http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

CONSTRUINDO *CORPORA* DE LEGENDAS: PASSO A PASSO METODOLÓGICO PARA PESQUISAS BASEADAS EM *CORPUS*

BUILDING SUBTITLES CORPORA: A STEP-BY-STEP METHODOLOGY FOR CORPUS-BASED RESEARCH

Janailton Mick Vitor da Silva
Instituto Federal Goiano – *Campus* Campos Belos, Brasil
janailtonm@gmail.com

RESUMO: O objetivo deste artigo é descrever o passo a passo metodológico para criação de *corpora* de legendas, retiradas de obras audiovisuais, como filmes e séries de TV, que pode servir a pesquisadores que trabalham no campo da Tradução Audiovisual e dos Estudos da Tradução Baseados em *Corpus.* O ponto de partida para a descrição ora introduzida baseou-se na pesquisa de Silva (2018), na qual foram apresentados passos para compilação, edição e preparação de *corpora* de legendas da série de TV *Star Trek: Enterprise,* utilizando-se programas razoavelmente conhecidos, como o Google Chrome, Bloco de Notas, Subtitle Edit, Microsoft Word e Excel. No presente artigo, entende-se que os passos aqui apresentados se estabelecem como sugestões de futuros percursos metodológicos a serem seguidos em pesquisas nas áreas supramencionadas.
PALAVRAS-CHAVE: construção de *corpora* de legendas; passos metodológicos; Tradução Audiovisual; Estudos da Tradução Baseados em *Corpus.*

ABSTRACT: The purpose of this article is to describe the step-by-step process for the creation of subtitles corpora, extracted from audiovisual works, such as films and TV series, that may be useful for researchers in the field of Audiovisual Translation and Corpus-Based Translation Studies. The starting point for such description is based on the research of Silva (2018), in which steps for compiling, editing and preparing corpora subtitles of the TV Series Star Trek: Enterprise were presented, using such reasonably known programs as Google Chrome, Notepad, Subtitle Edit, Microsoft Word and Excel. In the present article, it is understood that the steps introduced herein are suggestions intended as future methodological steps to be followed in research done in the aforementioned areas.
KEYWORDS: subtitles corpora compilation; methodological steps; Audiovisual Translation; Corpus-Based Translation Studies.

1 Introdução

Este artigo tem o objetivo de descrever o passo a passo metodológico da criação de *corpora* de legendas que, na pesquisa de Silva (2018), foram produzidas para a série de TV *Star Trek: Enterprise.* Esse percurso metodológico pode servir como possível modelo para futuras pesquisas que façam uso de *corpus/corpora* de legendas. No presente texto, entende-se *corpus* (singular da palavra *corpora*) como uma coleção de

http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

textos em formato eletrônico, que são passíveis de análises automáticas e/ou semiautomáticas (BAKER, 1996). Sendo assim, foram criados *corpora* de estudo e de referência, em inglês e português, que serviram, na prática, para realização de pesquisa de mestrado do autor do presente artigo (SILVA, 2018), defendida em julho de 2018, no Programa de Pós-Graduação em Estudos da Tradução, na Universidade de Brasília, e financiada pela Capes.

A motivação para construção de *corpora* de legendas parte do entrelaçamento aqui feito entre a Tradução Audiovisual (TAV) com os Estudos da Tradução Baseados em *Corpus* (ETBC), para assim ser possível analisar dados linguísticos, oriundos de legendas previamente criadas e dispostas na tela da TV, com o auxílio de ferramentas da Linguística de *Corpus* (LC) disponibilizadas pelo programa WordSmith Tools (WST), a saber, o Concord e o WordList (SCOTT, 2018). Dentro desse escopo, o que interessa a este artigo é a etapa anterior à manipulação de dados linguísticos pelas ferramentas da LC, isto é, aquela que se inicia no momento de acesso às legendas no *site* da Netflix e finaliza no momento em que são criados arquivos para posterior utilização no *WST*.

Neste artigo, legendas são aqui compreendidas como os textos que aparecem geralmente no fim da tela, em sincronia com imagem e diálogo, fornecendo uma tradução semanticamente adequada do diálogo da língua fonte (LF) e permanecendo na tela tempo suficiente para sua leitura e visualização pela audiência (CHAUME, 2004; DÍAZ CINTAS; REMAEL, 2007; GEORGAKOPOULOU, 2009; GOTTLIEB, 1994; 1998; 2005a; 2005b; IVARSSON; CARROLL, 1998). Além disso, a área da legendagem, dentro do escopo da TAV, compreende uma dimensão semiótica ampla, cujos textos fontes (TFs) são de natureza polissemiótica, ou seja, constituídos por signos de natureza diversa (GOTTLIEB, 2005a), dada a correlação de signos verbais e não verbais que compõem as cenas das obras audiovisuais.

É preciso levar em conta, também, o espaço em que a utilização de *corpora* ocupa nos Estudos da Tradução (ET) desde os anos 1993, quando Baker havia inicialmente advogado por seu uso nos ET (OLOHAN, 2004). Segundo Olohan (2004), *corpora* servem como metodologia dentro dos ET, como é o caso da pesquisa de mestrado supracitada. Nessa investigação, fez-se uso de arquivos eletrônicos de textos fontes (TFs) e textos traduzidos (TTs), bem como se utilizou um programa computacional, como é o caso do WST, para fins de análise linguística sob uma perspectiva descritiva (BAKER, 1995; 1996; CAMARGO, 2007), o que levou a pesquisa a se enquadrar dentro dos ETBC. Outros programas, fora dos ETBC, foram utilizados como parte do desenvolvimento da pesquisa, como Google Chrome, Bloco de Notas, Subtitle Edit, Microsoft Word e Excel[1]. A partir do tópico seguinte, inicia-se a apresentação dos *corpora* compilados.

2 Apresentação dos *corpora* da pesquisa

Os textos que serviram como base para criação dos *corpora* da pesquisa de Silva

1 Faz-se relevante mencionar que, embora a revista na qual o presente artigo se encontra publicado fomente a utilização de *softwares* livres, e que os *softwares* utilizados na pesquisa aqui retratada tenham sido, em sua maioria, pagos, outros programas livres (ex.: pacote *LibreOffice*, *AntConc, ParaConc,* entre outros) podem servir a futuros pesquisadores para etapas de pesquisas que contribuam para a compilação e manipulação de *corpora.*

Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

(2018) foram coletados de episódios da primeira e segunda temporadas da série de TV *Star Trek: Enterprise,* disponibilizada pela Netflix para assinantes. Com esse material, foi possível compilar os seguintes *corpora*: 1) um *corpus* de estudo (abreviado como CETAR), constituído por TTs (legendas) em Português Brasileiro (PB), feitos pela legendista responsável por traduzir os episódios (Talita Ribeiro), a partir de TFs em inglês americano; 2) um *corpus* de referência (abreviado como CROL), composto por TTs em PB feitos por outros legendistas, a partir de TFs em inglês americano; 3) um *corpus* paralelo (abreviado como CPTR), formado pelo CETAR e seus respectivos TFs. Tanto no CETAR quanto no CROL, os TFs consistiam em transcrições de áudio em inglês americano, no formato *Closed Caption* (*CC*), já disponibilizadas pela Netflix.

Para a seleção desses *corpora* supramencionados*,* foi necessário seguir alguns pré-requisitos básicos para a construção de qualquer *corpus* na área da LC, conforme Berber Sardinha (2004), tais como conter textos autênticos produzidos em linguagem natural, demonstrar que a escolha desses textos foi feita com critérios, e ser representativo, principalmente para os objetivos da investigação.

Entre as séries legendadas por Talita, a escolha por *Star Trek: Enterprise* se deu com base em três critérios. Primeiramente, dentre os filmes e séries do seu currículo, nem todos estavam disponíveis na Netflix, e, quando estavam, tinham legendas em apenas uma das línguas de interesse. Em segundo lugar, *Star Trek: Enterprise* foi a única série cujas temporadas, da primeira até a quarta, apresentavam ao menos duas traduções da legendista em cada temporada. No caso da pesquisa de Silva (2018), esse elemento era de extrema importância, uma vez que, nos estudos de estilo, o trabalho com várias traduções feitas pela mesma profissional, compostas por vários episódios de temporadas distintas, possibilitaria identificar um padrão de escolhas feitas por ela. Por fim, o último critério de escolha deu-se pelo fato de a primeira e a segunda temporadas de *Star Trek: Enterprise* serem criadas em um formato episódico (MITTELL, 2006), em que os conflitos da trama eram introduzidos e resolvidos em cada episódio. Sendo assim, não seria necessário assistir a todos os episódios de todas as temporadas para entender o que se desenvolvia na série.

Por fim, a representatividade dos *corpora* foi entendida com relação ao seu tamanho. Cada temporada tinha 26 episódios. O CETAR era composto pelos TTs por Talita, mais especificamente, por dois episódios da primeira temporada e dois episódios da segunda temporada. O CROL, por outro lado, era conforme sugerido por Berber Sardinha (2000), seis vezes maior do que o CETAR e era composto por onze episódios da primeira temporada e doze episódios da segunda temporada, claramente apresentando maior quantidade de caracteres. Nesse *corpus*, a seleção dos episódios foi feita aleatoriamente. Nessa escolha, no entanto, tentou-se selecionar episódios que precedessem e/ou sucedessem aqueles legendados por Talita. É importante adicionar que o CROL continha TTs por outros legendistas, cuja identificação não foi obtida na pesquisa de Silva (2018), tendo em vista a ausência de créditos nas legendas, ao fim dos episódios, nem outras fontes que pudessem indicar e atestar sua autoria.

A seguir, nos Quadros 1, 2 e 3, são apresentados os textos que fizeram parte dos *corpora* acima elencados. É relevante ressaltar que os títulos dos episódios em português são de autoria da Netflix.

http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

Quadro 1: Textos do CETAR.

COMPOSIÇÃO DO *CORPUS* DE ESTUDO			
TEMPORADA 1		**TEMPORADA 2**	
Nº / DURAÇÃO	**EPISÓDIOS**	**Nº / DURAÇÃO**	**EPISÓDIOS**
10 (45’)	*Cold front* (Frente Fria)	16 (43’)	*Future Tense* (Futuro do Presente)
11 (45’)	*Silent Enemy* (Inimigo Silencioso)	24 (43’)	*First Flight* (Primeiro voo)

Fonte: Silva (2018, p. 59).

Quadro 2: Textos do CROL – temporada 1.

COMPOSIÇÃO DO *CORPUS* DE REFERÊNCIA			
TEMPORADA 1			
Nº / DURAÇÃO	**EPISÓDIOS**	**Nº / DURAÇÃO**	**EPISÓDIOS**
1 (87’)	*Broken Bow (Parts 1, 2)* (Broken Bow – partes 1, 2)	21 (45’)	*Vox Sola* (Vox Sola)
2 (45’)	*Fight or flight* (Luta ou fuga)	22 (45’)	*Fallen Hero* (Herói em decadência)
3 (45’)	*Strange new world* (Explorar novos mundos)	23 (45’)	*Desert Crossing* (Travessia no Deserto)
4 (45’)	*Unexpected* (Inesperado)	24 (45’)	*Two days and two nights* (Dois dias e duas noites)
5 (45’)	*Terra Nova* (Terra Nova)	25 (45’)	*Shockwave (Part I)* (Onda de choque – parte 1)
20 (45’)	*Detained* (Detidos)		

Fonte: Silva (2018, p. 59).

Quadro 3: Textos do CROL – temporada 2.

COMPOSIÇÃO DO *CORPUS* DE REFERÊNCIA			
TEMPORADA 2			
Nº / DURAÇÃO	**EPISÓDIOS**	**Nº / DURAÇÃO**	**EPISÓDIOS**
1 (43’)	*Shockwave (Part 2)* (Onda de choque – parte 2)	11 (43’)	*Precious Cargo* (Carga preciosa)
2 (43’)	*Carbon Creek* (Carbon Creek)	18 (43’)	*The Crossing* (A travessia)
3 (43’)	*Minefield* (Campo Minado)	19 (43’)	*Judgment* (Julgamento)

http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

Nº / DURAÇÃO	EPISÓDIOS	Nº / DURAÇÃO	EPISÓDIOS
4 (43’)	*Dead Stop* (Ponto Morto)	20 (43’)	*Horizon* (Horizon)
9 (43’)	*Singularity* (Singularidade)	21 (43’)	*The Breach* (A ruptura)
10 (43’)	*Vanishing Point* (Ponto de Fuga)	26 (43’)	*The expanse* (A expansão)

Fonte: Silva (2018, p. 60).

3 A construção de *corpora* de legendas

Nesta seção, apresentam-se os procedimentos metodológicos da pesquisa de Silva (2018), por meio de instruções que serviram como base para compilação, edição e preparação dos *corpora* da sua pesquisa. É importante ressaltar que as instruções a seguir são meras sugestões que podem ser seguidas e/ou adaptadas à realidade dos objetivos de futuras pesquisas que façam uso de *corpora* de legendas.

3.1 Compilando os arquivos de legendas

Neste tópico, apresenta-se o passo a passo para obtenção dos TFs e dos TTs dos episódios de duas temporadas de *Star Trek: Enterprise.*

3.1.1 Obtendo o texto fonte de cada episódio *online*

O TF é a transcrição cronometrada de áudio em inglês para cada um dos episódios. Para obtê-la, foi preciso ativar a função *CC* quando o episódio era exibido pela Netflix, fazendo com que as legendas em inglês fossem acionadas e aparecessem na tela. Assistiu-se a cada episódio para confirmar que as legendas eram a transcrição literal do áudio; caso contrário, seria preciso fazer possíveis alterações na etapa de análise. No ato de assistir, observou-se que as legendas em inglês, sendo primordialmente destinadas ao público surdo e ensurdecido, apresentavam a identificação de sons e de personagens em formato de legendas.

Para fazer o *download* de cada TF, como foi feito para cada episódio do CETAR e do CROL, utilizou-se o Google Chrome (a versão utilizada por Silva (2018) foi a 61.0.3163.100 – 64 bits), seguindo os passos abaixo:

i) Abrir a página do episódio na Netflix;

ii) Acessar as ‘Ferramentas do desenvolvedor’ na aba ‘Mais ferramentas’ do Google Chrome e certifiquar-se de que a guia superior ‘*Network’* está pressionada;

iii) Adicionar o código ‘?o=’ no campo ‘*Filter’;*

iv) Ativar, na página do episódio, a legenda na língua desejada, ex.: ‘inglês

http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

[CC]';

v) Observar que, assim que a legenda é ativada e o episódio é reproduzido, uma lista de solicitações recuperadas através da guia '*Network*' carrega rapidamente na parte média-inferior da tela, numa área chamada de '*Requests Table',* que consiste em seis colunas[2]: '*Name'*, '*Status'*, '*Type'*, '*Initiator'*, '*Size'*, '*Time'*, '*Waterfall'*';

vi) Pesquisar e selecionar, na lista de recuperações na coluna '*Name',* o nome do arquivo que começa com '?o='. Observar que uma nova guia ao lado do arquivo abrirá, na aba '*Response',* exibindo as legendas enumeradas.

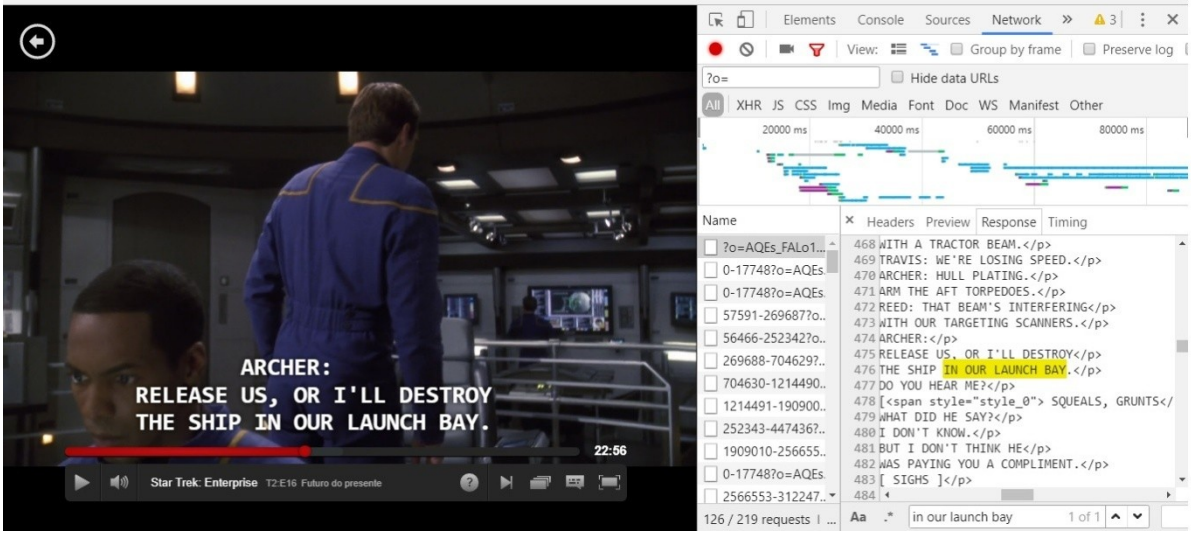


Figura 1: Obtendo a transcrição em Inglês pelo Google Chrome: exemplo de *Star Trek, Season 2, Episode 16* (CETAR).
Fonte: Silva (2018, p. 70).

vii) Clicar com o botão direito sobre o arquivo e selecionar a opção '*Open in a new tab*' (abrir em uma nova guia). O *download* do arquivo começará automaticamente e será salvo com nome 'download';

viii) Renomear o arquivo baixado com o nome desejado, como, por exemplo, 'ST_S02_EP16_T.xml' (arquivo de legenda da Figura 1) e salvá-lo em extensão *.xml.*

3.1.2 Obtendo o texto traduzido de cada episódio *online*

O TT corresponde às legendas cronometradas em português. Na pesquisa de Silva (2018), elas foram obtidas também pelo Google Chrome, seguindo-se a maioria dos passos para a obtenção dos TFs descritos em 3.1.1. Contudo, algumas diferenças podem ser notadas nas etapas abaixo, como é sugerido a seguir:

i) Selecionar a legenda para o episódio na língua desejada, ex.: 'português'.

2 Provavelmente, outras funções além daquelas aqui nomeadas estão disponíveis nas colunas da 'Requests Table', mas, na versão do Google Chrome utilizada na pesquisa de Silva (2018), foram apresentadas seis colunas.

http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

Após adotar os procedimentos citados nos itens 5 e 6 da subseção 3.1.1, observar as legendas enumeradas na aba '*Response'*, listadas conforme Figura 2:

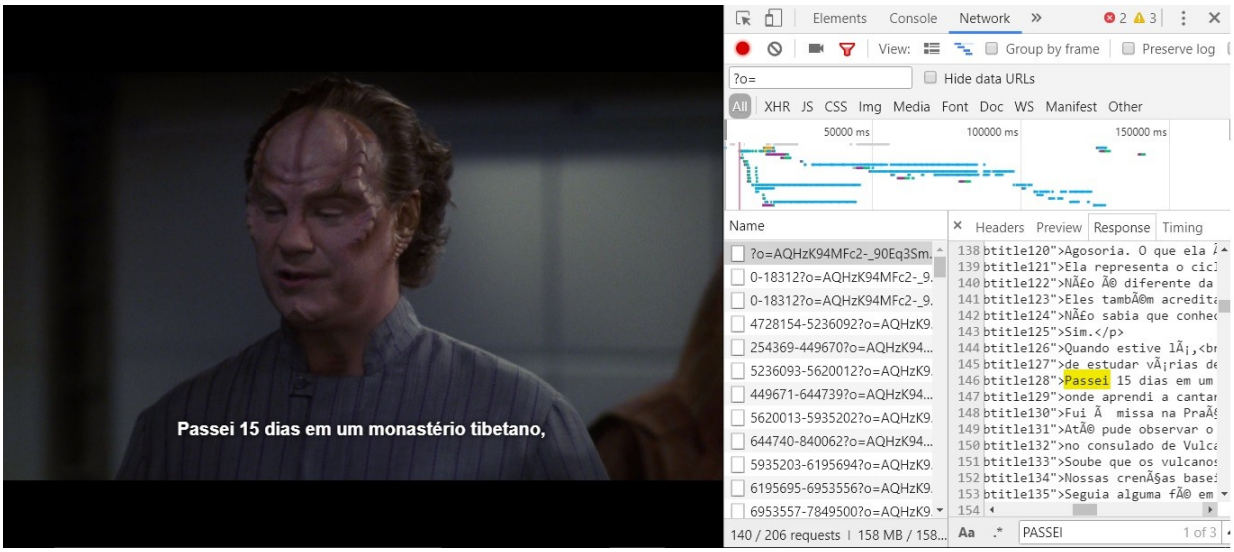


Figura 2: Obtendo as legendas em Português pelo Google Chrome: exemplo de Star Trek, Season 1, Episode 10 (CETAR).
Fonte: Silva (2018, p. 71).

ii) Salvar o arquivo de legendas conforme o passo 7 acima descrito e alterar seu nome, salvando-o em extensão *.xml,* a exemplo de: 'ST_T01_E10_T.xml' (arquivo de legenda da Figura 2).

3.2 Editando os arquivos de legendas

Após obter todos os arquivos de legendas para os episódios escolhidos, é preciso, conforme Silva (2018), fazer a sua edição em um programa de legendagem, como, por exemplo, o *Subtitle Edit*[3]. Nesta etapa metodológica, embora o programa tenha muitas outras funções, ele serviu apenas para eliminar etiquetas de identificação sonora e algumas etiquetas HTML, de todos os arquivos de legendas em inglês e em português, bem como adicionar novas legendas e etiquetas.

3.2.1 Editando os textos fontes

Para edição dos TFs, podem ser seguidos os passos abaixo:

i) Utilizar o programa *Subtitle Edit* para fazer o *upload* do arquivo em *.xml*. A disposição desse arquivo é observada na Figura 3.

3 A versão utilizada por Silva (2018) foi a 3.5.5. No entanto, a versão mais atual é a 3.5.9 e está disponível para *download* gratuito em: http://www.nikse.dk/subtitleedit/. Acesso em: 02 jul. 2019.

http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

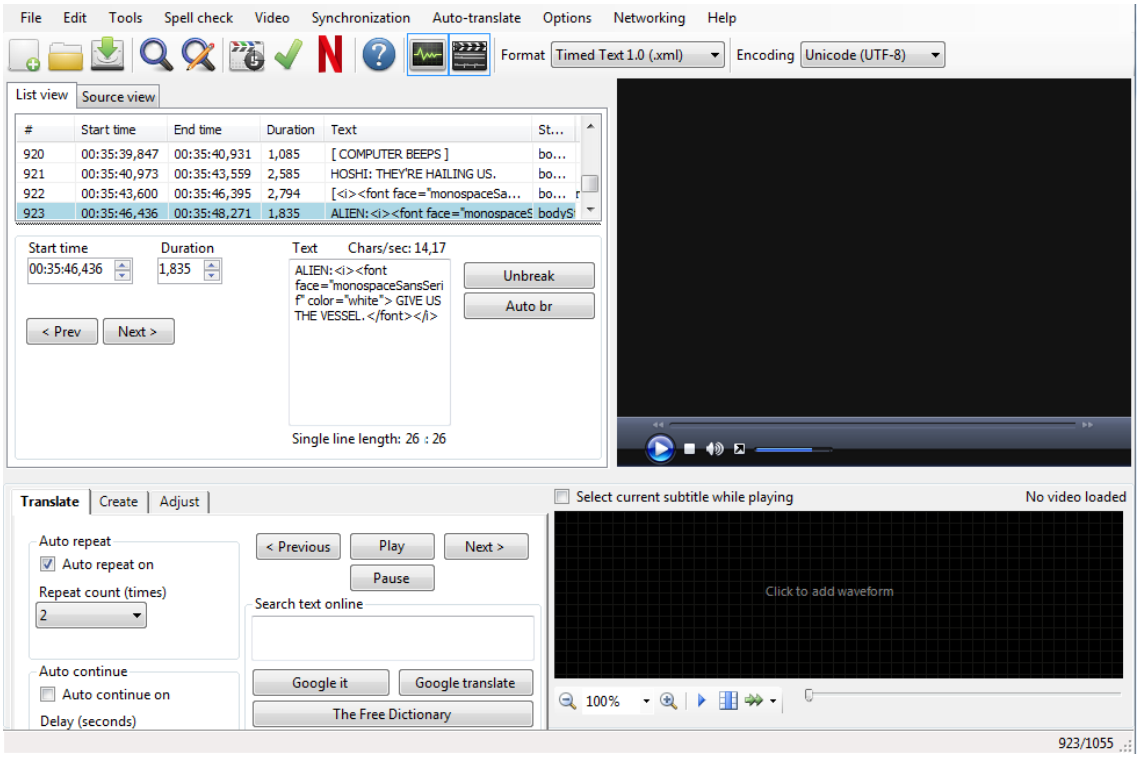


Figura 3: Exibindo o arquivo de transcrição em inglês, em formato de legendas, no Subtitle Edit: exemplo em *Star Trek, Season 2, Episode 16* (CETAR).
Fonte: Silva (2018, p. 73).

ii) Remover, do tipo de legendas para surdos e ensurdecidos, as etiquetas de identificação sonora, tais como textos entre colchetes (ex.: [SCANNER BEEPING]) e que contêm símbolos musicais '♪'. É possível fazer esse procedimento ao acessar a guia '*Remove texts for hearing impaired*' na aba '*Tools*', localizada na parte superior do programa. Conforme visto na Figura 5, a coluna '*Before'* se refere à legenda no modo atual, sem alterações, enquanto a coluna à direita, '*After',* indica como ela ficará, após as devidas mudanças. Ressalta-se que essa mesma guia oferece a opção de excluir o nome de personagens. No entanto, na pesquisa de Silva (2018), por ser relevante manter a identificação do falante na legenda em inglês, uma vez que cada personagem também apresenta um estilo particular de usar a língua, optou-se por não se excluir o texto antes de dois pontos ('*Remove text before a colon'*).

http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

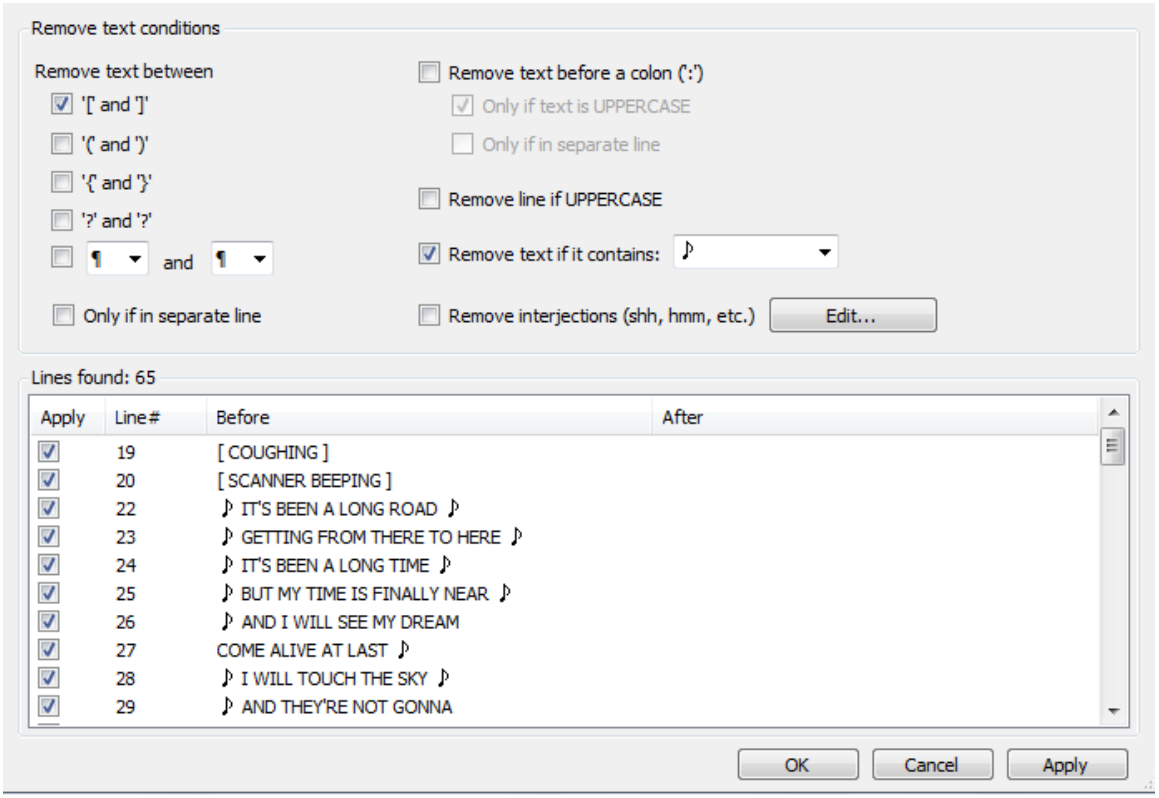


Figura 4: Eliminando a identificação sonora das legendas em inglês: exemplo em *Star Trek, Season 2, Episode 16* (CETAR).
Fonte: Silva (2018, p. 74).

iii) Remover outras etiquetas HTML dos arquivos de legendas que não sejam relevantes para sua análise, tais como *''*, *'*, *''* e *''.* É possível fazer esse procedimento ao acessar a guia '*Multiple replace*' na aba '*Edit'*, localizada na parte superior do programa. A Figura 5 ilustra essa etapa.

http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

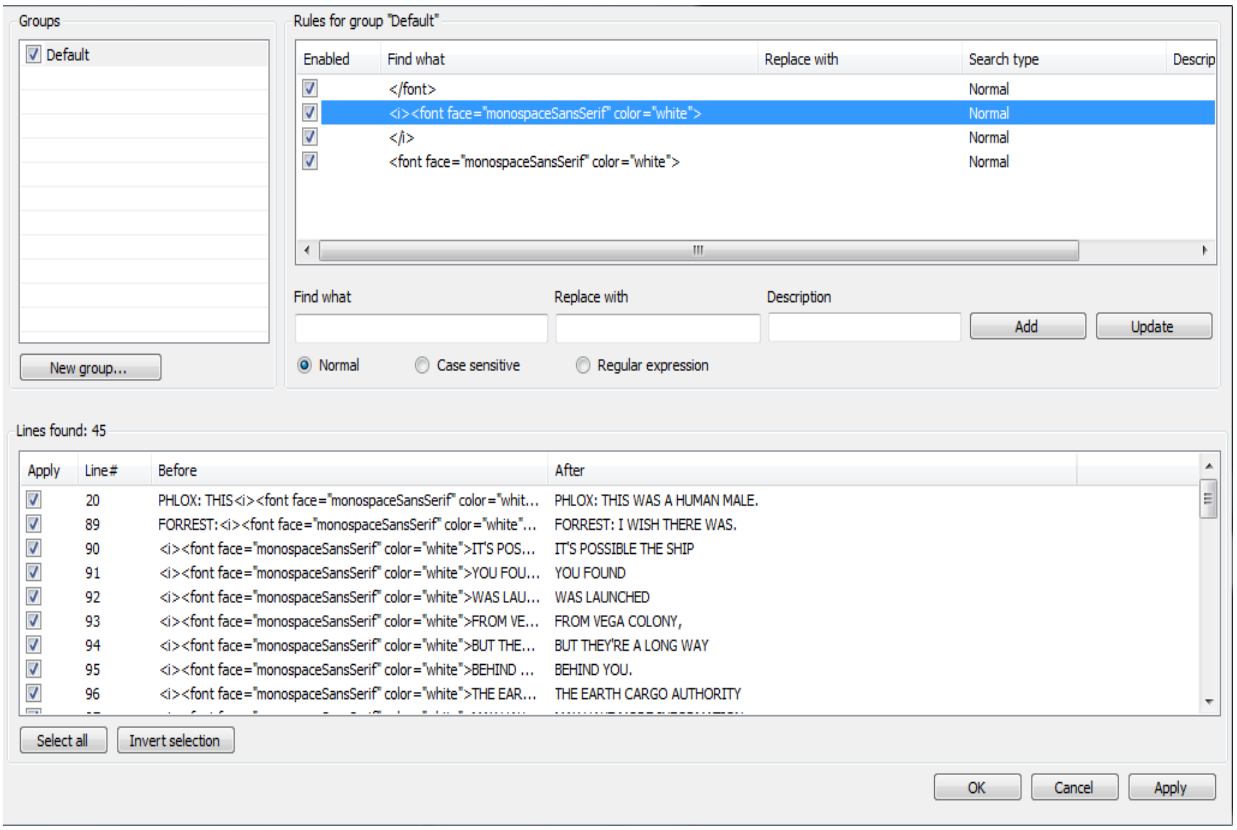


Figura 5: Eliminando etiquetas HTML das legendas em inglês: exemplo em *Star Trek, Season 2, Episode 16* (CETAR).
Fonte: Silva (2018, p. 75).

iv) Colocar o nome de cada falante dentro de um par de angulares. Por exemplo, o nome 'ARCHER:' passaria a ser apresentado como '<ARCHER:>', conforme Figura 6. Essa alteração foi feita no nome de cada personagem e nos demais termos que identificam o falante (ex.: MAN, WOMAN) apenas nos arquivos em inglês, no CETAR e no CROL. Essa mudança foi necessária na pesquisa de Silva (2018), uma vez que se preferiu evitar que o WordList do WST contabilizasse o nome do personagem, pois isso poderia alterar as estatísticas e as listas de palavras e, por conseguinte, os dados finais. Esse procedimento foi feito utilizando-se a função '*Multiple Replace'* explicada no item 3.

http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

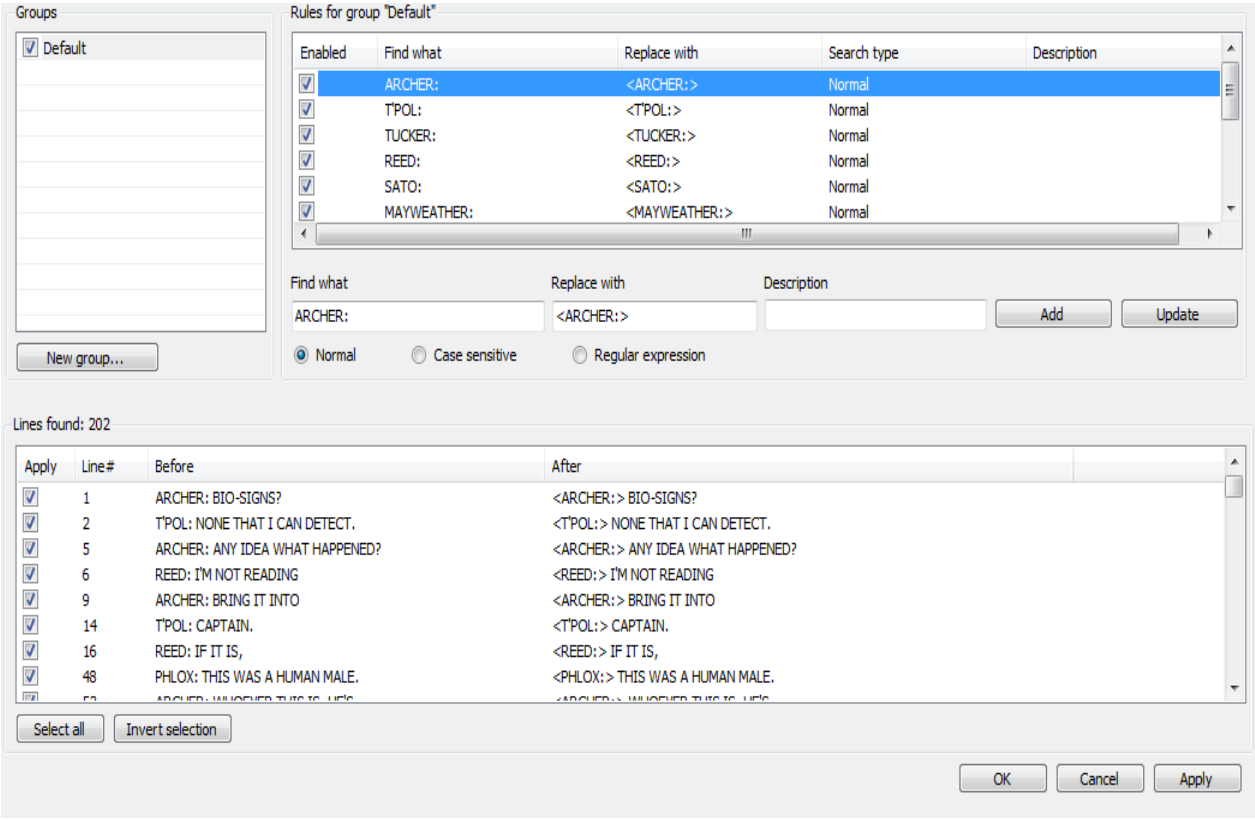


Figura 6: Adicionando angulares aos nomes dos falantes: exemplo em *Star Trek, Season 2, Episode 16* (CETAR).
Fonte: Silva (2018, p. 76).

v) Transcrever o texto disposto na imagem, chamado de intertítulo, que serve como base para a criação da legenda do tipo *FN,* e dispô-lo entre as etiquetas *'<fn>' '</fn>',* conforme visualizado na Figura 7. Esse tipo de legenda, em sua maioria, apresentava o título da série e do episódio. Para realizar essa etapa, foi necessário também adicionar tempos de entrada e saída para cada legenda criada, que, nesses casos, corresponderam ao tempo de sua reprodução no episódio.

http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

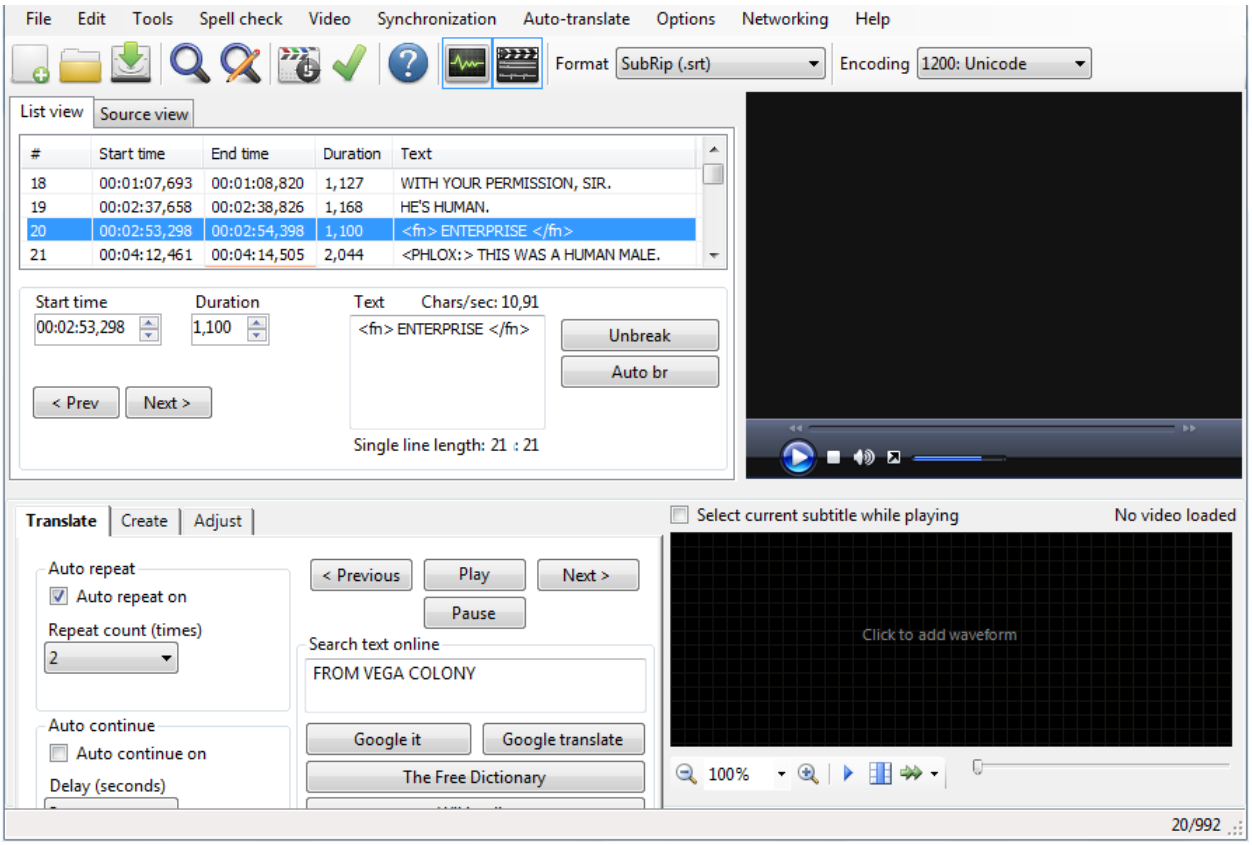


Figura 7: Adicionando textos dispostos nas imagens entre etiquetas: exemplo em *Star Trek, Season 2, Episode 16* (CETAR).
Fonte: Silva (2018, p. 77).

vi) Salvar as legendas em inglês em extensão *.srt* e em padrão *Unicode* após as devidas edições terem sido concluídas. Na Figura 8, observa-se o texto final após as edições feitas.

http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

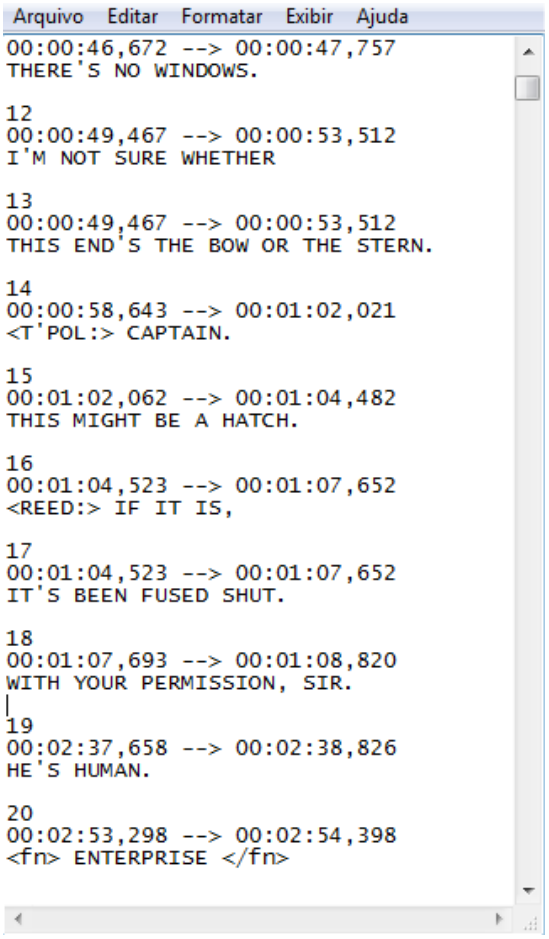


Figura 8: Visualizando no bloco de notas o arquivo de legendas em inglês após edições: exemplo em *Star Trek, Season 2, Episode 16* (CETAR).
Fonte: Silva (2018, p. 78).

3.2.2 Editando os textos traduzidos

Nesta etapa, similarmente à anterior, fez-se uso do Subtitle Edit. Aqui, devido à ausência de identificação sonora, não foi necessário remover etiquetas desse tipo. No entanto, foi preciso excluir algumas etiquetas HTML, pois não seriam necessárias para análise, e adicionar outras, uma vez que dariam mais informações na etapa da análise, como será informado abaixo. Por fim, os arquivos foram salvos em extensão *.srt,* padrão *Unicode.* Os passos estão descritos a seguir:

i) Fazer o *upload* dos arquivos de legendas em português, ainda em extensão *.xml,* no Subtitle Edit*,* conforme visualizado na Figura 9:

http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

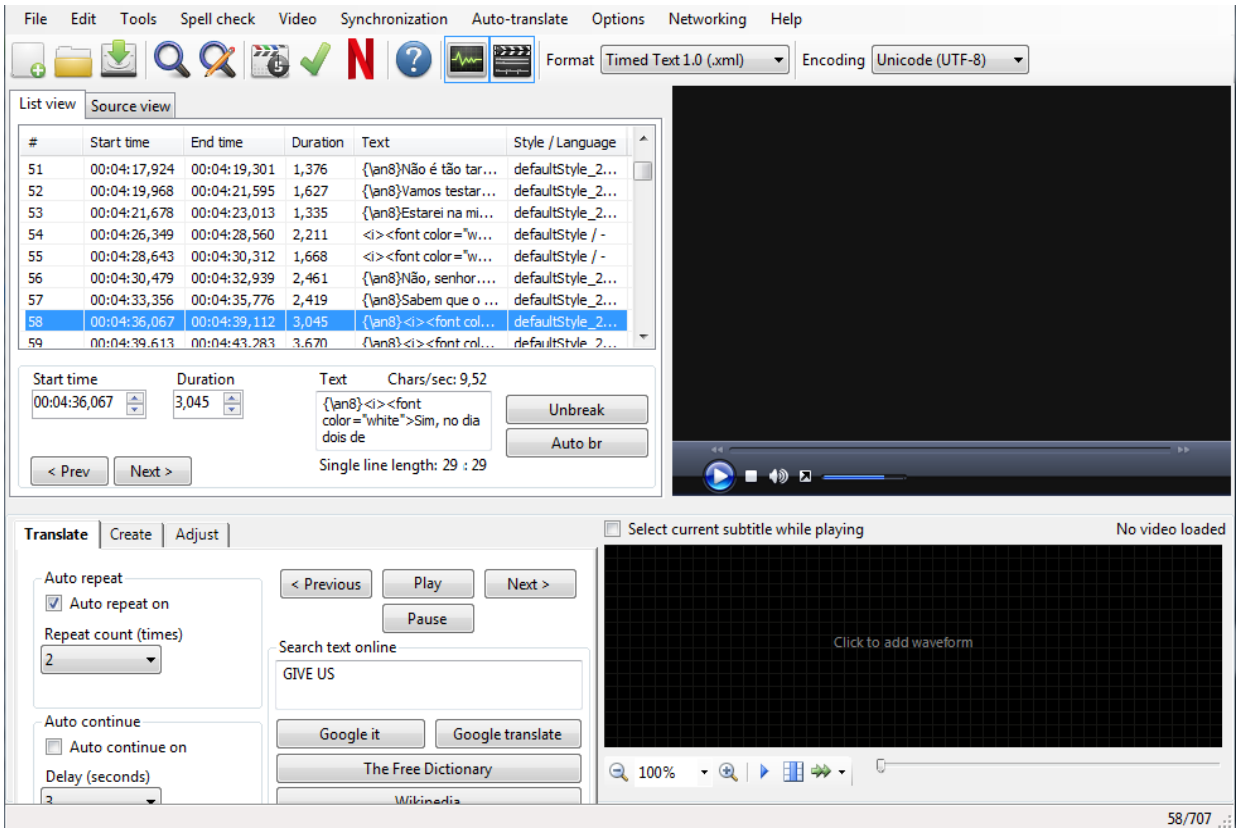


Figura 9: Exibindo o arquivo de legendas em português, em formato de legendas, no Subtitle Edit*:* exemplo em *Star Trek,* Temporada 1, Episódio 11 (CETAR).
Fonte: Silva (2018, p. 79).

ii) Remover etiquetas HTML dos arquivos de legendas, tais como *'{\an8}'*, *''* e *'',* conforme Figura 10*.* Na pesquisa de Silva (2018), não foram removidas as etiquetas *''* e *'',* que indicam o uso do itálico, pois esse recurso, embora usado diferentemente em legendas em comparação a textos literários, pode indicar um aspecto estilístico do tradutor, como visto em pesquisas de outros autores, como Magalhães e Blauth (2015).

http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

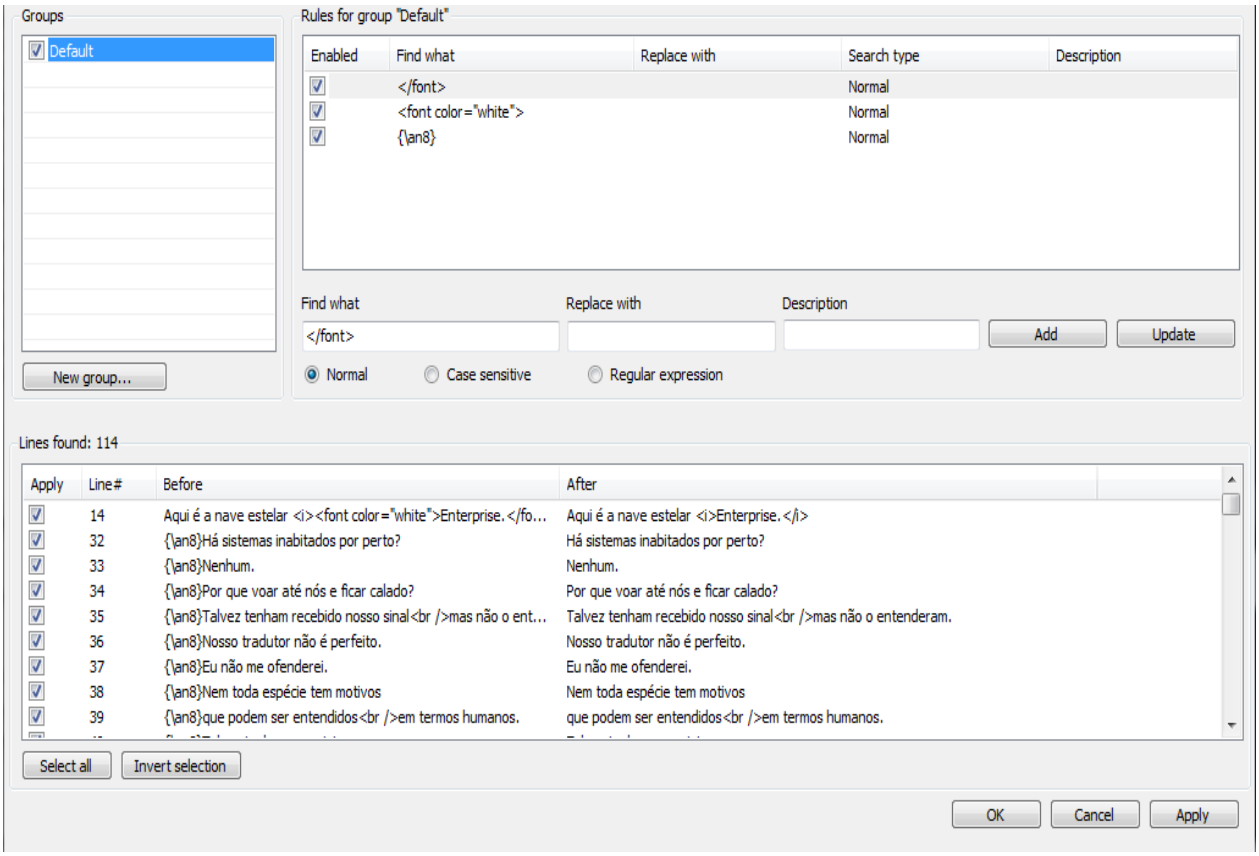


Figura 10: Eliminando etiquetas HTML das legendas em português: exemplo em *Star Trek*, Temporada 1, Episódio 11 (CETAR).
Fonte: Silva (2018, p. 80).

iii) Adicionar etiquetas *'<fn>' '</fn>'* antes e depois das traduções dos intertítulos, respectivamente. Na pesquisa de Silva (2018), esse procedimento foi necessário tendo em vista que as legendas em português eram traduções de outros itens que não somente o áudio, como elementos dispostos na tela. Portanto, foi preciso especificar quais eram os TFs nesses casos, ou seja, as *FN*. Ainda assim, na ferramenta Concord do WST, é possível realizar buscas por etiquetas, como *<fn>* e *</fn>,* para saber, por exemplo, quantas e quais legendas se encaixam nessa especificidade, informações estas que poderiam ser úteis posteriormente. Em sua maioria, na pesquisa supramencionada, os textos *FN* eram frequentemente a tradução do nome da série e de cada episódio, exs.: "JORNADA NAS ESTRELAS: ENTERPRISE"; "INIMIGO SILENCIOSO".

http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

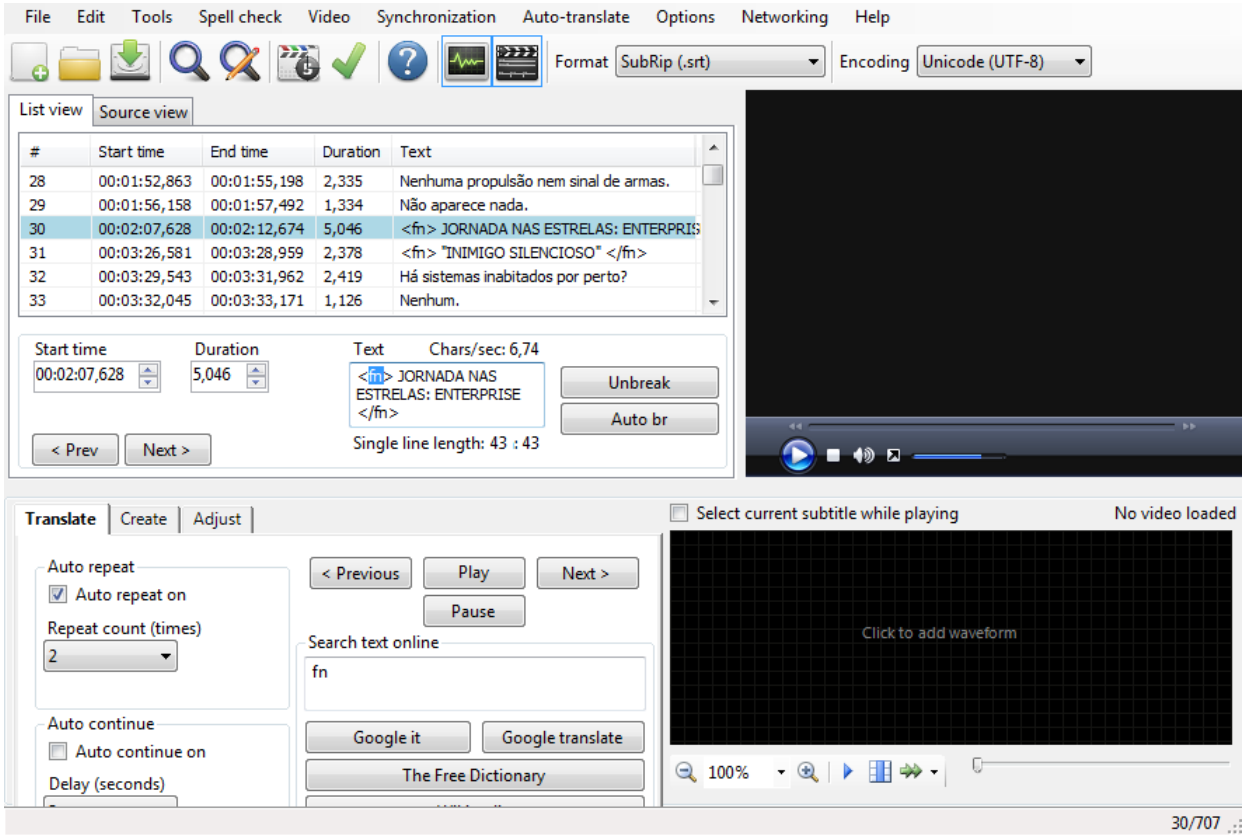


Figura 11: Adicionando etiquetas FN nas legendas em português: exemplo em *Star Trek*, Temporada 1, Episódio 11 (CETAR).
Fonte: Silva (2018, p. 81).

iv) Salvar as legendas em português em extensão *.srt* e em padrão *Unicode*, após as devidas edições terem sido concluídas. Na Figura 12, observa-se o texto final após as edições.

http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

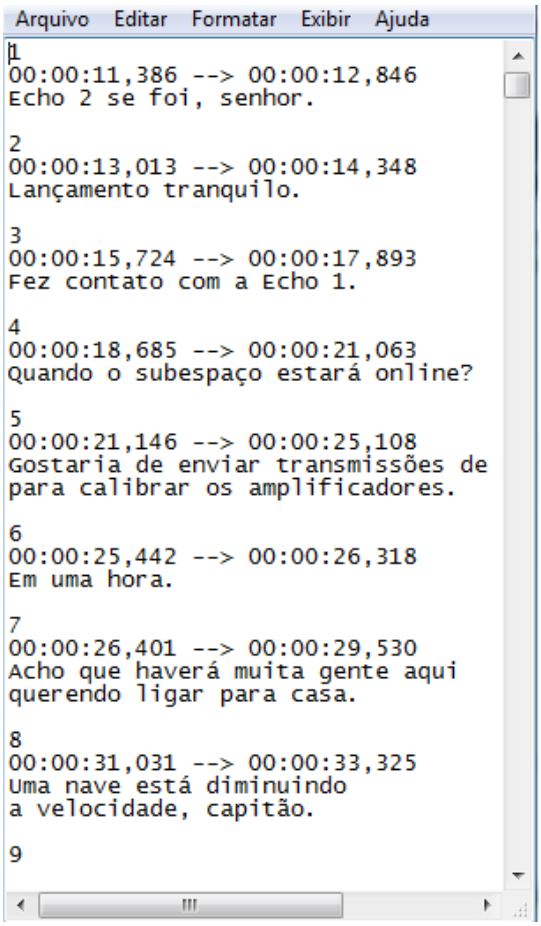


Figura 12: Visualizando no bloco de notas arquivo editado de legendas em português: exemplo em *Star Trek*, Temporada 1, Episódio 11 (CETAR).
Fonte: Silva (2018, p. 82).

3.3 Preparando as legendas

Na investigação de Silva (2018), os passos empreendidos nesta etapa da pesquisa foram seguidos no intuito de preparar cada arquivo de legenda para posterior uso no WST e para análise dos dados. Esses passos consistiram, resumidamente, da edição e formatação das legendas dos arquivos em extensão *.srt* utilizando o Microsoft Office Word 2007 e, *a posteriori*, do alinhamento de alguns textos no Microsoft Office Excel 2007. Na fase de edição e formatação, todos os arquivos de legendas foram processados. No entanto, o alinhamento só se deu nas legendas do CETAR.

3.3.1 Editando e formatando as legendas no Microsoft Office Word 2007

Como observado nas Figuras 8 e 12, que exibem os arquivos de legendas após as edições finais, cada bloco de legenda é organizado em um conjunto de três a quatro linhas. Cada linha corresponde a uma informação específica e cada bloco é separado por um espaço em branco entre si. A primeira linha indica a sequência numérica da legenda, a segunda os tempos de entrada e saída do bloco da legenda e a terceira e quarta as

http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

legendas propriamente ditas. Apesar de essa organização ser bastante útil para melhor visualização das legendas, há outras formas de organização que podem igualmente auxiliar o pesquisador. No presente caso, no intuito de melhor visualizar as legendas e poder trabalhar com elas mais eficientemente no WST e na análise dos dados, editou-se cada arquivo de legenda, em ambas as línguas, no Word 2007, de modo que cada bloco de legenda estivesse em apenas uma linha. Esse procedimento foi feito para cada arquivo de legenda em ambos os *corpora*. O passo a passo dessa etapa encontra-se descrito a seguir, exemplificado pelo TF do Episódio 10, Temporada 1 (CETAR).

i) Copiar o texto do Bloco de Notas e cole-o no Word 2007.
ii) Ativar a marcação de parágrafos, representada pelo símbolo '¶'
iii) Substituir '2 ¶' por '2 ¶ <', conforme Figura 13:

^p^p → ^p^p<

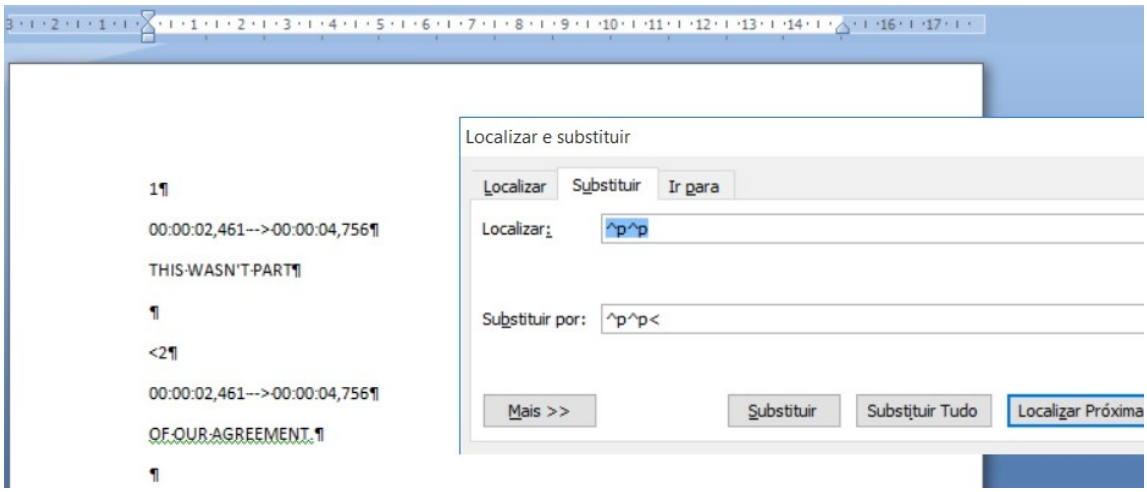


Figura 13: Substituindo '2 parágrafos' por '2 parágrafos <'.
Fonte: Silva (2018, p. 84).

iv) Substituir '2 ¶' por '#', como visto na Figura 14:

^p^p → #

http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

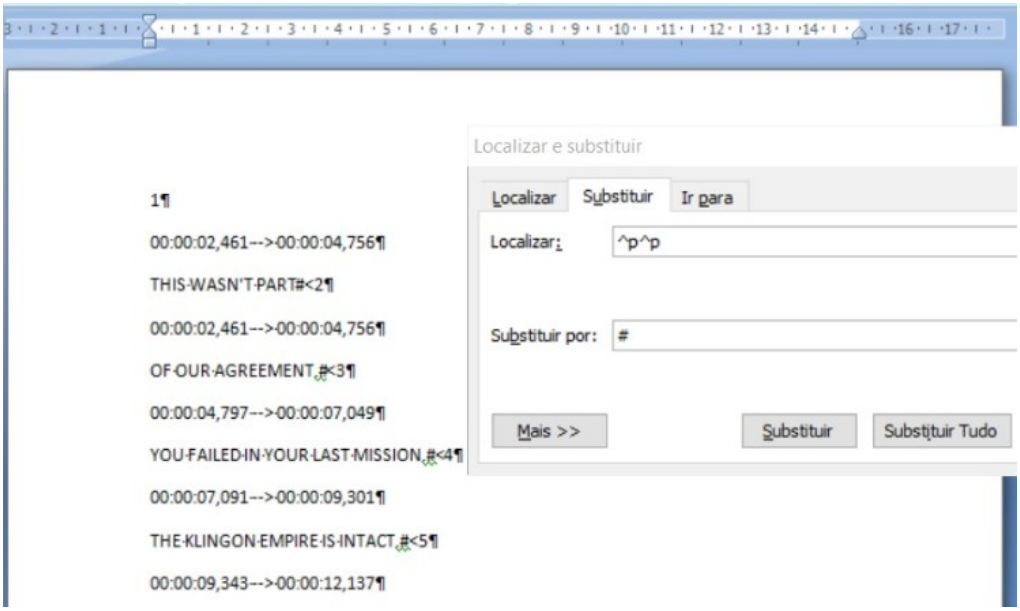


Figura 14: Substituindo '2 parágrafos' por '#'.
Fonte: Silva (2018, p. 84).

v) Substituir '¶ simples' por símbolo de tabulação '→', conforme Figura 15:

^p → ^t

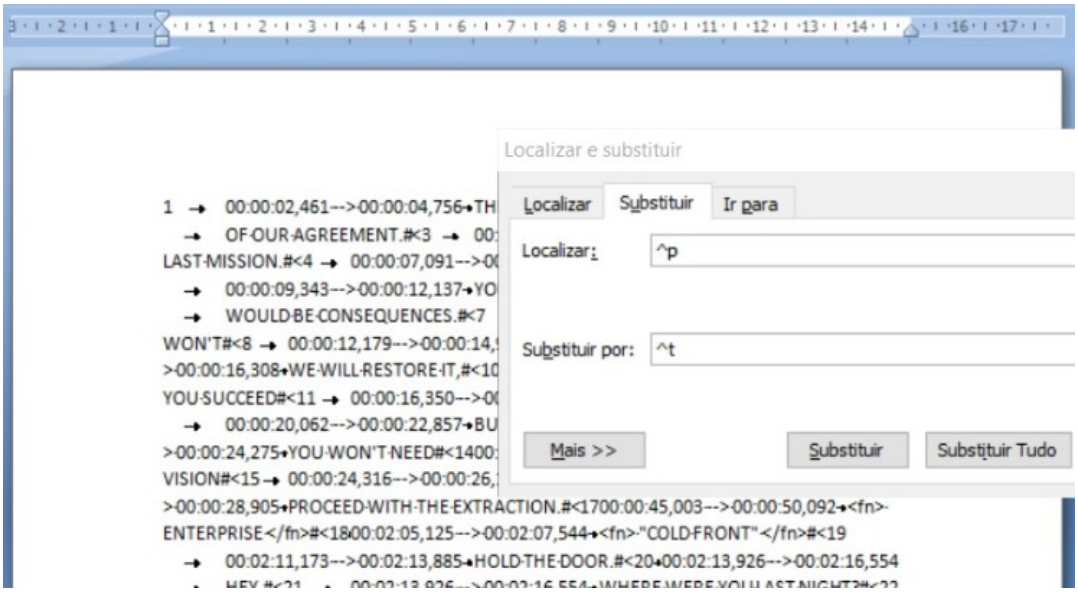


Figura 15: Substituindo 'parágrafo simples' por '^t' (tabulação).
Fonte: Silva (2018, p. 85).

vi) Substituir '#' por '1 ¶ simples', conforme Figura 16:

→ ^p

http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

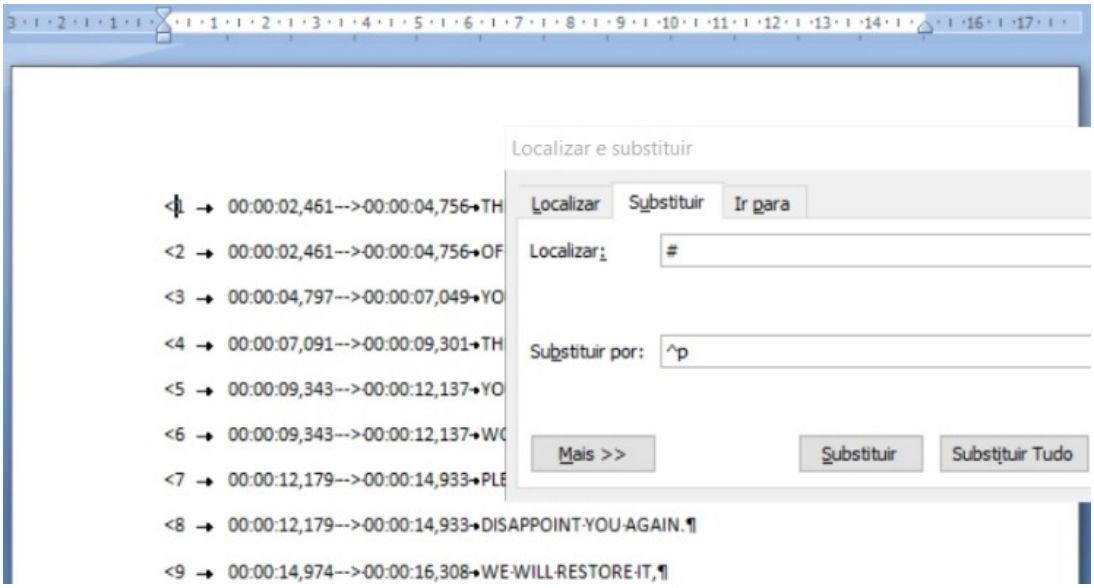


Figura 16: Substituindo '#' por 1 'parágrafo simples'.
Fonte: Silva (2018, p. 85).

vii) Converter texto em tabela ('Inserir' → 'Tabela' → 'Converter texto em tabela'). Selecionar, na última opção 'Texto separado em', o item 'Tabulações'. O texto em 3 colunas é visto conforme Figura 17:

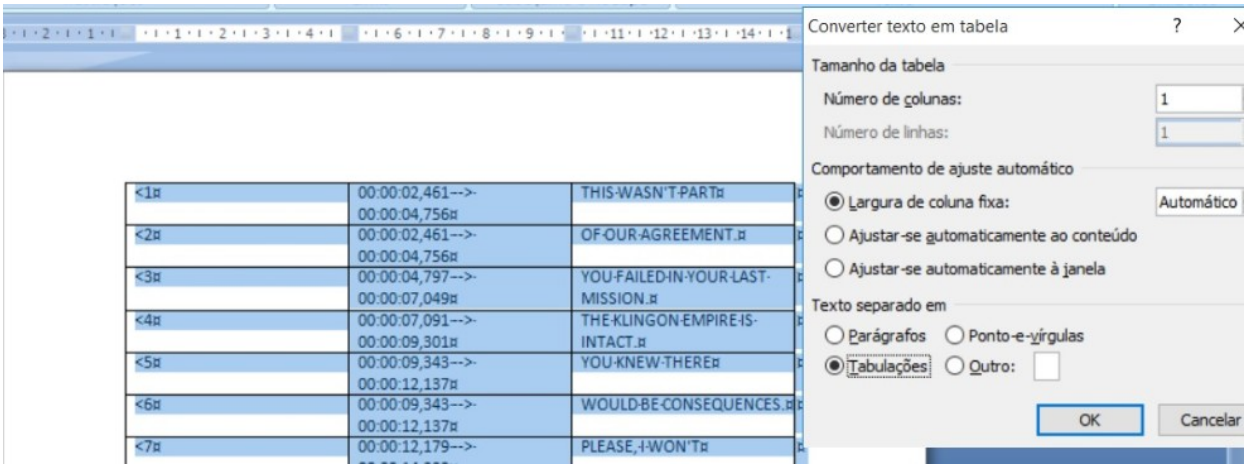


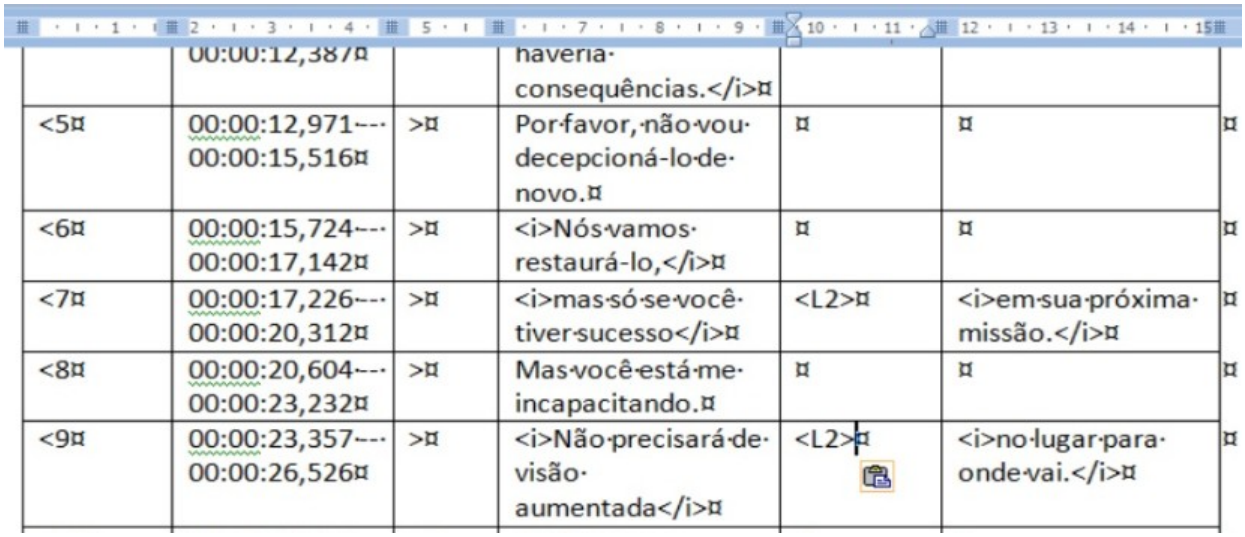
Figura 17: Convertendo texto em tabela.
Fonte: Silva (2018, p. 86).

viii) Adicionar coluna após a minutagem e acrescentar o símbolo de '>' em todas as linhas. No caso da edição das legendas em português, adicionar nova coluna após a 4ª e acrescentar 'L2' para indicar uma segunda linha de legenda ou uma quebra de linha, conforme Figuras 18 e 19.

http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

<1¤	00:00:02,461--> 00:00:04,756¤	>¤	THIS·WASN'T·PART¤
<2¤	00:00:02,461--> 00:00:04,756¤	>¤	OF·OUR·AGREEMENT.¤
<3¤	00:00:04,797--> 00:00:07,049¤	>¤	YOU·FAILED·IN·YOUR·LAST·MISSION.¤
<4¤	00:00:07,091--> 00:00:09,301¤	>¤	THE·KLINGON·EMPIRE·IS·INTACT.¤
<5¤	00:00:09,343--> 00:00:12,137¤	>¤	YOU·KNEW·THERE¤
<6¤	00:00:09,343--> 00:00:12,137¤	>¤	WOULD·BE·CONSEQUENCES.¤
<7¤	00:00:12,179--> 00:00:14,933¤	>¤	PLEASE,·I·WON'T¤
<8¤	00:00:12,179--> 00:00:14,933¤	>¤	DISAPPOINT·YOU·AGAIN.¤
<9¤	00:00:14,974--> 00:00:16,308¤	>¤	WE·WILL·RESTORE·IT,¤

Figura 18: Adicionando coluna após a minutagem.
Fonte: Silva (2018, p. 87).



	00:00:12,387¤		haveria· consequências.¤		
<5¤	00:00:12,971--· 00:00:15,516¤	>¤	Por·favor,·não·vou· decepcioná-lo·de· novo.¤	¤	¤
<6¤	00:00:15,724--· 00:00:17,142¤	>¤	Nós·vamos· restaurá-lo,¤	¤	¤
<7¤	00:00:17,226--· 00:00:20,312¤	>¤	mas·só·se·você· tiver·sucesso¤	<L2>¤	em·sua·próxima· missão.¤
<8¤	00:00:20,604--· 00:00:23,232¤	>¤	Mas·você·está·me· incapacitando.¤	¤	¤
<9¤	00:00:23,357--· 00:00:26,526¤	>¤	Não·precisará·de· visão· aumentada¤	<L2>¤	no·lugar·para· onde·vai.¤

Figura 19: Adicionando nova coluna após a 4ª e acrescentando 'L2' nos arquivos em português.
Fonte: Silva (2018, p. 87).

ix). Substituir '-->' por '--', conforme Figura 20.

http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

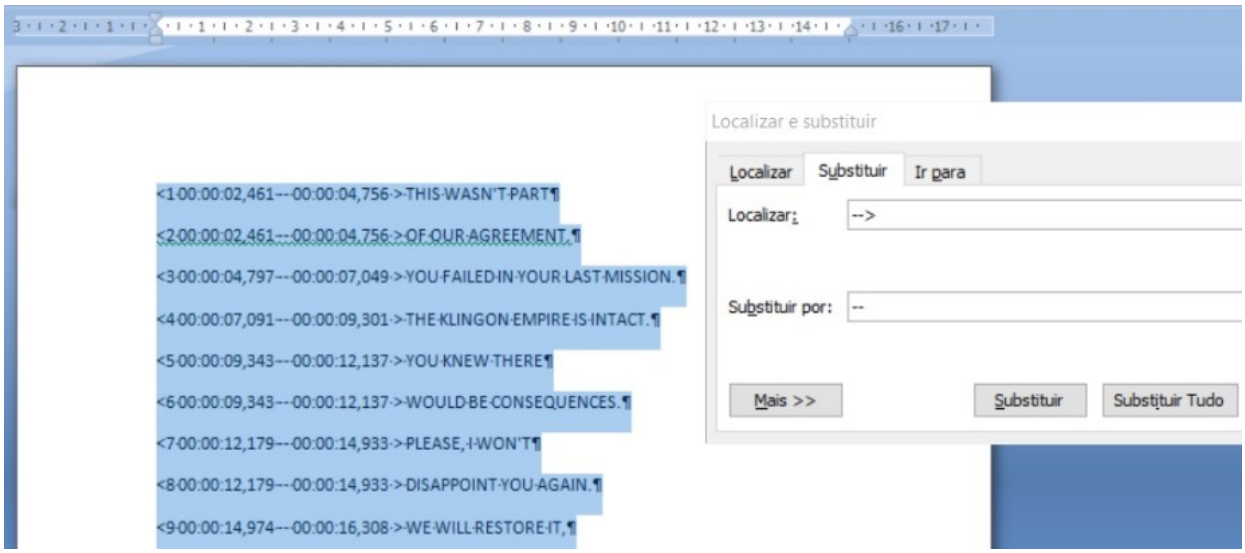
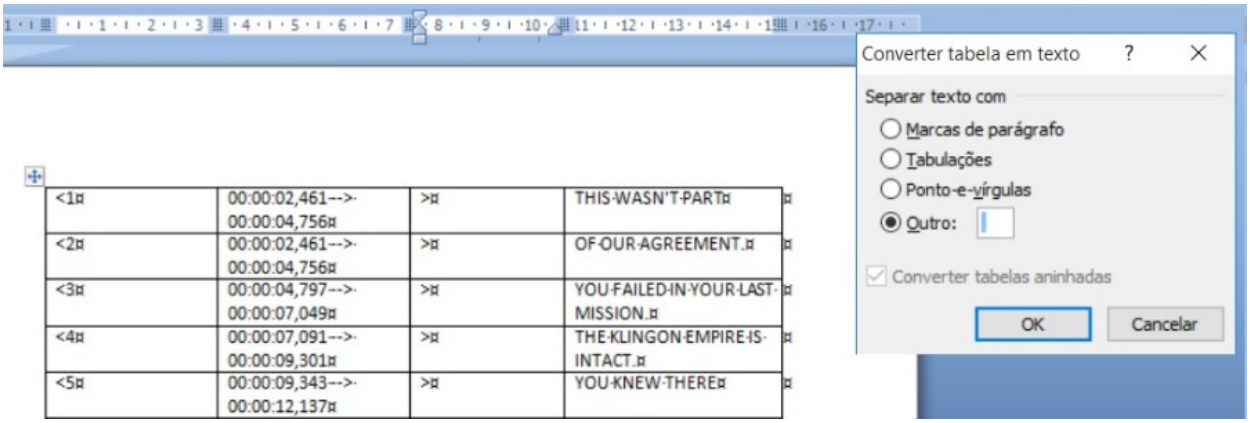


Figura 20: Substituindo '-->' por '--'.
Fonte: Silva (2018, p. 88).

x) Converter a tabela em texto, na aba Layout, 'Converter em Texto'. No item 'Outro', acrescentar um espaço em branco, conforme Figura 21.



<1¤	00:00:02,461--> 00:00:04,756¤	>¤	THIS·WASN'T·PART¤
<2¤	00:00:02,461--> 00:00:04,756¤	>¤	OF·OUR·AGREEMENT.¤
<3¤	00:00:04,797--> 00:00:07,049¤	>¤	YOU·FAILED·IN·YOUR·LAST·MISSION.¤
<4¤	00:00:07,091--> 00:00:09,301¤	>¤	THE·KLINGON·EMPIRE·IS·INTACT.¤
<5¤	00:00:09,343--> 00:00:12,137¤	>¤	YOU·KNEW·THERE¤

Figura 21: Convertendo a tabela em texto.
Fonte: Silva (2018, p. 88).

xi) Salvar o texto resultante em extensão *.txt* (texto sem formatação). Abri-lo no Bloco de Notas e salvá-lo em *Unicode,* conforme visto na Figura 22.

http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

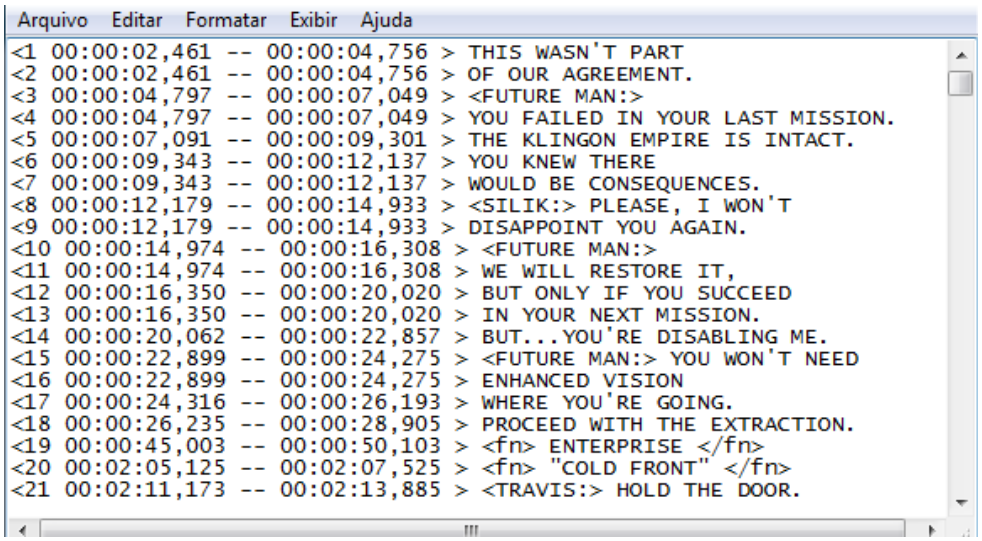


Figura 22: Visualizando o texto em extensão .txt no bloco de notas.
Fonte: Silva (2018, p. 89).

3.3.2 Alinhando as legendas traduzidas por Talita Ribeiro

Na pesquisa de Silva (2018), esta etapa consistiu na criação de um *corpus* paralelo (CPTR) a partir do alinhamento manual dos textos do *corpus* de estudo e seus respectivos TFs, com a utilização do Microsoft Office Excel 2007, conforme exemplo na Figura 23.

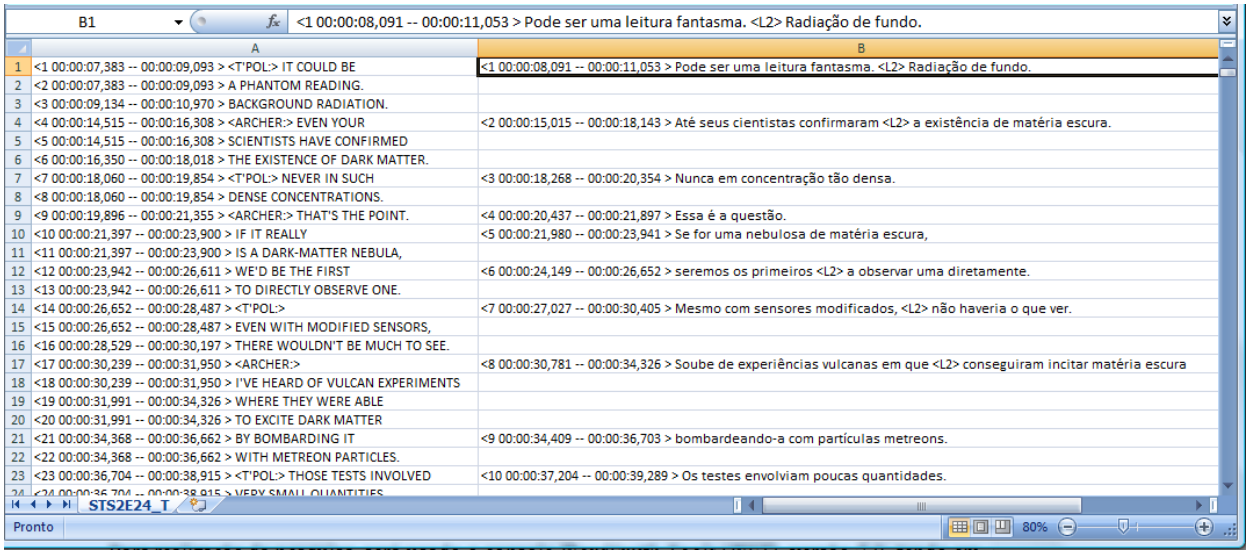


Figura 23: Alinhando os TFs com os TTs: criando o CPTR.
Fonte: Silva (2018, p. 90).

Não foi necessário o alinhamento do *corpus* de referência, na pesquisa de Silva (2018), pois preferiu-se priorizar o *corpus* de estudo, embora esse mesmo programa permitisse um alinhamento manual. Todavia, quando foi preciso contrastar TFs e TTs do *corpus* de referência, uma possibilidade foi abrir dois arquivos no Bloco de Notas, lado a lado. Na pesquisa mencionada, esse *corpus* foi utilizado apenas para checagem manual

http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

das legendas na etapa de análise, pois, ao converter Excel em Word e Word em Texto Sem Formatação, o alinhamento não era disposto organizadamente no WST nesse novo arquivo em *.txt.*

Como observado na Figura 23, a transcrição em inglês ficou disposta na coluna esquerda do programa, enquanto as legendas em português ficaram na coluna direita. Para conseguir essa formatação, alguns passos podem ser seguidos. É preciso copiar do Bloco de Notas as transcrições e as legendas e colá-las em duas colunas no Excel, uma de cada vez. Em seguida, deve-se proceder ao alinhamento manual das legendas de acordo com seu respectivo TF. É importante observar que as legendas em inglês podem ocupar mais linhas por dois motivos: o TF é geralmente maior comparado à quantidade de legendas na língua-alvo; cada fala do personagem no TF, no contexto da pesquisa, estava disposta em mais de um bloco. Para lidar com esse desnível, sugere-se adicionar linhas em branco abaixo das legendas em português, para que assim o TT possa ficar relativamente em paralelo ao TF. Sendo assim, cada legenda em português é colocada na primeira linha correspondente ao início das transcrições em inglês, enquanto que as linhas em branco marcam a continuação dos blocos dos TFs.

Vale salientar que há outras ferramentas disponíveis para alinhamento de TFs e TTs. No entanto, na pesquisa de Silva (2018), tentou-se utilizar outras ferramentas para alinhamento automático, como o Viewer & Aligner, do WST, e o Wordfast Aligner[4]. Contudo, essas ferramentas não produziram alinhamentos confiáveis no caso dos arquivos de legendas. O desafio com esse utilitário do WST é que ele alinha textos apenas no nível da sentença e, como afirma Olohan (2004), nem sempre há correspondência um-a-um entre as sentenças dos textos fonte e traduzido. Esse fenômeno é ainda mais evidente na legendagem, em que uma sentença do TF pode ser traduzida em mais de uma nas legendas, ou mais de uma sentença do TF traduzida em apenas uma. Sendo assim, muitas pós-edições manuais deveriam ser feitas de qualquer modo. Outro desafio para o Wordfast Aligner é que, ao colar os arquivos no *site*, o alinhamento não é feito em todo o texto, apenas em parte dele. Descobriu-se que os erros estavam acontecendo porque o programa ainda não está apto a trabalhar com a minutagem das legendas nem com símbolos como os angulares.

3.3.3 Nomeando os arquivos eletrônicos

Após seguir todos os passos acima e compilar todos esses *corpora*, é relevante nomear cada arquivo com nomes específicos, o que, na pesquisa de Silva (2018), foi feito com siglas, como observado nos Quadros 4 e 5 a seguir.

4 Disponível em: https://www.wordfast.net/?go=align. Acesso em: 06 jul. 2019.

http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

Quadro 4: Listando os nomes dos arquivos eletrônicos do *corpus* de estudo.

***CORPUS* DE ESTUDO**					
TEMPORADA 1			**TEMPORADA 2**		
EPISÓDIOS	**INGLÊS**	**PORTUGUÊS**	**EPISÓDIOS**	**INGLÊS**	**PORTUGUÊS**
10	STS1E10_T	STT1E10_T	16	STS2E16_T	STT2E16_T
11	STS1E11_T	STT1E11_T	24	STS2E24_T	STT2E24_T

Fonte: Silva (2018, p. 91).

Quadro 5: Listando os nomes dos arquivos eletrônicos do *corpus* de referência.

***CORPUS* DE REFERÊNCIA**					
TEMPORADA 1			**TEMPORADA 2**		
EPISÓDIOS	**INGLÊS**	**PORTUGUÊS**	**EPISÓDIOS**	**INGLÊS**	**PORTUGUÊS**
1	STS1E1_O	STT1E1_O	1	STS2E1_O	STT2E1_O
2	STS1E2_O	STT1E2_O	2	STS2E2_O	STT2E2_O
3	STS1E3_O	STT1E3_O	3	STS2E3_O	STT2E3_O
4	STS1E4_O	STT1E4_O	4	STS2E4_O	STT2E4_O
5	STS1E5_O	STT1E5_O	9	STS2E9_O	STT2E9_O
20	STS1E20_O	STT1E20_O	10	STS2E10_O	STT2E10_O
21	STS1E21_O	STT1E21_O	11	STS2E11_O	STT2E11_O
22	STS1E22_O	STT1E22_O	18	STS2E18_O	STT2E18_O
23	STS1E23_O	STT1E23_O	19	STS2E19_O	STT2E19_O
24	STS1E24_O	STT1E24_O	20	STS2E20_O	STT2E20_O
25	STS1E25_O	STT1E25_O	21	STS2E21_O	STT2E21_O
			26	STS2E26_O	STT2E26_O

Fonte: Silva (2018, p. 91).

Legendas das siglas:

ST = *Star Trek*
S = *Season* (número da temporada)
T = Temporada
E = *Episode* / Episódio (número do episódio)
T = Talita
O = Outros legendistas

Em atenção às siglas criadas e exibidas nos dois últimos quadros, cada uma segue uma lógica de abreviação. As duas primeiras letras (ST) referem-se ao nome da série (*Star Trek*); a terceira letra, em combinação com um número, dizem respeito à temporada

http://periodicos.letras.ufmg.br/index.php/textolivre
Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

(*Season*) e seu respectivo número; a quarta letra, em combinação com um número, dizem respeito ao episódio (*Episode*) e seu número; a última letra, seguida do tracejado (_), refere-se à autoria das legendas daquele episódio, seja Talita (T), sejam os outros legendistas (O). Vale salientar que, para diferenciar os arquivos dos *corpora* em inglês e português, a marcação utilizada deu-se apenas na letra referente à temporada: 'S' para *season*, em inglês; 'T' para temporada, em português. O exemplo a seguir ilustra essa organização: 'STS1E10_T' se refere ao Episódio 10 da Temporada 1 da Série *Star Trek: Enterprise*, legendado por Talita. Outro exemplo remete à sigla STS1E1_O, que corresponde ao Episódio 1 da Temporada 1 da Série *Star Trek: Enterprise*, legendado por outros legendistas. Essas siglas auxiliam especialmente na análise, para poder identificar de onde os excertos linguísticos são retirados, por exemplo.

Tendo essa seção sido finalizada, o tópico a seguir confere as considerações finais deste artigo.

4 Considerações finais

O presente artigo teve como objetivo descrever o passo a passo metodológico da criação de *corpora* de legendas, em PB e em inglês americano, de uma série de TV, *Star Trek: Enterprise,* na pesquisa de Silva (2018). Salienta-se, aqui, que a intenção foi a de oferecer um possível modelo metodológico que, ao fazer uso de diferentes *softwares,* possa servir a futuros pesquisadores em suas investigações, principalmente naquelas que estabeleçam relações entre TAV e ETBC.

A investigação ora anunciada buscou, por intermédio dos *corpora* de legendas produzidos e ilustrados neste artigo, realizar um estudo das peculiaridades linguísticas e tradutórias encontradas em material produzido por uma legendista em particular, as quais apontaram para a constituição de um estilo próprio. A principal vantagem em utilizar os *corpora* de estudo e de referência é de estabelecer pontos de contraste entre eles para assim poder analisar determinadas peculiaridades linguísticas. Outro fator relevante foi a criação e utilização de um *corpus* paralelo, o CPTR, por meio do qual foi possível ter acesso aos TFs. Esse tipo de *corpus* é de extrema importância em pesquisas sobre estilo do tradutor, pois, quando há comparação entre TFs e TTs, é possível identificar prováveis influências nas escolhas feitas pelos legendistas e explorar, quando possível, explicações para tais escolhas.

A ilustração dos caminhos a serem seguidos neste artigo apontam para a utilização de programas já conhecidos por usuários que usam dispositivos informatizados para realizarem estudos, pesquisas, ou até mesmo navegações básicas, como é o caso do Google Chrome, Bloco de Notas, Subtitle Edit e do Microsoft Word e Excel. Algumas das funções desses programas permitem a criação de *corpora* de legendas e auxiliam o pesquisador no desenrolar metodológico de suas investigações. Após manuseio desses *softwares* e compilação dos *corpora*, é então possível utilizar ferramentas mais específicas da LC, como o WordList e do Concord, do WST, mencionados em Silva (2018).

Por fim, o passo a passo metodológico aqui descrito resultou, portanto, na criação de *corpora* de legendas que, por enquadrarem-se no campo dos ETBC, conforme Olohan (2004) observa, podem apresentar certas características recorrentes e levar

Belo Horizonte, v. 12, n. 3, p. 168-195, set.-dez. 2019 – ISSN 1983-3652
DOI: 10.17851/1983-3652.12.3.168-195

pesquisadores a, dentre várias possibilidades: terem interesse no estudo descritivo de TTs, na forma em que a língua é usada nesses textos, que pode revelar o que é provável e típico da tradução e, a partir disso, interpretar o que é incomum; realizarem a combinação de análises quantitativa e qualitativa na descrição dos dados, focando-se em características lexicais, sintáticas e discursivas; entre outras. Sendo assim, tais instruções podem servir como possível caminho para o desenvolvimento de *corpora* de legendas em pesquisas nas áreas apresentadas.

Agradecimentos

Agradecemos à CAPES o financiamento da pesquisa; a Alessandra Harden e a Elisa D. Teixeira, a orientação deste artigo; e a Talita Ribeiro, tradutora que produziu as legendas utilizadas como base para construção dos *corpora* ora apresentados.

Referências

BAKER, M. Corpora in translation studies: an overview and some suggestions for future research. *Target*, Amsterdam/Philadelphia, v. 7, n. 2, p. 223-243, 1995.

BAKER, M. Corpus-based translation studies: the challenges that lie ahead. In: SOMERS, H. (ed.) *Terminology, LSP and translation:* studies in language engineering in honour of Juan C. Sager. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1996. p. 177-186.

BERBER SARDINHA, T. Comparing corpora with WordSmith Tools: how large must the reference corpus be? In: Annual Meeting of the Association for Computational Linguistics, 38, 2000, Hong Kong. *Proceedings of the Workshop on Comparing Corpora…* Hong Kong: Hong Kong University of Science and Technology, 2000. p. 7-13. Disponível em: https://dl.acm.org/citation.cfm?id=1117731. Acesso em: 26 mar. 2018.

BERBER SARDINHA, T. *Lingüística de corpus*. Barueri: Manole, 2004.

CAMARGO, D. C. *Metodologia de pesquisa em tradução e linguística de corpus.* São Paulo: Cultura Acadêmica/São José do Rio Preto: Laboratório Editorial do IBILCE/UNESP, 2007.

CHAUME, F. Film Studies and Translation Studies: two disciplines at stake in audiovisual translation. *Meta: Translators' Journal,* Montréal, v. 49, n. 1, p. 12-24, 2004.

DÍAZ CINTAS, J.; REMAEL, A. *Audiovisual translation:* subtitling. Manchester: St. Jerome, 2007.

GEORGAKOPOULOU, P. Subtitling for the DVD industry. In: DÍAZ CINTAS, J.; ANDERMAN, G. (org.). *Audiovisual translation:* language transfer on screen. Great-Britain: Palgrave Macmillan, 2009. p. 21-36.

GOTTLIEB, H. Subtitling: diagonal translation. *Perspectives:* studies in Translatology, UK,

v. 2, n.1, p. 101-121, 1994.

GOTTLIEB, H. Subtitling. In: BAKER, M. (ed.). *Routledge Encyclopedia of Translation Studies.* London/New York: Routledge, 1998. p. 244-248.

GOTTLIEB, H. Multidimensional translation: semantics turned semiotics. In: MuTra: Challenges of Multidimensional Translation, 1, 2005, Saarbrücken. *Conference proceedings...* Saarbrücken: Saarland University, 2005a. p. 1-29. Disponível em: http://www.euroconferences.info/proceedings/2005_Proceedings/2005_Gottlieb_Henrik.pdf. Acesso em: 22 fev. 2017.

GOTTLIEB, H. Texts, translation and subtitling – in theory, and in Denmark. In: GOTTLIEB, H. (ed.). *Screen Translation.* Eight studies in subtitling, dubbing and voice-over. Copenhagen: University of Copenhagen, 2005b. p. 1-40.

IVARSSON, J.; CARROLL, M. *Subtitling.* Simrishamn: TransEdit, 1998.

MAGALHÃES, C. M.; BLAUTH, T. P. Estilo do tradutor: um estudo do uso do itálico, palavras estrangeiras e itens culturais específicos por seis tradutores do português de Heart of Darkness. In: VIANA, V.; TAGNIN, S. (org.). *Corpora na tradução*. 1. ed. Paulo: Hub Editoria, 2015. p. 171-209.

MITTELL, J. Narrative complexity in contemporary American television. *The Velvet Light Trap,* Austin, n. 58, p. 29-40, 2006.

OLOHAN, M. *Introducing corpora in Translation Studies*. London/New York: Routledge, 2004.

SCOTT, Mike. *WordSmith Tools manual.* Stroud: Lexical Analysis Software Ltd., 2018.

SILVA, J. M. V. da. *Que espaço a legendista ocupa? Um estudo sobre estilo do tradutor*. 2018. 176f. Dissertação (Mestrado em Estudos da Tradução) – Universidade de Brasília, Brasília, 2018.

STAR TREK: Enterprise (Season 1). Produção de Rick Berman e Brannon Braga. Hollywood: Paramount Network Television, 2001-2002.

STAR TREK: Enterprise (Season 2). Produção de Rick Berman e Brannon Braga. Hollywood: Paramount Network Television, 2002-2003.

Recebido em dia 02 de julho de 2019.
Aprovado em dia 20 de outubro de 2019.