

A LINGUÍSTICA DE CORPUS COMO FERRAMENTA PARA TRADUÇÃO DE RESUMOS CIENTÍFICOS: ÁREA DE CINEMA

Evelyn Mary Ann Faeda¹

Jane Marian²

RESUMO

O objetivo da pesquisa foi analisar a linguística de corpus como uma abordagem metodológica para a tradução de resumos de artigos científicos, visto que é uma exigência da maioria das revistas científicas. Esta é uma pesquisa empírica de abordagem quantitativa e qualitativa. Para o estudo foram selecionados dezoito resumos de artigos científicos, sendo eles na língua portuguesa e inglesa (2.604 e 2.610 palavras, respectivamente) da área de cinema. Para as análises foi utilizado programa computacional *Sketch Engine*. Os resultados revelaram que o tradutor precisa de conhecimentos linguísticos e ferramentas que possam auxiliá-lo neste processo de busca de equivalentes tradutórios para obter o sucesso de uma tradução que possa ser compreendida para o leitor nativo.

Palavras-chave: Linguística de Corpus. Tradução. Resumos Científicos. Corpus Paralelo. Cinema.

¹ Aluna do 6º período do curso de Letras – Português e Inglês da FAE Centro Universitário. Voluntária do Programa de Apoio à Iniciação Científica (PAIC 2019-2020). *E-mail*: evelyn.faeda@mail.fae.edu.com

² Orientadora da Pesquisa. Doutora em Linguística. Professora da FAE Centro Universitário. *E-mail*: jane.marian@fae.edu

INTRODUÇÃO

A linguística de corpus há muito tempo é usada por pesquisadores e cientistas que estudam a linguagem. Segundo Frérot (2016, p. 37) o uso de corpora com propósito de utilizar na aplicação da tradução já é utilizada desde meados dos anos 90. No entanto, ainda não é conhecida por muitos estudantes, tradutores e professores desta área no Brasil.

Sabe-se que dominar a língua materna é um processo difícil, e quando se trata de uma segunda língua nem sempre o aprendiz ou profissional tem 100% de segurança. Neste caso, ele precisa de um banco de dados que possa ser confiável e consultado constantemente com rapidez. Assim, a linguística de corpus possibilita ao usuário aumentar seu leque de investigação em textos produzidos por falantes nativos, assim como, diminui as chances de erros ou equívocos na hora de traduzir um texto.

Na escrita científica, a maioria das revistas exigem a tradução dos resumos. Isso ocorre porque o inglês é a língua utilizada na ciência e permite que os pesquisadores de qualquer lugar do mundo se comuniquem e compartilhem o conhecimento. A informação tem avançado muito por causa das tecnologias e facilidades de interação entre os pesquisadores. Muitas revistas publicam a obra completa no idioma inglês. No entanto, como nosso idioma não é o inglês, mas sim o português, traduzir bem o resumo já é um sinal de cortesia e avanço para a divulgação do conhecimento. No resumo o leitor tem pistas sobre a essência do artigo de forma geral e breve. E, conseqüentemente, após a leitura ele decidirá se irá traduzir e ler o artigo completo ou não. Portanto, uma boa tradução é vital para que seu artigo seja visto, lido e citado internacionalmente.

O objetivo do presente projeto é compilar, processar e analisar um corpus composto por resumos bilíngues (português e respectivas traduções para o inglês) de artigos científicos da área de cinema. Serão utilizadas as ferramentas computacionais de análises linguísticas do programa *Sketch Engine* para extrair as palavras-chave da área e verificar como os termos são traduzidos e se o pesquisador fez uma tradução e revisão textual. Sabe-se que suportes de tradução, tais como ferramentas de tradução automática, por exemplo: o *Google Translator*, podem parecer rápidos e seguros, porém, para textos de áreas específicas as conseqüências de utilizar contextos incorretos podem ser altas, prejudicando assim a naturalidade da tradução.

1 FUNDAMENTAÇÃO TEÓRICA

1.1 A LINGUÍSTICA DE CORPUS

A linguística estuda a linguagem, seja ela verbal, gramatical ou até mesmo toda a evolução e história. Ela é subdividida em teórica e aplicada. A teórica visa estudar a linguagem de forma conceitual, é conhecida por ser uma pesquisa qualitativa. Em contrapartida, a aplicada tem como objetivo o estudo prático do idioma, neste caso, a pesquisa tem um viés mais quantitativo ou misto.

Já a linguística de corpus é uma das categorias que se fragmenta a partir da linguística aplicada. Ela se ocupa do estudo de textos, chamados de corpora. Esses são dados coletados especificamente para o propósito de uma pesquisa de uma língua ou sua variedade linguística. Essa ciência é dedicada à exploração da linguagem de forma empírica, Oliveira (2009, p. 49) acrescenta ainda que como a linguagem é um ato social ela deve ser analisada no contexto real da fala, ou seja, no discurso, lugar onde o significado é negociado.

Enquanto ainda se discutia o pensamento humano e como a linguagem ocorre na mente do indivíduo, Noam Chomsky, em 1957, publica seu livro *Syntactic Structures*, que trouxe uma grande discussão para o campo da linguística. O autor acreditava que a linguagem era inata ao ser humano, que dependendo das condições iniciais da criança, exposições e experiências, ela se tornava adulta com e possuía vários sistemas de linguagem, para ele as pessoas não aprendiam a língua, já nasciam com um sistema pronto de linguagem, apenas desenvolviam ou não de acordo com vivências (CHOMSKY; LIGHTFOOT, 2002). Porém, a língua não é tão simples assim, alguns pesquisadores se questionam sobre detalhes específicos da língua, observando os dilemas e as diferenças reais da linguagem.

O estudo da linguagem de forma quantitativa e empírica foi muito criticado por Chomsky nesta época. Mas, em 1959, Randolph Quirk gerenciou o primeiro centro de pesquisa da Europa que envolvia estudos empíricos da língua. Para isso, foi compilado o primeiro corpus, chamado Survey of English Usage. O projeto começou no final dos anos 50 e se iniciou com o objetivo de servir como referência para qualquer um que usasse ou estudasse a língua inglesa britânica. Nesta época, o corpus foi todo preparado à mão e organizado em cartões/fichas.

Com os avanços das tecnologias, em 1964, teve-se a notícia de que seria lançado o primeiro corpus eletrônico, o Corpus Brown, pioneiro em compilação de corpus eletrônico. Todos os dados que estavam no formato manual foram transferidos para o

computador por meio de cartões perfurados (FRANCIS, KUCERA, 1979, p 2). Atualmente, há diversos programas computacionais que auxiliam nesta coleta e no processamento dos dados, assim como, várias universidades têm pesquisas baseadas em corpora eletrônicos e centros de pesquisa em linguística aplicada (SARDINHA, 2000, p. 4).

A Linguística de Corpus surgiu para facilitar a investigação das estruturas linguísticas, seja evidenciando as diferentes funções que as palavras podem exercer dentro de um texto ou diferentes significados e interpretações que o leitor pode extrair. Ou seja, o conjunto de dados selecionado é compilado e sistematizado, dessa forma, o pesquisador poderá analisar os elementos lexicais do corpus. Primeiro extraíndo os itens do texto, geralmente, conforme os dados estatísticos apresentam, e na sequência análise qualitativa destes dados em seu contexto real da fala ou escrita.

Teubert e Cermáková (2004, p. 37) afirmam que dentro do conceito da linguística de corpus a linguagem é vista como um fenômeno social. Os autores sugerem igualmente que o significado consiste em um fenômeno, já que é algo que pode ser discutido com a comunidade do discurso. Para eles não há fórmula secreta nem na linguagem natural, nem em cálculos formais que contém o significado da palavra ou frase, por exemplo, a frase: “arma de destruição em massa” pode ter diferente significado quando duas pessoas diferentes falam a mesma frase. Portanto, a mensagem precisa ser analisada em seu contexto de uso.

1.2 AS POSSIBILIDADES PARA A TRADUÇÃO

Em pesquisas acadêmicas, as possibilidades de tradução para artigos científicos são grandes, visto que a ciência deve ser divulgada mundialmente. Tradutores precisam transmitir a informação da língua de partida para a língua de chegada sem modificar o sentido. Portanto, a tradução científica deve ser fiel e não modificada como nos textos literários que podemos adaptar à cultura em que o texto será traduzido. As discussões que compõem esse tema são, entre outros elementos, a fidelidade textual, a dicotomia entre fiel ou livre, a linha divisória entre tradução literária e tradução técnica e a tecnologia como aliada ou inimiga do tradutor, (OLIVEIRA, 2015, p 2). O tradutor tem papel relevante na comunidade científica, pois sua função é ler, interpretar e reproduzir o conhecimento sem perder a essência do conteúdo e informação.

De um falante ingênuo à um tradutor ingênuo, pode haver perdas de convenções importantíssimas nos trabalhos de traduções. Perdas em níveis morfológicos, semânticos, pragmáticos, discursivos e culturais são algumas delas que podem prejudicar o trabalho de um tradutor. A presença de um corpus de estudo na pesquisa é uma ferramenta

de linguagem extremamente eficaz. A compilação de corpus dá ao tradutor inúmeros exemplos de aceitabilidade e convencionalidade que podem ser empregados nos seus textos traduzidos (EVER; FINATTO, 2015, p 22).

As línguas naturais não podem ser simplesmente traduzidas uma para outra, palavra por palavra. É necessário checar todas as colocações e colocados possíveis do texto e verificar qual o contexto e sentido elas assumem no texto original para que ao traduzir não sejam perdidas suas funções na língua alvo. O inglês é uma língua em que as palavras compostas ou *clusters* podem, facilmente, mudar o significado. Sendo assim, o tradutor deve prestar atenção na construção da linguagem do texto de chegada, garantindo assim, a naturalidade do texto, neste sentido, a compilação e pesquisa em um corpus são instrumentos válidos e necessários.

Para o tradutor, outro investimento é a construção do corpus paralelo. Um corpus paralelo é resultado de textos alinhados com suas respectivas traduções, para posterior processamento em *softwares* computacionais. Com a comparação desses textos o pesquisador pode monitorar sua prática de tradução (TEUBERT, CERMÁKOVÁ, 2004, p 119). Considerando que os dicionários bilíngues tradicionais ou traduções automáticas, como o *Google Translator*, são ferramentas que o tradutor deve utilizar, mas não se limitar a linguística de corpus é uma área que possui ferramentas essenciais para o tradutor do século XXI.

1.3 ANÁLISE DO DISCURSO

A análise do Discurso é uma área que deriva da Linguística em que estuda a estrutura de um texto ou fala, pois cada texto tem seu contexto atrelado, no qual ele vai ser desenvolvido. Há três concepções de linguagem no transcorrer de sua história: uma representação do mundo e do pensamento, um instrumento de comunicação ou forma de ação e interação (SOUZA; BARBA, 2014, p. 4). A segunda concepção será a mais usada neste trabalho.

Ela seria uma linguagem centrada na comunicação, transmitindo mensagens, com um emissor e receptor. O emissor dessa mensagem a compartilha com um contexto único, e para chegar no seu receptor é necessário transmitir sua fala de forma coesa e coerente. A análise do discurso não é uma metodologia é uma disciplina de interpretação (CAREGNATO; MUTTI; 2005, p. 4), pois tem-se a pretensão de analisar os sentidos do texto.

Portanto, a linguagem para a análise do discurso, vai além do texto, e encontra com seu contexto. Por isso que para se analisar um texto, ou fala, para uma tradução é necessário trazer seu contexto e sentido, pois são eles que podem moldar totalmente sua mensagem que será passada para o receptor.

2 METODOLOGIA

O presente projeto de pesquisa é um estudo empírico de abordagem quantitativa e qualitativa. Para a linguística de corpus a linguagem é um sistema probabilístico, quanto maior a quantidade de palavras no corpus maior a probabilidade de determinado traço ocorrer “Assim, algumas palavras têm frequência de ocorrência muito rara e, para que haja probabilidade de ocorrerem no corpus, é necessário incorporar uma quantidade grande de palavras” (SARDINHA, 2000, p. 23).

Dessa forma, foram pesquisados e compilados 18 resumos de artigos científicos em língua portuguesa e suas respectivas traduções para a língua inglesa na área de Cinema. Ou seja, foi necessário extrair dos artigos científicos somente os resumos científicos e separar os textos escritos em português dos escritos em inglês para posterior alinhamento. O corpus de português contendo 2.604 palavras e o de inglês com 2.610 palavras. A base de dados utilizada para a seleção dos artigos científicos foram o Scielo e o buscador do Google Acadêmico.

Na sequência, por ser um corpus cujo objetivo era analisar seus equivalentes tradutórios foi preciso alinhar os resumos em português e suas traduções para o inglês. Um texto alinhado significa dizer que cada frase da língua portuguesa ficará linear com seu correspondente tradutório para futura manipulação dos dados. Para o alinhamento foi utilizado o programa *YouAlign*, neste *software* os textos são salvos em formato TMX.

E, finalmente, no processamento e análise dos dados foi utilizado o programa *Sketch Engine*, um sistema computacional de análises linguísticas. Neste, foram utilizadas cinco ferramentas principais: *parallel concordance*, *concordancer*, *wordlist* e *keywords*. A primeira, *parallel concordance*, mostra o corpus paralelo (frase em português e respectivos correspondentes tradutórios). A segunda, *concordancer*, apresenta linhas de concordância (a palavra nódulo fica centralizada e os contextos ao redor do nódulo). A terceira, o *wordlist*, exibe uma lista de palavras por ordem de frequência. A quarta, *keywords*, extrai palavras-chave do corpus de estudo comparando estatisticamente com um corpus de referência. E, por último, o *word sketch*, que ajudou no reconhecimento das combinações das palavras, sendo assim possível analisar quais palavras gramaticais e colocações acompanham cada termo.

3 ANÁLISE DOS RESULTADOS

O objetivo desta pesquisa foi verificar como a construção de um corpus paralelo, utilizando a perspectiva da linguística de corpus, poderia ser um diferencial para a tradução-humana. Para isso, foi compilado e alinhado um corpus da área de cinema e, na sequência, esse corpus foi processado pelo programa *Sketch Engine* e analisado pelas autoras.

Pôde-se perceber que algumas palavras são repetidas frequentemente em ambos os textos, sejam elas em português ou na língua inglesa. Após o processamento do corpus pela ferramenta *wordlist* e *keywords* selecionamos uma palavra para compor a amostra. O item lexical cinema foi encontrado 33 vezes no corpus em português e foi utilizado como exemplo nesta pesquisa, pois se trata de uma palavra-chave frequentemente traduzida pelo tradutor neste contexto.

Conhecido como a Sétima Arte, o cinema é de grande valor histórico. Vê-se que o Cinema é esse processo de realização de filmes e também os filmes em si. Então, foi necessário analisar se todos os textos mantinham uma tradução que compunha todo o contexto apresentado no artigo da língua original.

A partir do *Parallel Concordance* pôde-se analisar as traduções alinhadas. Assim, investigou-se as diferentes formas em que o texto poderia ser traduzido observando o seu contexto. Cinema apresentou três traduções possíveis, sendo elas: *film* (8), *cinema* (23), *cinematograph* (2).

FIGURA 1 – Corpus Paralelo Português/Inglês - Cinema

<> Este artigo pretende discutir as relações entre artes visuais, cinema e os novos desafios impostos às diversas produções artísticas contemporâneas, examinando-os à luz de uma determinada tradição artística que, a seu tempo, buscou constituir uma imagística que não dissociasse forma artística de matéria histórica. </>	<> this article intend to discuss the relation between visual art, cinema , and the new challenge impose on contemporary artistic production, examine them under the light of an artistic tradition that aim at create an imagery that would not dissociate artistic form from historic matter. </>
<> Este artigo apresenta o projeto de extensão cinema , Subjetividade e Sociedade: A Sétima Arte na Produção de Saberes, da Universidade Federal de São Paulo (UNIFESP) – campus Baixada Santista, que utiliza a produção cinematográfica para a reflexão, questionamento, produção e aprofundamento de questões ligadas à constituição das configurações subjetivas na contemporaneidade. </>	<> this article present the extension project Cinema , Subjectivity and Society: the seventh art in the production of knowledge of the Federal University of São Paulo – campus Santos, which use film production for reflection, question, production and explo- ration of issue relate to constitution of subjective configuration in contemporary time. </>
<> Em seu loft na Segunda Avenida – Babylonests – e em sua moradia seguinte – Hendryxsts – na Christopher Street, conceberam e realizaram projetos de cinema e fotografia. </>	<> in his loft on Second Avenue – Babylonests – and at his next home – Hendryxst – on Christopher Street, they conceive and carry out film and photography project. </>
<> Esse trabalho abrange uma visão crítica sobre cinema como peça artística ou mercadológica estabelecendo uma relação com as características estruturais das produtoras. </>	<> this work include a critical view of cinema as artistic or marketing piece establish a relationship with the structural characteristic of producer. </>
<> A pesquisa objetiva entender as formas como as empresas de cinema do Rio Grande do Sul se estruturam procurando uma correspondência com o tipo de cinema produzido. </>	<> the research aim to understand the way in which film company in Rio Grande do Sul be structure for a match with the type of film produce. </>
<> A pesquisa objetiva entender as formas como as empresas de cinema do Rio Grande do Sul se estruturam procurando uma correspondência com o tipo de cinema produzido. </>	<> the research aim to understand the way in which film company in Rio Grande do Sul be structure for a match with the type of film produce. </>

FONTE: As Autoras (2020)

A análise dos corpora paralelos permitem ao tradutor compreender os possíveis equivalentes tradutórios por diferentes ângulos e analisar o contexto em que a palavra se insere. O termo Cinema deriva da palavra Cinematógrafo, a invenção para contar histórias pelos irmãos Lumière. O Cinema nada mais é do que um conjunto de imagens, que quando juntas se transformam em uma narrativa, seja ela real ou fictícia dando divertimento ao público (MASCARELLO, 2015). O fato de a palavra *cinematograph* ter sido traduzida apenas duas vezes nos chamou a atenção e fizemos a busca da palavra no corpus em inglês (FIG. 2).

FIGURA 2 – Corpus Paralelo Inglês/Português - *Cinematograph*

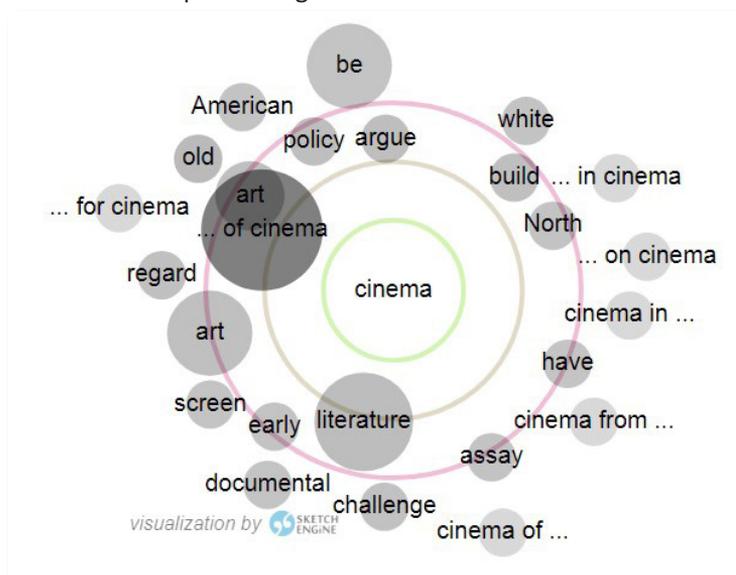
<p><s> This thesis compiles and analyses the chronicle regarding the kinetoscope/cinematograph published in carioca's newspapers and magazines from 1894 to 1922 - from the beginning of these media up to the establishment of the cinema as both a solid industry and as an art form. </s></p>	<p><s> Esta tese compila e analisa a prosa cronística a respeito do kinetoscópio/cinematógrafo, publicada em periódicos cariocas de 1894 a 1922 - dos primórdios dos media até o estabelecimento do cinema enquanto indústria sólida e arte. </s></p>
<p><s> The goal of the research is to establish how the kinetoscope and the cinematograph gave form to the way writers observed the city in general, and specially the streets of Rio de Janeiro in those years when the unbreathable alleys of the provincial city turned into long and elegant avenues built up as open stages - inviting the "Society" to give up private habits to public ones. </s></p>	<p><s> O eixo da pesquisa visa a estabelecer como o kinetoscópio e o cinematógrafo moldaram o olhar dos cronistas à cidade e, em especial, às ruas do Rio de Janeiro, naqueles anos em que as vielas acanhadas da cidade provinciana transformavam-se em vias elegantes abertas como palcos, a convidarem a "sociedade" a trocar o âmbito privado pela flânerie pública. </s></p>

FONTE: As Autoras (2020)

Observou-se que quando a palavra em português se referenciava ao cinematógrafo os tradutores utilizaram o termo equivalente *cinematograph*. No entanto, essa palavra não é muito utilizada na atualidade, o termo era utilizado para descrever o profissional que operava a máquina que transmitia ao público curtas em tela. Essa é a origem do que nós conhecemos hoje como Cinema (COSTA, 1998), dessa forma, o termo deixou de ser usado, e a sua tradução *cinematograph* passou a ser usada como equivalente de Cinema. A compilação do corpus e processamento dos dados trazem informações ao tradutor que ajudam na compreensão dos equivalentes tradutórios.

Então, a partir disso entende-se que Cinema tem a tarefa de contar histórias, retirando momentos reais da vida ou criando novos cenários, dessa forma podendo tratar-se dos processos da criação da linguagem do Cinema. Com a ferramenta *Word Sketch* foi possível analisar quais são as palavras que acompanham as traduções equivalentes, dessa forma, tendo noção do que se tratam e como elas aparecem no texto.

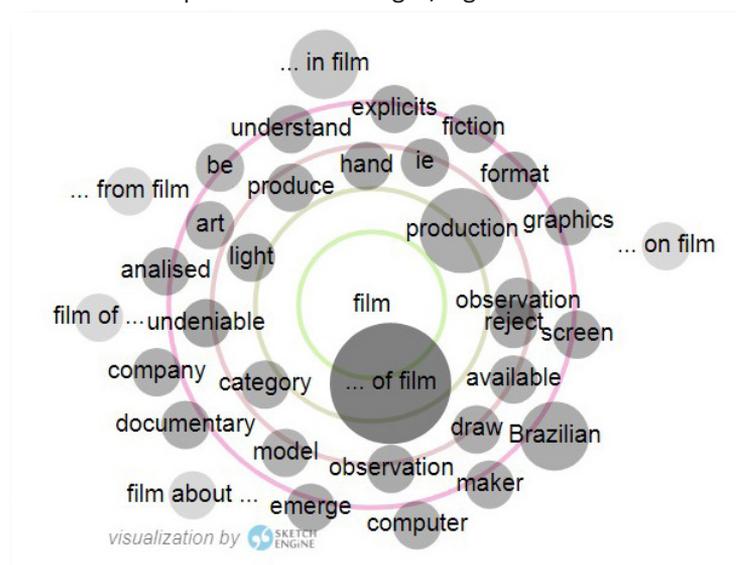
FIGURA 3 – Corpus em Inglês – *Cinema*



FONTE: As Autoras (2020)

A palavra *Cinema* apresenta várias combinações que podem trazer sentido para a frase, e apresentando modificadores e proposições que agregam para seu sentido, se tratando de produções cinematográficas, a arte dessa linguagem de contar histórias. (FIG. 3). Para a palavra *Film* percebeu-se que seus modificadores são substantivos, por se tratar do produto final de uma produção, sendo possível perceber os formatos de filmes que são na maioria das vezes abordados, como Cinema Documental, Ficção e entre outros. (FIG. 4).

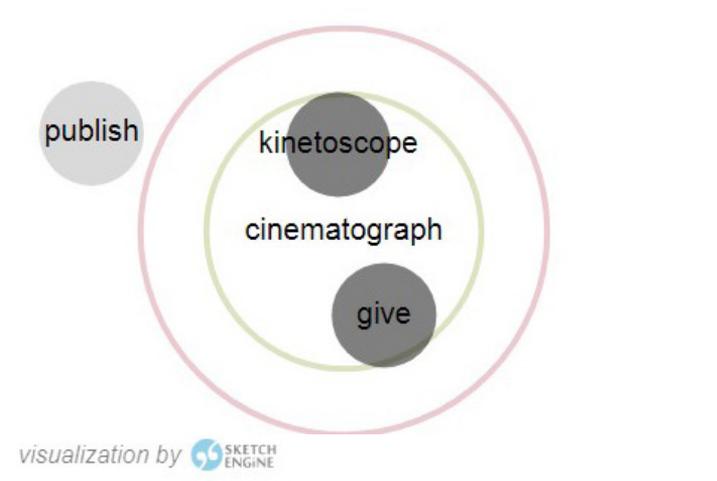
FIGURA 4 – Corpus Paralelo Português/Inglês - *Film*



FONTE: As Autoras (2020)

Com *cinematograph* os resultados não foram diferentes, já que ela se refere a invenção feita pelos irmãos Lumieré, o cinematógrafo, derivada do Cinetoscópio, criação de Thomas Edison, que permitia apenas que a pessoa assistisse imagens em movimento (BRITANNICA, 2020). Sendo assim, a palavra seguiu um caminho de desconstrução com os acontecimentos do passado (FIG. 5). Pois agora, dá significado ao Cinema como conhecemos nos dias atuais.

FIGURA 5 – Corpus Paralelo Português/Inglês - *Cinematograph*



FONTE: As Autoras (2020)

Com essa ferramenta de busca do *Sketch Engine* conseguimos analisar os dados por diferentes ângulos o que auxilia o tradutor na compreensão da linguagem, probabilidade de ocorrências e frequências em que aparecem nos textos, assim como, possibilidades de equivalentes tradutórios (MORA; SIERRA, 2013). Investigar o sentido das palavras dentro das frases ajuda a desenvolver um conhecimento sobre suas colocações e os contextos empregados. Como a palavra *Film*, que além de ser uma equivalência para o produto final de uma produção cinematográfica, o filme, também é visto em frases sobre companhias de cinema (FIG. 6).

FIGURA 6 – Corpus Paralelo Português/Inglês - *Film*

<p>① doc#0</p> <p><s> The research aims to understand the ways in which film companies in Rio Grande do Sul are structured for a match with the type of film produced. </s></p>	<p><s> A pesquisa objetiva entender as formas como as empresas de cinema do Rio Grande do Sul se estruturaram procurando uma correspondência com o tipo de cinema produzido. </s></p>
<p>① doc#0</p> <p><s> The research aims to understand the ways in which film companies in Rio Grande do Sul are structured for a match with the type of film produced. </s></p>	<p><s> A pesquisa objetiva entender as formas como as empresas de cinema do Rio Grande do Sul se estruturaram procurando uma correspondência com o tipo de cinema produzido. </s></p>

FONTE: As Autoras (2020)

Portanto, é preciso identificar em qual meio a palavra pode ser utilizada, analisando as particularidades da língua. Visto que, a palavra Cinema, em língua

portuguesa, pode corresponder, resumidamente, como a Sétima Arte e a produção a sua volta, como fotografar e filmar eletronicamente, e os filmes propriamente ditos. Vemos essas características aplicadas em três vocábulos diferentes na língua inglesa. Sendo respectivamente, *Cinema*, *Film* e *Cinematograph*. Ou seja, é necessário compreender o contexto da frase para poder aplicar um dos três vocábulos equivalentes. Na língua portuguesa se tem o costume de usar Cinema como denominador comum para corresponder suas propriedades, sendo assim uma variedade linguística comparada à língua inglesa, que diferencia suas características (FIG. 7).

FIGURA 7 – Corpus Paralelo Português/inglês - Cinema

① doc#0	<s> A relação entre literatura e cinema é antiga e ao comparar questões da criação literária com os dilemas dos realizadores cinematográficos, encontraremos muitos pontos em comum. </s>	<s> The relationship between literature and cinema is old and when comparing issues of literary creation with the intricacies of film makers, we find many points in common. </s>
① doc#0	<s> Mas a influência da literatura sobre o cinema é inegável e pode ser facilmente comprovada no terreno das adaptações. </s>	<s> But the influence of literature on film is undeniable and can be easily established on the ground of the adjustments. </s>
① doc#0	<s> Pretendemos, a partir daqui, mostrar que essa relação é muito mais rica e até acontece no sentido inverso, isto é, o cinema também pode influenciar a literatura. </s>	<s> We intend, from here, show that this relationship is much richer and even true in reverse, ie the film can also influence the literature. </s>
① doc#0	<s> Optando pela modalidade narrativa, o cinema retira da literatura parte significativa da tarefa de contar histórias. </s>	<s> Opting for narrative format, the film draws from the literature significant part of the task of storytelling. </s>
① doc#0	<s> A narratividade continua a ser o traço hegemônico da cinematografia, apesar da grande diferença entre a página de um livro e a tela branca do cinema . </s>	<s> The narrative remains the hegemonic dash of cinematography , despite the great difference between the book and page of a white screen cinema . </s>
① doc#0	<s> Ambas as artes, Cinema e Literatura, apesar de se constituírem sistemas semióticos distintos, possuem o nobre ofício de alimentar e trazer para as mídias o prestígio da grande arte ou, no dizer de alguns, tornar a arte erudita acessível ao grande público. </s>	<s> Both arts, cinema and literature, although they constitute distinct semiotic systems have the noble craft of food and bring the media to the prestige of high art or, in the words of some, making high art accessible to the general public. </s>

FONTE: As Autoras (2020)

Em alguns casos, foi analisado o conteúdo das frases pelo *Google Translator*, para que pudesse ser certificado que essa ferramenta, que também consiste na base da Linguística de Corpus é confiável. Pode-se perceber que muitas frases tinham seu conteúdo de certa forma perdido se colocado para traduzir pela ferramenta, visto que palavras e frases eram traduzidas separadamente, não era levada em conta a análise da frase em geral, mas palavra por palavra. Sendo assim, em algumas ocorrências perdesse a essência do texto, e não leva em consideração o uso da linguagem natural.

CONSIDERAÇÕES FINAIS

O objetivo inicial do projeto era analisar a linguística de corpus como uma abordagem metodológica para a tradução de resumos de artigos científicos. Sabe-se que é uma exigência da maioria das revistas que o resumo seja traduzido para o inglês, língua utilizada pela ciência para a divulgação e disseminação do conhecimento.

Notou-se que as traduções da palavra cinema (em português) para *cinema*, *film* e *cinematograph* (em inglês) dependem de características relativas as divisões da área específica de cinema. O corpus pode revelar informações que, normalmente, não são percebidas pelo pesquisador sem as informações estatísticas do programa e agrupamento de palavras para verificação e análise das ocorrências.

O conhecimento deste instrumento por novos pesquisadores e estudiosos é uma porta de entrada para dissertações e investigações deste cunho. Com o uso de um corpus paralelo pode-se explorar o modo que uma ideia foi expressada em uma língua natural, pode ser entendida da mesma maneira em outra língua (XIAO, WEL, 2014, p. 2).

REFERÊNCIAS

- BRITANNICA ESCOLA. **Os irmãos Lumieré**. Disponível em: <<https://escola.britannica.com.br/artigo/irm%C3%A3os-Lumi%C3%A8re/483343>>. Acesso em: 24 jun. 2020.
- CANDEL-MORA, Miguel A.; VARGAS-SIERRA, Chelo. An analysis of research production in corpus linguistics applied to translation. **Procedia-Social and Behavioral Sciences**, Valencia, v. 95, n. 1, p. 317-324, nov. 2013. Disponível em: <<https://www.sciencedirect.com>>. Acesso em: 24 jun. 2020.
- CAREGNATO, Rita Catalina Aquino; MUTTI, Regina. Pesquisa qualitativa: análise de discurso versus análise de conteúdo. **Texto & Contexto-Enfermagem**, Porto Alegre, v. 15, n. 4, p. 679-684, out. 2006. Disponível em: <<https://www.scielo.br/pdf/tce/v15n4/v15n4a17>>. Acesso em: 24 jun. 2020.
- CHOMSKY, Noam. **Sobre natureza e linguagem**. São Paulo: Martins Fontes, 2006.
- COSTA, Renato da Gama-Rosa. Os cinematógrafos do Rio de Janeiro (1896-1925). **História, Ciências, Saúde-Manguinhos**, Rio de Janeiro, v. 5, n. 1, p. 153-168, jun. 1998. Disponível em <https://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-59701998000100010&lng=pt&tlng=pt>. Acesso em: 24 jun. 2020.
- EVERS, Aline; FINATTO, Maria José Bocorny. Linguística de corpus e tradução. **GTLex**, v. 1, n. 2, p. 271-295, 2015.
- FRANCIS, Kucera. Brown Manuals. **Brown University**, 1979. Disponível em: <<http://korpus.uib.no/icame/manuals/BROWN/INDEX.HTM>> Acesso em: 24 jun. 2020.
- FRÉROT, Cécile. Corpora and corpus technology for translation purposes in professional and academic environments. Major achievements and new perspectives. **Cadernos de Tradução**, Florianópolis, v. 36, n. Spe, p. 36-61, fev. 2016. Disponível em: <<https://www.scielo.br/pdf/ct/v36nspe/2175-7968-ct-36-spe1-0036.pdf>> Acesso em: 24 jun. 2020.
- OLIVEIRA, Lúcia Pacheco de. Linguística de Corpus: teoria, interfaces e aplicações. **Matraga**, Rio de Janeiro v. 16, n. 24, p. 01-29, jun. 2009. Disponível em: <<https://www.e-publicacoes.uerj.br/index.php/matraga/article/view/27796>> Acesso em: 24 jun.2020
- OLIVEIRA, Joacyr Tupinambás de. **A Linguística de Corpus na formação do tradutor: compilação e proposta de análise de um corpus paralelo de aprendizes de tradução**. 2015. 147 f. Tese (Doutorado em Letras) – Universidade de São Paulo, 2015. Disponível em: <https://www.teses.usp.br/teses/disponiveis/8/8147/tde-26052015-104749/publico/2015_JoacyrTupinambasDeOliveira_VCorr.pdf>. Acesso em: 24 jun. 2020.
- MASCARELLO, Fernando. **História do cinema mundial**. Campinas: Papirus, 2015. Disponível em: <<https://books.google.com.br/books?hl=pt-BR&lr=&id=GniADwAAQBAJ&oi=fnd&pg=PT2&dq=hist%C3%B3ria+do+cinema&ots=RRGcT3jl8J&sig=-15jmtBvBNFFOY8HlyVFLeoX1oc#v=onepage&q=hist%C3%B3ria%20do%20cinema&f=false>>. Acesso em: 24 jun. 2020
- SARDINHA, Tony Berber. Linguística de corpus: histórico e problemática. **Delta**, São Paulo v. 16, n. 2, p. 323-367, fev. 2000. Disponível em: <<https://www.scielo.br/pdf/delta/v16n2/a05v16n2.pdf>>. Acesso em: 24 jun. 2020.

SOUZA, Marinho; BARBA, Clarides. Análise do Discurso: O que é? Como se faz? E para quê serve?. **Gestão Universitária**, São Paulo, n. 51, mar. 2014. Disponível em: <<http://www.gestaouniversitaria.com.br/artigos/analise-do-discurso-o-que-e-como-se-faz-e-para-que-serve>>. Acesso em: 24 jun. 2020.

TEUBERT, Wolfgang; CERMÁKOVÁ, Anna. Corpus linguistics: a short introduction. **Continuum**, Nova York, v. 0, n. 1, p. 01-176, abr. 2007.

XIAO, Richard; WEI, Naixing. Translation and contrastive linguistic studies at the interface of English and Chinese: Significance and implications. **Corpus Linguistics and Linguistic Theory**, Berlim, v. 10, n. 1, p. 1-10, maio 2014. Disponível em: <https://eprints.lancs.ac.uk/id/eprint/64329/1/01_Xiao_and_Wei.pdf>. Acesso em: 24 jun. 2020.