>

<sub800><2L>
 <t>02:44:09,619 --> 02:44:12,383</t>
 <cpl35>Eduardo era presidente da Assoc. de <PROSEGL><SP_prep+SV>
 <cpl26>Investidores de Harvard...
 <veloc_leg 61c/2.764s 22.069>
 </sub800>

<sub893><2L>
 <t>02:48:16,999 --> 02:48:19,968</t>
 <cpl15>Alegam roubo de <PROSEGL><SP_prep+SN>
 <cpl24>propriedade intelectual.
 <veloc_leg 39c/2.969s 13.136>
 </sub893>

<sub896><2L>
 <t>02:48:23,339 --> 02:48:26,274</t>
 <cpl28>Dizem que roubamos o site de <PROSEGL><SP_prep+SN>
 <cpl29>D. Narendra e dos Winklevoss.
 <veloc_leg 57c/2.935s 19.421>
 </sub896>

<sub937><2L>
 <t>02:50:21,490 --> 02:50:24,891</t>
 <cpl33>Fiz um Juramento aqui e não quero <PROSEGL><SV_composto>
 <cpl17>cometer perjúrio.
 <veloc_leg 50c/3.401s 14.702>
 </sub937>

<sub949><2L>
 <t>02:51:03,632 --> 02:51:07,193</t>
 <cpl23>São 1 2h45. Por que não <PROSEGL><SV_verbo+SAdv>
 <cpl22>paramos para o almoço?
 <veloc_leg 45c/3.561s 12.637>
 </sub949>

<sub978><2L>
 <t>02:53:18,100 --> 02:53:20,591</t>
 <cpl24>Como se seus pensamentos <PROSEGL><SN_subst+SAdj>
 <cpl12>"geniais"...
 <veloc_leg 36c/2.491s 14.452>
 </sub978>

<sub1005><2L>
 <t>02:54:36,678 --> 02:54:39,408</t>
 <cpl17>e seus 5% saem da <PROSEGL><SP_prep+SN>
 <cpl19>minha participação.
 <veloc_leg 36c/2.73s 13.187>
 </sub1005>

<sub1048><2L>

<t>02:56:44,640 --> 02:56:47,768</t>
 <cpl27>e a federal, expandindo pra <PROSEGL><SP_prep+SN>
 <cpl29>Connecticut, NY e Califórnia.
 <veloc_leg 56c/3.128s 17.903>
 </sub1048>

<sub1093><2L>
 <t>02:58:40,322 --> 02:58:45,055</t>
 <cpl29>Criei um site pra se baixar e <PROSEGL><COORD_conj+oração>
 <cpl30>compartilhar músicas de graça.
 <veloc_leg 59c/4.733s 12.466>
 </sub1093>

<sub1142><2L>
 <t>03:00:48,283 --> 03:00:52,413</t>
 <cpl33>Um prédio 100 anos mais velho que
 <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl29>este país, portanto, cuidado.
 <veloc_leg 62c/4.13s 15.012>
 </sub1142>

<sub1167><2L>
 <t>03:01:57,018 --> 03:02:00,146</t>
 <cpl23>Respeitar a propriedade <PROSEGL><SN_subst+SAdj>
 <cpl20>pública e privada...
 <veloc_leg 43c/3.128s 13.747>
 </sub1167>

<sub1168><2L>
 <t>03:02:00,322 --> 03:02:02,313</t>
 <cpl25>e que roubo e apropriação <PROSEGL><SN_subst+SAdj>
 <cpl10>indébita--
 <veloc_leg 35c/1.991s 17.579>
 </sub1168>

<sub1199><2L>
 <t>03:03:15,697 --> 03:03:18,495</t>
 <cpl28>Francamente, não sei se está <PROSEGL><SV_verbo+SAdv>
 <cpl14>apto a julgar.
 <veloc_leg 42c/2.798s 15.011>
 </sub1199>

<sub1250><2L>
 <t>03:05:51,152 --> 03:05:53,985</t>
 <cpl22>Um jantar marcado pela <PROSEGL><SP_prep+SN>
 <cpl24>minha namorada na época.
 <veloc_leg 46c/2.833s 16.237>
 </sub1250>

<sub1372><2L>

<t>03:11:25,653 --> 03:11:29,350</t>
 <cpl23>mas antes deu sua maior <PROSEGL><SN_subst+SAdj>
 <cpl23>contribuição à empresa.
 <veloc_leg 46c/3.697s 12.443>
 </sub1372>

<sub1390><2L>
 <t>03:12:22,310 --> 03:12:27,646</t>
 <cpl32>65% para M. Zuckerberg, 30% para <PROSEGL><SP_prep+SN>
 <cpl34>E. Saverin e 5% para D. Moskovitz?
 <veloc_leg 66c/5.336s 12.369>
 </sub1390>

<sub1395><2L>
 <t>03:12:37,792 --> 03:12:41,421</t>
 <cpl32>que possa ser considerado motivo <PROSEGL><SN_subst+SAdj>
 <cpl23>legítimo para rescisão?
 <veloc_leg 55c/3.629s 15.156>
 </sub1395>

<sub1465><2L>
 <t>03:15:51,118 --> 03:15:54,849</t>
 <cpl31>Não precisamos disso, e sim, de <PROSEGL><SP_prep+SN>
 <cpl30>anunciantes. E também conheço.
 <veloc_leg 61c/3.731s 16.350>
 </sub1465>

<sub1497><2L>
 <t>03:17:40,895 --> 03:17:44,695</t>
 <cpl30>E o que o estagiário precisará <PROSEGL><SV_composto>
 <cpl25>fazer bêbado no trabalho?
 <veloc_leg 55c/3.8s 14.474>
 </sub1497>

<sub1521><2L>
 <t>03:18:54,001 --> 03:18:59,006</t>
 <cpl28>Mas eu pensei: o que poderia <PROSEGL><SV_composto>
 <cpl23>sair errado em 3 meses?
 <veloc_leg 51c/5.005s 10.190>
 </sub1521>

<sub1573><2L>
 <t>03:21:50,477 --> 03:21:53,344</t>
 <cpl20>Constrangido de ir a <PROSEGL><SP_prep+SN>
 <cpl24>lojas de departamento...
 <veloc_leg 44c/2.867s 15.347>
 </sub1573>

<sub1574><2L>
 <t>03:21:53,514 --> 03:21:57,610</t>

<cpl28>pensou numa butique fina que <PROSEGL><SUBORD_conj/pron_rel+oração>
 <cpl33>não te fizesse parecer um tarado.
 <veloc_leg 61c/4.096s 14.893>
 </sub1574>

<sub1591><2L>
 <t>03:22:45,566 --> 03:22:50,594</t>
 <cpl33>A aluna que eu curtia ficou com o <PROSEGL><SN_especif+SN>
 <cpl35>vice-capitão do time de lacrosse...
 <veloc_leg 68c/5.028s 13.524>
 </sub1591>

<sub1614><2L>
 <t>03:23:51,932 --> 03:23:55,265</t>
 <cpl23>É uma puta de uma ideia <PROSEGL><SN_subst+SAdj>
 <cpl16>única na vida...
 <veloc_leg 39c/3.333s 11.701>
 </sub1614>

<sub1681><2L>
 <t>03:28:16,830 --> 03:28:19,822</t>
 <cpl16>E em Oxford e na <PROSEGL><SP_prep+SN>
 <cpl30>Escola de Economia de Londres.
 <veloc_leg 46c/2.992s 15.374>
 </sub1681>

<sub1692><2L>
 <t>03:28:47,594 --> 03:28:49,789</t>
 <cpl26>Mas me incomoda, e deveria <PROSEGL><SV_(verbo)+oblíquo+SV>
 <cpl15>te incomodar...
 <veloc_leg 41c/2.195s 18.679>
 </sub1692>

<sub1695><2L>
 <t>03:28:55,735 --> 03:28:59,193</t>
 <cpl19>junto ao Conselho e <PROSEGL><COORD_conj+oração>
 <cpl27>ao presidente da faculdade.
 <veloc_leg 46c/3.458s 13.302>
 </sub1695>

<sub1782><2L>
 <t>03:33:06,786 --> 03:33:09,186</t>
 <cpl25>Muda pra cá ou vai acabar <PROSEGL><SV_composto>
 <cpl17>ficando pra trás.
 <veloc_leg 42c/2.4s 17.500>
 </sub1782>

<sub1814><2L>
 <t>03:34:44,250 --> 03:34:47,219</t>
 <cpl28>com uma estrutura de capital <PROSEGL><SN_subst+SAdj>

<cpl28>aberta a novos investidores.
 <veloc_leg 56c/2.969s 18.862>
 </sub1814>

<sub1841><2L>
 <t>03:36:09,736 --> 03:36:13,399</t>
 <cpl29>É pra pegar as vadias do Vale <PROSEGL><SN_núcleo>
 <cpl32>do Silicio, nas visitas ao Mark!
 <veloc_leg 61c/3.663s 16.653>
 </sub1841>

<sub1924><2L>
 <t>03:39:52,925 --> 03:39:55,723</t>
 <cpl23>"ciente de que era seu <PROSEGL><SN_especif+SN>
 <cpl20>atestado de óbito?"
 <veloc_leg 43c/2.798s 15.368>
 </sub1924>

<sub1925><2L>
 <t>03:40:01,200 --> 03:40:05,261</t>
 <cpl29>Foi burrice não ter submetido <PROSEGL><SV_verbo+SAdv>
 <cpl24>tudo aos meus advogados.
 <veloc_leg 53c/4.061s 13.051>
 </sub1925>

<sub1960><2L>
 <t>03:42:00,119 --> 03:42:03,350</t>
 <cpl25>tinha feito uma oferta de <PROSEGL><SP_prep+SN>
 <cpl25>investimento irrecusável.
 <veloc_leg 50c/3.231s 15.475>
 </sub1960>

<sub2043><2L>
 <t>03:46:28,420 --> 03:46:33,187</t>
 <cpl32>Falando nisso, vou dar uma festa <PROSEGL><SN_subst+SAdj>
 <cpl29>pós-festa na Kappa Eta Sigma.
 <veloc_leg 61c/4.767s 12.796>
 </sub2043>

<sub2070><2L>
 <t>03:48:10,022 --> 03:48:13,719</t>
 <cpl20>Vai levar a câmera e <PROSEGL><COORD_conj+oração>
 <cpl25>exibir aos amigos online.
 <veloc_leg 45c/3.697s 12.172>
 </sub2070>

<sub2145><2L>
 <t>03:52:36,755 --> 03:52:41,215</t>
 <cpl26>Sy e os outros foram a uma <PROSEGL><SN_especif+SN>
 <cpl30>churrascaria na University Av.

<veloc_leg 56c/4.46s 12.556>
</sub2145>