

LEGENDAGEM PARA SURDOS E ENSURDECIDOS (LSE): A SEGMENTAÇÃO NA LSE DE FILMES BRASILEIROS EM DVD.

CHAVES, Élide Gama
ARAÚJO, Vera Lúcia Santiago
POSLA - UECE

Resumo: A tradução audiovisual tem recorrido aos estudos baseados em corpora com o objetivo de encontrar regularidades presentes no corpus bem como sistematizar padrões de análises de modo rápido e preciso. Pesquisas em legendagem para surdos e ensurdecidos (LSE) e audiodescrição já demonstraram a eficiência dessa metodologia por meio da descrição sistematizada de corpora, como os estudos realizados por Hurtado (2007), Salway (2007), Kalantzi (2008), Perego (2003 e 2009), Chaves (2009), e Feitosa (2010). Para ampliar as discussões no âmbito da tradução audiovisual e da lingüística de corpus proponho descrever os problemas de segmentação encontrados na LSE de filmes brasileiros em DVD com auxílio do programa de análise lexical *Wordsmith Tools 5.0*. Segundo Diaz Cintas e Remael (2007) a segmentação na legendagem diz respeito à divisão dos diálogos, da fala, etc. em seções ou segmentos, chamados legendas. Se a segmentação não for cuidadosa, os leitores farão esforço para decodificar o texto e poderão se cansar mais rapidamente e perder o prazer proporcionado pelo programa. Os resultados demonstram problemas de segmentação entre linhas numa mesma legenda e entre duas ou mais legendas.

Palavras chave: Legendagem para surdos e ensurdecidos (LSE), Segmentação, Linguística de corpus.

INTRODUÇÃO

As pesquisas em tradução audiovisual (TAV), mais especificamente em legendagem para surdos e ensurdecidos (LSE), tem demonstrado que a LSE não se limita a uma transcrição de diálogos, efeitos sonoros, sons do ambiente e demais informações que não podem ser compreendidos por pessoas com deficiência auditiva como preconizado pela portaria nº 310, de 27 de junho de 2006. A LSE é considerada tradução, por estudiosos da área, e sofre as influencias das suas restrições assim como qualquer outra modalidade. As pesquisas de (FRANCO e ARAÚJO, 2003; ARAÚJO, 2004, 2005 e 2007) objetivaram encontrar quais seriam os parâmetros de legendagem a serem utilizados para que o surdo brasileiro pudesse assistir confortavelmente a produções audiovisuais. Os resultados geraram um modelo de legendagem, definido por surdos do CAS de Fortaleza (Centro de Atendimento ao Surdo), que está sendo testado em todas as regiões do Brasil.

O modelo de legenda definido pelos surdos do CAS de Fortaleza se configura da seguinte forma: legendas condensadas, velocidade da legenda de 145 ppm, legendas de no máximo duas linhas, cor amarela com borda preta, segmentação (divisão da legenda) pelos cortes das cenas, pelo fluxo da fala e pela sintaxe, a identificação de falante e de efeito sonoro deve ser feita sempre que necessário pelo uso de colchetes, dentre outros (ARAÚJO, 2007).

Ainda de acordo com pesquisas em LSE no Brasil, o parâmetro da segmentação na legendagem é um dos fatores que contribui para a boa recepção de legendas por parte dos surdos. De acordo Reid (1990) a segmentação pode ocorrer de varias formas: visual, ou seja, sempre que há um corte ou uma mudança de cena, uma nova legenda deve ser confeccionada. Quando não há corte, procura-se seguir o fluxo da fala, caracterizando uma segmentação retórica, em que a pausa dos falantes determina o fim de determinada legenda. Finalmente, entre cortes e pausas no diálogo, o parâmetro gramatical deve ser adotado. Cada legenda deve ser um todo coerente com unidades semânticas, que não podem ser separadas, sob pena de o espectador não conseguir integrar os diálogos, o

texto da legenda e o áudio do filme (REID, 1990, p. 100). Segundo Diaz Cintas e Remael (2007) a segmentação na legendagem diz respeito à divisão dos diálogos, da fala, etc. em seções ou segmentos, chamados legendas. Se a segmentação não for cuidadosa, os leitores farão esforço para decodificar o texto e como consequência disso poderão se cansar mais rapidamente e perder o prazer proporcionado pelo filme ou programa. As legendas precisam ter um pensamento completo para serem entendidas no curto espaço de tempo em que são exibidas. Karamitroglou (1998) propõe que a legenda deve ser segmentada no mais alto nível sintático possível. Isso significa que cada inserção de legenda deve conter uma sentença completa. Quando isso não é possível e a sentença tem que continuar na segunda linha de legenda ou mesmo em outra legenda, a segmentação de cada linha deve ser distribuída de modo a coincidir com o mais alto nível sintático possível.

A partir do proposto por Karamitroglou (1998), Perego (2008) analisa os casos de má segmentação gramatical em legendas diversas, pautada nos mais altos níveis sintáticos. Perego (2008) encontrou casos de má segmentação no corpus, nos quais as estruturas internas das unidades sintáticas foram divididas. Através dos casos de má segmentação encontrados foram definidas as seguintes categorias: quebra do sintagma nominal, quebra do sintagma preposicionado, quebra do sintagma verbal, quebra das orações coordenada e subordinada.

Baseando-se nas categorias de má segmentação propostas por Perego (2008) foi possível criar etiquetas discursivas para análise de segmentação em LSE, que serão mostradas mais adiante. As etiquetas possibilitam encontrar a ocorrência e a frequência com que cada categoria gramatical aparece no corpus. Portanto diante desse panorama sobre a legendagem para surdos no Brasil e diante da importância da segmentação para a boa recepção às legendas, é que proponho analisar a segmentação das legendas para surdos lançadas em DVD do filme brasileiro *Nosso Lar* (2010) com auxílio do programa de análise lexical *Wordsmith Tools 5.0*. O presente estudo apresenta os resultados iniciais de uma pesquisa maior em desenvolvimento, minha pesquisa de mestrado, na qual o corpus compreende 4 legendagens de filmes brasileiros em DVD, são eles: *Irmãos de Fé* (Drama, Moacyr Góes, 2004), *Nosso Lar* (Drama, Wagner de Assis, 2010), *Corisco e Dadá* (Drama, Rosemberg Cariry, 1996) e *O Grão* (Drama, Petrus Cariry, 2010).

Esta pesquisa está fundamentada pelos estudos descritivos da tradução audiovisual e pelos estudos baseados em corpora. Como nova tendência dos estudos descritivos de tradução a tradução audiovisual tem recorrido aos estudos baseados em corpora com o objetivo de encontrar regularidades presentes no corpus bem como sistematizar padrões de análises de modo rápido e preciso. Pesquisas em legendagem para surdos e ensurdecidos (LSE) e audiodescrição para deficientes visuais já demonstraram a eficiência dessa metodologia por meio da descrição sistematizada de corpora, como os estudos realizados por Hurtado (2007), Salway (2007), Kalantzi (2008), Perego (2003 e 2009), Chaves (2009), e Feitosa (2010).

METODOLOGIA

Contexto da pesquisa

A presente pesquisa pode ser definida como descritiva, de natureza qualitativa. Com relação aos procedimentos se caracteriza como baseada em corpus por envolver procedimentos característicos como: anotação do corpus (etiquetagem) e uso de programas de análise lexical como os concordanciadores. A pesquisa conta com

apoio financeiro da CAPES e está sendo desenvolvida no LATAV – Laboratório de Tradução Audiovisual da UECE onde são realizadas pesquisas em Acessibilidade Audiovisual pelo grupo de estudos LEAD (Legendagem e Audiodescrição) vinculado ao grupo de pesquisa *Tradução e Semiótica*, registrado no CNPq sob a coordenação da Profa. Vera Lúcia Santiago Araújo.

Instrumentos

Serão instrumentos da pesquisa os programas de computador: *SubRip* – para extração das legendas dos filmes em DVD; *Word* – para edição de texto e anotação do corpus; *WordSmith Tools 5.0* - programa de análise lexical para auxílio na análise dos dados.

Tratamento do corpus

O tratamento do corpus diz respeito aos procedimentos que antecedem e possibilitam a análise dos dados, são eles: extração das legendas ou compilação do corpus, etiquetagem e por último, análise dos dados a partir da interpretação e utilização do programa de análise lexical WordSmith Tools 5.0.

Extração das Legendas

Primeiramente as legendas foram extraídas via OCR (*Optical Character Recognition*) com auxílio do programa *SubRip*, para o formato “.srt”, que permite rápida identificação de todas as informações da legenda (número da legenda, tempos de entrada e saída, e legenda propriamente dita) quando aberto com o *Bloco de Notas* do *Windows*. A partir dos arquivos de legendas em “.srt” é possível fazer a anotação do corpus, procedimento explicado mais adiante.

Etiquetagem

A etiquetagem, processo autoral deste trabalho, auxilia na desambiguação lexical e permite a descrição de padrões léxico-gramaticais. Em outras palavras a etiquetagem consiste numa codificação do corpus passível ou não de análise eletrônica através de um concordanciador (BERBER SARDINHA, 2004, p. 145). Para esta pesquisa o concordanciador utilizado foi o *Concord* do conjunto de ferramentas do WordSmith Tools 5.0. A etiquetagem deve ser feita com auxílio de editores de texto e salva em formato “.txt” (texto sem formatação), podendo desta forma ser compreendida pelo *WordSmith Tools*. Como o corpus desta pesquisa são legendas em formato “.srt” tive que abrir os arquivos “.srt” com o editor de textos *Word* para começar a etiquetagem. Neste caso utilizei o editor de textos *Word* por considerá-lo de fácil e prático manuseio, mas poderia ter optado pelo *Bloco de Notas* ou qualquer outro editor de textos.

As etiquetas foram criadas a partir das categorias de análise da má segmentação textual em legendas, desenvolvidas por Perego (2008), que analisa a quebra das seguintes estruturas sintáticas: 1) Sintagma Nominal, 2) Sintagma Preposicionado, 3) Sintagma Verbal, 4) Oração Coordenada, 5) Oração Subordinada. A partir dessas categorias pude criar as seguintes etiquetas:

- 1) Para análise de sintagma nominal o modelo padrão se configura da seguinte forma <*Sintagma Nominal_estrutura sintática quebrada*> e os modelos específicos:

- <SN_det+subst> = Quando há quebra da estrutura nominal *determinante + substantivo*. Ex: o/documentário, a/casa
 - <SN_nome próprio> = Quando há quebra da estrutura nominal *nome próprio*. Ex: João/Paulo
 - <SN_título+nome próprio> = Quando há quebra da estrutura nominal *título + nome próprio*. Ex: Sra./Maria de Lurdes
 - <SN_subst composto> = Quando há quebra da estrutura nominal *substantivo + substantivo*. Ex: creche/escola, cidade Duas/Barras
 - <SN_expressões/idiom/conv> = Quando há quebra de expressões idiomáticas, convencionais, etc. Ex: cartão/de crédito, boa/ação do dia
- 2) Para análise de sintagma preposicional o modelo padrão se configura da seguinte forma <Sintagma Prepositivo_estrutura sintática quebrada> e os modelos específicos:
- <SP_prep+subst> = Quando há quebra da estrutura preposicionada *preposição+substantivo*. Ex: de/Paris, com/amor
 - <SN_ambiguidade> = Quando há quebra de estrutura preposicionada de duplo sentido no lugar errado. Ex: Papa *abençoa/fiéis* do hospital. Dependendo do lugar onde a expressão for quebrada a frase pode assumir mais de um sentido. Pela quebra da linha acima entende-se que o Papa abençoou os fiéis que estavam no hospital, porém o certo seria que do hospital, o Papa abençoou os fiéis, desta forma a quebra do texto deveria ter sido feita da seguinte forma: Papa *abençoa fiéis*/do hospital.
- 3) Para análise de sintagma verbal o modelo padrão se configura da seguinte forma <Sintagma Verbal_estrutura sintática quebrada> e os modelos específicos:
- <SV_verbo+verbo> = Quando há quebra da estrutura verbal *verbo+verbo*. Ex: estava/brincando, vai tentar/falar
 - <SV_verbo+adv> = Quando há quebra da estrutura verbal *verbo+advérbio*. Ex: não pertencia/mais
 - <SV_colocações> = Quando há quebra de colocações verbais. Ex: *marcar um/encontro, manter-se em/forma, levar em/consideração*.
- 4) Para análise de oração coordenada o modelo padrão se configura da seguinte forma <Oração Coordenada_estrutura sintática quebrada> e os modelos específicos:
- <COORD_conj+oração> = Quando há quebra da oração coordenada após a conjunção. Ex: venha comigo *e/sinta* a emoção
 - <COORD_negativa> = Quando há quebra da oração coordenada negativa após a negação. Ex: Se teu pai te abandonou *não/significa* que ele é ruim
- 5) Para análise de oração subordinada o modelo padrão se configura da seguinte forma <Oração Subordinada_estrutura sintática quebrada> e os modelos específicos:
- <SUBORD_conj+oração> = Quando há quebra da oração subordinada após a conjunção. Ex: Eu senti *que/você* gostou
 - <SUBORD_se> = Quando há quebra da oração subordinada após a partícula apassivadora “se”. Ex: Eu te dou um brinquedo *se/você* passar de ano.
 - <SUBORD_comparativa> = Quando há quebra da oração subordinada após a conjunção adverbial comparativa. Ex: Eu o admiro *tanto quanto/desprezo* seu irmão, Chorava *como/uma* criança

Além das informações gramaticais, foi necessário codificar outras informações no corpus para fins de organização, como: Identificação da legenda - a etiqueta de identificação da legenda ajuda a encontrar uma legenda específica no corpus, pela sua

numeração, e tem como modelo padrão a etiqueta <sub n° da legenda> *legenda* </sub n° da legenda> e como exemplo específico <sub1> [Música suave] </sub1>. Esse tipo de etiqueta com dupla estrutura marca o intervalo a ser analisado, delimita aonde começa e termina o trecho lingüístico analisado. A palavra “sub” é a abreviação da palavra *subtitle* que significa “legenda”, em inglês. Optei pelo uso de “sub” pelo fato de ser um termo mais reconhecido globalmente que o português “legenda” e de ser reconhecido em diversos contextos, especialmente no âmbito da tradução audiovisual.

Linhas de legendas – a etiqueta de linhas de legendas informa se a legenda é de uma, duas ou mais linhas. O modelo padrão dessa etiqueta é <n° de linhas da legenda + a letra maiúscula “L” de linha>, se a legenda tiver 1 linha a etiqueta será <1L>, se a legenda tiver 2 linhas a etiqueta será <2L>.

Tempos inicial e final – a etiqueta de tempo mostra os tempos inicial e final de cada legenda. Isso ajuda, dentre outras coisas, a retomar a legenda em sua posição exata no vídeo. O modelo é <T>tempo x</T>. Um exemplo em uso pode ser <T>00:01:13,706 --> 00:01:21,112</T>.

Velocidade da legenda, variando entre baixa - até 13caracteres por segundo (cps), ideal - de 14 a 15cps, alta – a partir de 16cps. A etiqueta de velocidade de legenda ajuda a analisar se a densidade lexical das legendas é baixa, ideal ou alta para o tempo em que são exibidas. Para obter a duração de cada legenda e o total de caracteres é preciso consultar esses valores em programas de legendagem que trabalhem com duração e total de caracteres. O programa de legendagem utilizado nessa pesquisa, para obter esses dados, foi o Subtitle Workshop 2.51. A figura 1 ilustra a interface do Subtitle Workshop, em que o número 1 indica a duração da legenda e o número 2 indica o total de caracteres por linha e o total de caracteres da legenda.

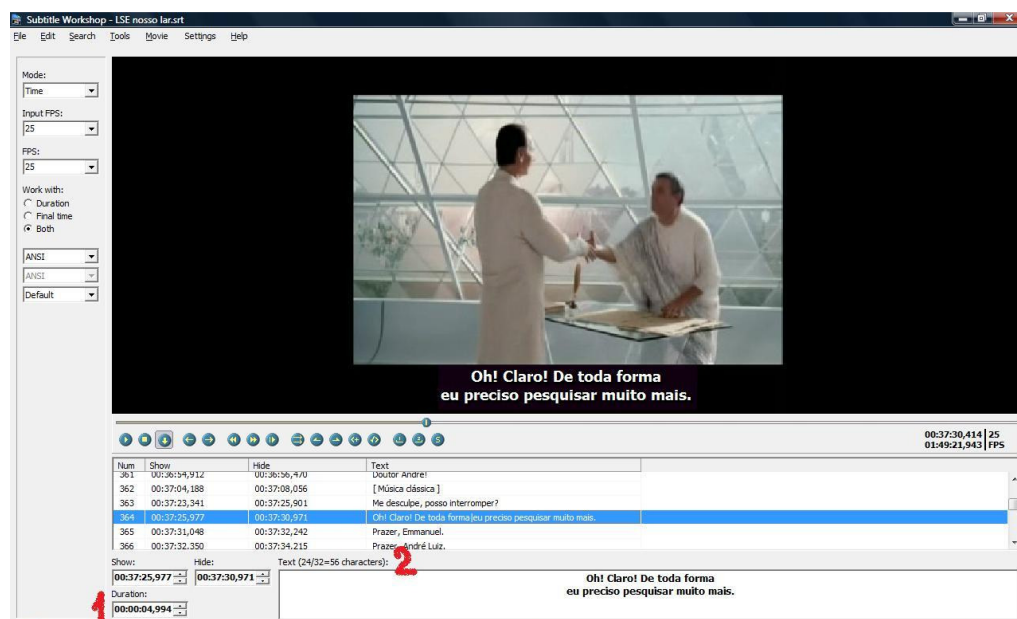


Figura 1: Interface do programa *Subtitle Workshop*

Caracteres por linha de legenda – a etiqueta de caracteres por linha de legenda ajuda a analisar, em legendas de duas linhas, se a má segmentação era decorrente da restrição de no máximo 35 caracteres por linha, ou se a má segmentação era inconsciente ou decorrente de uma decisão pessoal por parte do tradutor. O modelo padrão dessa etiqueta é <caracteres por linha n° de caracteres da linha> e um exemplo específico é <cpl 35>.

Problema de segmentação - a etiqueta de problema de segmentação codifica os casos onde há má segmentação dentro da mesma legenda e entre duas ou mais legendas, representada pelas respectivas etiquetas *problema de segmentação gramatical* <PROSEGG> e *problema de segmentação retórica* <PROSEGR>.

Em seguida, um trecho do corpus sem etiqueta e com etiqueta ilustra o processo de etiquetagem:

Trecho sem etiqueta	Trecho com etiqueta
3 00:02:46,766 --> 00:02:49,530 [Gritos]	<sub3><1L>3 <T>00:02:46,766 --> 00:02:49,530</T> [Gritos] <velocidade da legenda_baixa 10c/2,7s ></sub3>
4 00:02:53,673 --> 00:02:56,437 [Homem] Foi como acordar de um longo sonho.	<sub4><2L>4 <T>00:02:53,673 --> 00:02:56,437</T> <cpl9>[Homem] <cpl35>Foi como acordar de um longo sonho. <velocidade da legenda_alta 44c/2,7s ></sub4>
5 00:02:59,012 --> 00:03:02,607 E logo descobri que não pertencia mais ao mundo dos vivos.	<sub5><2L>5 <T>00:02:59,012 --> 00:03:02,607</T> <cpl33>E logo descobri que não <u>pertencia</u> <PROSEGG><SV_verbo+adv> <cpl24> <u>mais</u> ao mundo dos vivos. <velocidade da legenda_alta 57c/3,6s ></sub5>

Análise dos dados

A análise eletrônica dos dados foi viabilizada por duas das três ferramentas do programa *Wordsmith Tools 5.0*, o concordanciador *Concord* e a lista de palavras *WordList*. A *Concord* viabilizou a análise, através da visualização em concordância, dos padrões definidos nas etiquetas e a *WordList* viabilizou a análise da ocorrência de todas as palavras do corpus. A ferramenta *Concord* oferece uma opção de busca de palavras chamada *search word* que ajuda a contabilizar o total de palavras e/ou etiquetas pesquisadas no corpus. Por exemplo, se há interesse em saber a quantidade de erros de segmentação gramatical basta escrever no campo de busca a etiqueta <PROSEGG> e automaticamente o programa disponibilizará uma lista com todas as ocorrências. As figuras 2 e 3 ilustram as telas das ferramentas *WordList* e *Concord* respectivamente.

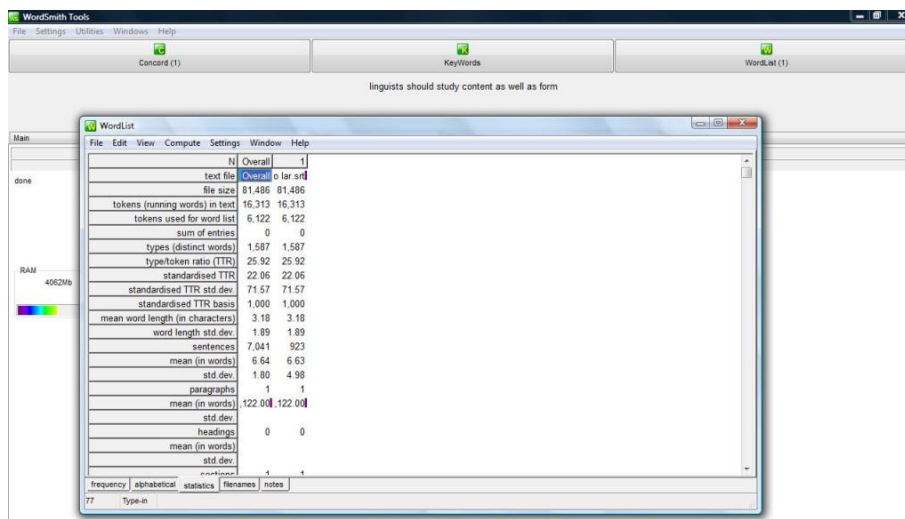


Figura 1: Interface da *WordList*

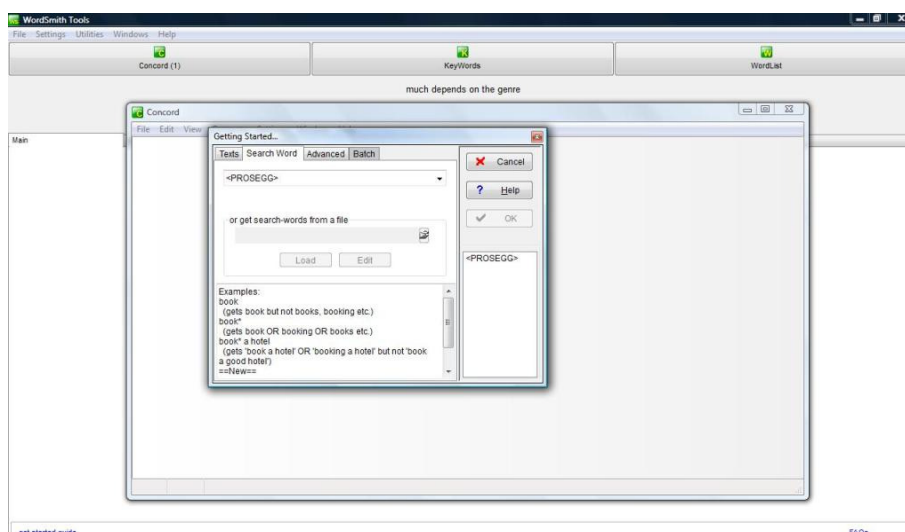


Figura 1: Interface da *Concord*

CONSIDERAÇÕES FINAIS

A narrativa do filme *Nosso Lar* narra a vida do espírito André Luís que ao despertar no Mundo Espiritual, se depara com criaturas assustadoras e sombrias. Além disso, ele se assusta por perceber que apesar de ter "morrido" ele ainda continua vivo e ainda sente fome, sede, frio e outras sensações materiais. Após um longo período de sofrimento ele é recolhido dessa zona de sofrimento e purgação de falhas do passado por espíritos do bem e é levado para a Colônia Espiritual *Nosso Lar*, de onde surge o nome do filme. A partir desse momento ele começa a conhecer melhor a vida no além-túmulo e a aprender lições e adquirir conhecimentos que mudarão completamente o seu modo de enxergar a vida.

O filme *Nosso Lar* contém 1.132 legendas, o que caracteriza um filme longo de muitos diálogos. Os resultados obtidos através da *WordList* mostraram que o total de palavras (tokens) da legendagem do filme *Nosso Lar* foi de 16.313 palavras e o total de palavras distintas (types) foi de 1.587. Os resultados obtidos através da *Concord* mostraram que houve um total de 125 problemas de segmentação num total de 1.132

legendas, em outras palavras podemos afirmar que 11% das legendas estão com problemas de segmentação. Alguns exemplos podem ilustrar esses resultados.

LEGENDA SEM ETIQUETA	LEGENDA COM ETIQUETA
PROBLEMAS DE QUEBRA DO SINTAGMA NOMINAL	
142 00:23:30,275 --> 00:23:32,743 Pelo menos tira aquela sensação de fome.	<sub142><2L>142 <T>00:23:30,275 --> 00:23:32,743</T> <cpl22>Pelo menos tira <u>aquela</u> <PROSEGG><SN_det+subst> <cpl17>sensação de fome. <velocidade da legenda_alta 39c/2,4s></sub142>
227 00:28:22,100 --> 00:28:24,534 Seu aparelho gastrointestinal estava	<sub227><2L>227 <T>00:28:22,100 --> 00:28:24,534</T> <cpl12>Seu <u>aparelho</u> <PROSEGG><SN_subst composto> <cpl23>gastrointestinal <u>estava</u> <PROSEGG><SV_verbo+verbo> <velocidade da legenda_ideal 35c/2,4s></sub227> 52c/5s></sub252>
PROBLEMAS DE QUEBRA DE SINTAGMA VERBAL	
143 00:23:32,811 --> 00:23:35,177 Com certeza, o deixou mais bem disposto.	<sub143><2L>143 <T>00:23:32,811 --> 00:23:35,177</T> <cps21>Com certeza, o <u>deixou</u> <PROSEGG><SV_verbo+adv> <cps18> <u>mais bem</u> disposto. <velocidade da legenda_alta 39c/2,3s></sub143>
197 00:26:54,646 --> 00:26:57,740 O amigo parece ter compreendido o sentido da água,	<sub197><2L>197 <T>00:26:54,646 --> 00:26:57,740</T> <cpl18>O amigo <u>parece</u> <u>ter</u> <PROSEGG><SV_verbo+verbo> <cpl31> <u>compreendido</u> o sentido da água, <velocidade da legenda_alta 49c/3,1s></sub197>
PROBLEMAS DE QUEBRA DE SINTAGMA PREPOSICIONADO	
215 00:27:42,193 --> 00:27:46,220 guarda um histórico de todas as ações praticadas no mundo material.	<sub215><2L>215 <T>00:27:42,193 --> 00:27:46,220</T> <cpl28>guarda um histórico <u>de</u> <u>todas</u> <PROSEGG><SP_prep+subst> <cpl38> <u>as ações</u> praticadas no mundo material. <velocidade da legenda_alta 66c/4s></sub215>

<p>PROBLEMAS DE QUEBRA DE ORAÇÃO COORDENADA</p>	
<p>585 00:52:33,483 --> 00:52:36,384 vamos entrar e você fala com o governador.</p>	<p><sub585><2L>585 <T>00:52:33,483 --> 00:52:36,384</T> <cpl19>vamos entrar e <u>você</u><PROSEGG><COORD_conj+oração> <cpl22>fala com o governador. <velocidade da legenda_ideal 41c/2,9s></p>
<p>PROBLEMAS DE QUEBRA DE ORAÇÃO SUBORDINADA</p>	
<p>196 00:26:49,441 --> 00:26:53,309 Todo o ceticismo termina quando se acorda no mundo espiritual.</p>	<p><sub196><2L>196 <T>00:26:49,441 --> 00:26:53,309</T> <cpl31>Todo o ceticismo termina <u>quando</u><PROSEGG><SUBORD_conj+oração> o> <cpl30>se acorda no mundo espiritual. <velocidade da legenda_alta 61c/3,8></sub196></p>
<p>201 00:27:05,190 --> 00:27:09,126 Lísias, eu não vou parar enquanto não souber o que está acontecendo.</p>	<p><sub201><2L>201 <T>00:27:05,190 --> 00:27:09,126</T> <cpl33>Lísias, eu não vou parar <u>enquanto</u><PROSEGG><SUBORD_conjunção+oração> <cpl34><u>não souber o que está acontecendo.</u> <velocidade da legenda_alta 67c/4s></sub201></p>

Pode-se perceber uma maior ocorrência de problemas de segmentação gramatical no âmbito dos sintagmas verbal e nominal e preposicionado. Isso pode ser atribuído ao fato de que os sintagmas são unidades de difícil análise, podendo ser confundidos entre si, dependendo do contexto. Já as orações coordenadas e subordinadas são identificadas com mais facilidade e são intuitivamente mais fáceis de serem distribuídas numa legenda. Para uma maior descrição da segmentação desse corpus seria necessária uma expansão dos itens de análise. Mas esse estudo teve como objetivo apenas a descrição dos casos de má segmentação das legendas em DVD do filme *Nosso Lar*.

Uma das conclusões que podemos chegar é que diante de maior quantidade de legendas má segmentadas nos níveis do sintagma deve-se atribuir mais atenção na análise das frases nos níveis do sintagma do que nos níveis da oração. Maior e melhor treinamento pode ser dado a tradutores em formação visando a questão da quebra dos sintagmas.

REFERÊNCIAS

- ARAÚJO, Vera Lúcia Santiago. Por um modelo de legendagem para Surdos no Brasil. **Tradução e Comunicação**, Revista Brasileira de Tradutores, 2008, n. 17, p. 59–76.
- BERBER SARDINHA, Tony. **Linguística de Corpus**. Barueri, São Paulo: Manole, 2004, p. 145. ISBN: 85-204-1676-4.
- DIAZ-CINTAS, Jorge; REMAEL, Aline. **Audiovisual Translation: Subtitling**. Manchester, UK & Kinderhook (NY), UK: St. Jerome Publishing, 2007. ISBN: 978-1900650-95-3/1- 900650-95-9.
- CHAVES, Élide Gama. **Legenda para Surdos no Brasil: uma análise baseada em corpus**. Monografia (Graduação em Letras): Universidade estadual do Ceará, Brasil, 2009.
- FEITOSA, Marcos Pinheiro. **Legendagem comercial e legendagem pirata: um estudo comparado**. Tese (Doutorado): Faculdade de Letras da Universidade Federal de Minas Gerais, Brasil, 2010.
- FRANCO, Eliana; ARAUJO, Vera L. S. Reading Television: Checking deaf people's Reactions to Closed Subtitling in Fortaleza, Brazil. In **The Translator**, v. 9, n. 2, 2003, p.249-267.
- HURTADO, Catalina Jiménez. Una gramática local del guión audiodescrito. Desde la semántica a la pragmática de un nuevo tipo de traducción. In **Catalina Jiménez Hurtado (ed.) Traducción y accesibilidad**. Peter Lang Internationaler Verlag der Wissenschaften, 2007, p. 55 – 80.
- KALANTZI, Dimitra. **Subtitling for the Deaf and Hard of Hearing: A corpus-based methodology for the analysis of subtitles with a focus on segmentation and deletion**. Tese (Doutorado): School of Languages, Linguistics and Cultures of The University of Manchester, UK, 2008.
- KARAMITROGLOU, Fotios. A Proposed Set of Subtitling Standards in Europe. In: **Translation Journal**, nº 2, v. 2, p. 1 – 15. 1998. Disponível em: <http://translationjournal.net/journal//04stndrd.htm> acessado em: 07/07/11.
- PEREGO, Elisa. Evidence of explicitation in subtitling: towards a characterization. In **Across Languages and Cultures**, volume 4, n 1, 63-88, 2003.
- PEREGO, Elisa. What Would We Read Best? Hypotheses and Suggestions for the Location of Line Breaks in Film Subtitles. In **The Sign Language Translator and Interpreter**. UK. St. Jerome Publishing, Manchester, 2008, p. 35-63. ISSN 1750-3981.
- PEREGO, Elisa. The codification of non-verbal information in subtitled texts. In CINTAS, J. D. (ed) **New trends in audiovisual translation**. Bristol,UK; Multilingual Matters, 2009, 58-69.
- REID, Hellene. Literature on the screen: subtitle translation for public broadcasting. In: Weterweel, Bart and Theo D'haen (Eds.) **Something understood. Studies in Anglo-Dutch literary translation**. Amsterdam: Rodopi, 1990, 97-107.
- SALWAY, Andrew. A corpus-based analysis of audio description. In **Jorge Diaz Cintas, Pilar Orero, Aline Remael (eds.) Media for All - Subtitling for the Deaf, Audio Description, and Sign Language**. Amsterdam - New York, NY: Rodopi, 2007, p. 151 – 174.