

UNIVERSIDADE FEDERAL DE UBERLÂNDIA  
INSTITUTO DE LETRAS E LINGUÍSTICA

GUILHERME RODRIGUES FERREIRA

**PROTÓTIPO DE UM VOCABULÁRIO BILÍNGUE INGLÊS/PORTUGUÊS DE *STAR WARS* BASEADO EM *CORPUS***

UBERLÂNDIA

2022

GUILHERME RODRIGUES FERREIRA

**PROTÓTIPO DE UM VOCABULÁRIO BILÍNGUE INGLÊS/PORTUGUÊS DE *STAR WARS* BASEADO EM *CORPUS***

Trabalho de Conclusão de Curso como parte dos requisitos para a obtenção grau de licenciado em Letras – Inglês e Literaturas de Língua Inglesa - Instituto de Letras e Linguística (ILEEL).

Orientador: Prof. Dr. Guilherme Fromm

UBERLÂNDIA

2023

## AGRADECIMENTOS

Agradeço ao professor Guilherme Fromm, que despertou meu interesse por Linguística de *Corpus* durante o curso e gentilmente me acolheu como orientando.

À Emanuelle, que aceitou viver ao meu lado e sempre demonstrou verdadeiro e infinito entusiasmo com todas as etapas deste trabalho.

Aos meus pais, que sempre apoiaram a minha jornada universitária e as possibilidades do curso de Letras.

Aos meus amigos, que sempre mencionam o quanto adoram a ideia de um trabalho tão longo dedicado à minha paixão de infância, *Star Wars*.

E às minhas colegas de curso, Ana Cecília e Luyza, que diariamente me acompanharam nos cafés.

## RESUMO

O contínuo lançamento de novos títulos em determinadas franquias audiovisuais movimentou uma série de profissionais no ramo da tradução. Principalmente no caso de *Star Wars*, diversas obras são lançadas todos os anos, e com a solidificação da franquia na cultura popular, seus elementos são constantemente referenciados em outras mídias que também serão traduzidas. Sendo assim, este trabalho propõe a produção e disponibilização de um vocabulário bilíngue inglês/português de *Star Wars* no VoTec (Vocabulário Técnico Online), um ambiente web de gestão terminológica para tradutores iniciantes e profissionais. Para isso, foi levantado um *corpus* referente ao universo canônico de *Star Wars*, incluindo filmes, livros, séries e animações, com análise no *Wordsmith Tools 7*. Por fim, o trabalho avalia a complexidade da operação, as possibilidades de avanço futuro e a utilidade do vocabulário.

**Palavras-chave:** Terminologia; Terminografia; Linguística de *Corpus*; VoTec; *Star Wars*.

## ABSTRACT

The constant release of new installments of specific audiovisual franchises engages numerous professionals in the translation industry. Especially in the case of Star Wars, several entries are released every year, and with the consolidation of the franchise in popular culture, its elements are constantly referenced in other media that will also be translated. Therefore, this paper proposes the insertion of a bilingual English/Portuguese Star Wars vocabulary in VoTec (*Vocabulario Técnico Online*), a web environment for terminology management for professional translators and translation students. To implement this insertion, a corpus of the canonical universe of Star Wars, including films, books, series, and animated series, was created using Wordsmith Tools 7. Lastly, the study examines the complexity of the operation, the possibilities for future advancement and the usefulness of the vocabulary.

**Keywords:** Terminology; Terminography; Corpus Linguistics; VoTec; Star Wars.

## LISTA DE FIGURAS

Figura 1 - VHS de Ameaça Fantasma com novo título “ <i>Star Wars</i> ”. Ao lado, cartaz brasileiro anunciando “O Império Contra-Ataca” como "Guerra nas Estrelas".....	13
Figura 2 - Interface do <i>WordsmithTools</i> 7.....	18
Figura 3 - Interface de administrador do VoTec. ....	19
Figura 4 - Erro de codificação no corpus. ....	24
Figura 5 - Interface do SubtitleEdit durante o processo de limpeza de legendas.....	25
Figura 6 - Diferença do material antes e após a limpeza.....	26
Figura 7 - Processo de limpeza manual dos créditos de tradutores para análise no corpus. .....	27
Figura 8 - KeyWord List.....	28
Figura 9 - Contextos encontrados para "Ewok". ....	29
Figura 10 - Contextos cadastrados no VoTec. ....	30
Figura 11 - Interface do VoTec. ....	31
Figura 12- Processo de cadastro de um novo termo.....	32
Figura 13 - Cadastrando contextos. ....	33
Figura 14 - Cadastrando uma fonte de contexto.....	34
Figura 15 - Detalhando o termo.....	34
Figura 16 - Criando conceito final e definição. ....	35
Figura 17 - Interface do VoTec para o usuário final. ....	35

## LISTA DE QUADROS

Quadro 1 – <i>Corpus</i> de livros (Títulos originais em inglês)..	21
Quadro 2 – <i>Corpus</i> de animações (Títulos originais em inglês).	22
Quadro 3 – <i>Corpus</i> de filmes (Títulos originais em inglês).	23
Quadro 4 – <i>Corpus</i> de séries (Títulos originais em inglês).	23
Quadro 5 – Termos pré-selecionados.	28

## SUMÁRIO

<b>1 INTRODUÇÃO .....</b>	<b>9</b>
1.1 Objetivos .....	9
<b>2 SOBRE <i>STAR WARS</i> .....</b>	<b>11</b>
<b>3 REFERENCIAIS TEÓRICOS .....</b>	<b>14</b>
3.1 Terminologia .....	14
3.2 Teoria Comunicativa da Terminologia .....	15
3.3 Etnoterminologia .....	16
3.4 Terminografia .....	16
3.5 <i>Wordsmith Tools</i> .....	17
3.6 VoTec .....	18
<b>4 METODOLOGIA.....</b>	<b>20</b>
4.1 Tratamento do <i>Corpus</i> .....	24
4.2 Seleção de candidatos a termos .....	27
4.3 Exploração das linhas de concordância .....	29
<b>5 CRIAÇÃO DOS VERBETES.....</b>	<b>31</b>
5.1 Criação do projeto no VoTec.....	31
5.2 Elaboração dos verbetes.....	31
<b>6 CONSIDERAÇÕES FINAIS.....</b>	<b>37</b>
<b>REFERÊNCIAS .....</b>	<b>38</b>



## 1 INTRODUÇÃO

*Star Wars* é uma das franquias mais importantes e influentes da história do cinema. Desde o lançamento do primeiro filme em 1977, a saga criada por George Lucas se tornou um fenômeno e uma valiosa marca, hoje propriedade da *The Walt Disney Company*. Entre as diversas razões que motivaram a realização desta pesquisa de *corpus* baseada em *Star Wars*, a influência da saga foi uma das mais importantes. Trata-se de uma grande franquia multimídia em constante expansão desde os anos 70, que mobiliza, diariamente, equipes de tradução audiovisual que precisam encarar a difícil padronização exigida de um universo tão vasto. Atualmente, as altas demandas das empresas de dublagem e legendagem geraram um ritmo verdadeiramente frenético, com prazos extremamente curtos e projetos cada vez maiores. Entre filmes, animações e séries de TV, a franquia *Star Wars* tem, atualmente, dezenas de títulos em diferentes estágios de produção, que serão em algum momento finalizados e enviados à essas empresas para suas devidas traduções e adaptações. Considerando o trabalho dos tradutores de dublagem e legendagem, que lidam com complexos prazos e não necessariamente possuem domínio de um vasto universo ficcional como o de *Star Wars*, um auxílio terminológico é naturalmente bem-vindo e motiva o desenvolvimento deste trabalho.

Cada aspecto de *Star Wars* apresenta um vasto universo terminológico passível de ser explorado. A geografia da saga especifica diversos planetas, florestas e regiões, com nomes e características próprias. A política, por sua vez, envolve facções e organizações, como a República, a Aliança Rebelde, o Império e a Primeira Ordem; tendo cada instituição uma série de afiliações, entre patentes militares, tecnologias utilizadas e elementos culturais e econômicos. A franquia também se aprofunda em diferentes raças, crenças, filosofias, instituições e tecnologias, sendo cada uma dessas divisões portadoras de diversos termos, que podem ser estudados literariamente ou trabalhados no universo profissional e amador da tradução audiovisual a partir dos constantes lançamentos em cinema, TV, vídeo game e literatura.

### 1.1 Objetivos

O objetivo geral deste trabalho é criar uma proposta de vocabulário sobre a franquia *Star Wars* na plataforma VoTec (Vocabulário Técnico Online). A intenção, tanto deste trabalho quanto da plataforma, é auxiliar aprendizes de Tradução e tradutores profissionais já atuantes.

É esperado que, com a concretização desta proposta, o trabalho de tradutores que lidam com a franquia seja melhorado (eliminando ou diminuindo a possibilidade de múltiplas traduções para um mesmo termo) e facilitado (a partir da interface de busca terminológica da plataforma VoTec).

Os objetivos específicos são:

- Levantar os candidatos a termos em inglês utilizado nas legendas e livros;
- Levantar os candidatos a termos em português utilizado pelos tradutores;
- Analisar o vocabulário mais utilizado em inglês;
- Analisar esse vocabulário mais utilizado em português, a fim de refletir sobre questões de tradução;
- Disponibilizar a proposta de vocabulário gratuitamente on-line por meio do VOTEC <sup>1</sup> (FROMM, 2007).

Sendo assim, este trabalho parte do princípio de que existe uma comunidade expressiva de tradutores que podem ser beneficiados com a inclusão do vocabulário de *Star Wars* no VoTec. No próximo capítulo, serão expostos os referenciais teóricos que pautaram a pesquisa e a execução deste trabalho.

---

<sup>1</sup> Disponível em: <http://ic.votec.ileel.ufu.br/>

## 2 SOBRE *STAR WARS*

Quando lançado nos cinemas em 1977, *Guerra nas Estrelas* de George Lucas foi um fenômeno instantâneo. Na época, ganhou duas sequências (O Império Contra-ataca, 1980, e O Retorno de Jedi, 1983) que formaram o que é hoje conhecido como a trilogia clássica. Uma segunda trilogia foi lançada entre 1999 e 2005 e uma terceira entre 2015 e 2019. Estes nove filmes compõem a Saga Skywalker, que conta a história de uma família de pessoas sensíveis à Força que lutam pelo destino da galáxia, desde a ascensão e queda de Anakin Skywalker até a jornada da jovem Rey em encontrar seu lugar no universo. A franquia também conta com filmes independentes à saga, entre animações (*Star Wars: A Guerra dos Clones*, 2008), filmes produzidos para a televisão (*The Star Wars Holiday Special*, 1978; *Caravana da Coragem*, 1984 e *Ewoks - A Batalha de Endor*, 1985) e *spin-offs*<sup>2</sup> (*Rogue One: Uma História Star Wars*, 2016 e *Han Solo: Uma História Star Wars*, 2018). Lançado simplesmente como *Guerra nas Estrelas* (*Star Wars*), foi apenas com o sucesso do VHS que os filmes foram renomeados, finalmente aos moldes da visão original do autor George Lucas. O primeiro filme então passa a ser intitulado *Star Wars: Episódio IV – Uma Nova Esperança* (*Star Wars: Episode IV – A New Hope*), indicando que o futuro traria uma nova trilogia sobre o passado desse universo.

Segundo Lucas (2011), não era esperado que o filme se tornasse um sucesso, uma vez que diversos elementos da obra eram exatamente opostos às convenções do cinema de ficção científica e espacial da época. Inspirado por Akira Kurosawa, o primeiro ato de *Uma Nova Esperança* estabelece quem são alguns dos personagens mais poderosos e relevantes da trama, mas, em contrapartida, a história é definida a partir do ponto de vista das duas figuras de menor importância. No Episódio 4, estes personagens são os droides<sup>3</sup> C-3PO e R2-D2, enquanto Kurosawa faz a mesma dinâmica com Tahei e Matashichi, em *A Fortaleza Escondida*.

A mitologia de *Star Wars* foi construída, desde o primeiro filme, em torno de elementos comuns da vida real, como a política, a religião e a eterna polarização do bem contra o mal. A religião, por exemplo, é um dos temas centrais na saga, sendo representada pela Força, uma espécie de energia espiritual que conecta todos os elementos vivos do universo. Ela, então, é vista como uma religião para os Jedi, grupo que busca o equilíbrio da galáxia ao seguirem um

---

<sup>2</sup> “Spin Off” é o nome dado para obras que se originam de uma história ou um universo já existente. O objetivo de um spin off é, normalmente, explorar características e pessoas que eram secundárias no material original.

<sup>3</sup> Em todo o corpus em português, 2.819 contextos com a palavra “droide” foram encontrados. Para dróide, com acento agudo, foram encontrados apenas 194.

rigoroso código moral atrelado à princípios éticos e filosóficos. A política, por sua vez, também desempenha importante papel na saga, cuja narrativa é desenvolvida em uma galáxia com fraturas políticas e ideológicas, em que personagens constantemente disputam pelo domínio do poder. Temas como tirania, democracia, corrupção e opressão política são abordados ao longo das trilologias por meio de complexas questões políticas, como a corrupção no Senado e a tirania do Império Galáctico.

A partir da grandiosidade do universo estabelecido na trilogia clássica, uma enorme quantidade de material adicional passou a ser produzido desde o lançamento do primeiro filme em 1977. Entre romances, quadrinhos, enciclopédias, jogos e, mais recentemente, diversas séries de televisão, o universo *Star Wars* ganhou uma mitologia estabelecida em centenas de exemplares em diversas mídias, constituindo um vasto universo que engloba uma expressiva quantidade de planetas, veículos, grupos, tecnologias, armas e droides, por exemplo.

Com a constante expansão do universo *Star Wars*, a franquia se tornou muito cedo uma marca, presente mundialmente ao se tornar um símbolo da cultura pop. Um exemplo dos esforços, até então de George Lucas e a *20th Century Fox*, de fortalecer o caráter mercadológico da franquia, é evidente na forma como ela passa a ser reconhecida mundialmente com o fim da primeira trilogia. Os filmes, que até então tinham seus títulos traduzidos ao redor do mundo (*La guerre des étoiles*, na França; *Guerre stellari*, na Itália e *Guerra nas Estrelas*, no Brasil, por exemplo), passam a ser conhecidos mundialmente como *Star Wars* a partir do lançamento da segunda trilogia, em 1999. Devido a este fenômeno, este trabalho considera o nome “*Star Wars*” e não “*Guerra nas Estrelas*”.

Na figura abaixo, é possível ver como a marca era anunciada no Brasil antes e depois do lançamento da segunda trilogia.

Figura 1 - VHS de Ameaça Fantasma com novo título “*Star Wars*”. Ao lado, cartaz brasileiro anunciando “O Império Contra-Ataca” como “Guerra nas Estrelas”.



Fonte: Anos Dourados: Imagens & Fatos (2022) e VHS Mania (2016)

Desde o lançamento do primeiro filme, é seguro dizer que a franquia garantiu um importante espaço na cultura popular. Em 2015, com o lançamento de *Star Wars: Episódio VII – O Despertar da Força* (*Star Wars: Episode VII - The Force Awakens*), a franquia foi oficialmente reanimada e vive seu momento mais ativo, com diversas séries, animadas e em *live-action*, atualmente em atividade, sendo traduzidas, por meio de dublagem ou legendagem, ao redor de todo o mundo.

Sendo assim, é importante reforçar a complexidade desses trabalhos de tradução nas mais diversas mídias. Devido aos curtos prazos exigidos pelas agências, estúdios e empresas que solicitam os trabalhos de tradução e localização, o profissional pode encontrar dificuldades em localizar as especificidades do vocabulário técnico de franquias como *Star Wars*. Tais dificuldades facilitam a incidência de erros, já que um menor tempo de pesquisa pode resultar em equívocos terminológicos.

### 3 REFERENCIAIS TEÓRICOS

Neste capítulo, serão analisados os referenciais teóricos que sustentam a pesquisa. Partiremos dos estudos da Terminologia, da Teoria Comunicativa da Terminologia, de Cabré, da Terminografia, os recursos utilizados – como o *WordsmithTools 7* e o VoTec – e dos estudos da Etnoterminologia de Barbosa (2006), aplicados aqui, não apenas na literatura, mas no audiovisual.

#### 3.1 Terminologia

É chamado de “Terminologia” o estudo referente aos termos e vocabulários específicos de uma determinada área, sendo um campo interdisciplinar que oferece apoio à várias outras disciplinas. Segundo Lara (2004), “Terminologia” apresenta um significado polissêmico, referindo-se tanto à abordagem teórica e metodológica quanto à terminologia concreta. A perspectiva teórica oferece metodologias que descrevem, organizam e transferem conhecimento, estabelecendo princípios que regem a compilação, formação de termos, estruturação de campos conceituais, uso e administração de terminologias. A acepção da terminologia concreta, por sua vez, se refere a um conjunto de termos relacionados a uma língua especializada. (p. 234)

De acordo com Lara (2005), a unidade Terminológica é o objeto da Terminologia. Diferente de uma palavra, que abrange as várias possibilidades de uso da língua, o Termo se refere às palavras utilizadas em situações específicas de comunicação, ou seja, a língua em contextos especializados. Ainda de acordo com Lara (2005), a Terminologia apresenta objetivos teóricos e práticos. A abordagem teórica destaca a posição das terminologias no conhecimento e na prática dos especialistas. O objetivo é a análise, por meio de métodos específicos, do uso das terminologias em diferentes contextos. A abordagem prática, por sua vez, ocorre, por exemplo, nas áreas de “tradução, documentação, normalização linguística e planejamento linguístico” (LARA, 2005, p.6). É seguro dizer, então, que ela desempenha um papel fundamental na elaboração de recursos linguísticos como vocabulários e glossários, por exemplo.

O trabalho terminológico apresenta um crescente uso de ferramentas tecnológicas desenvolvidas no âmbito da Linguística Computacional e vem contando cada vez mais com a metodologia da Linguística de *Corpus*. O objetivo, de acordo com Zamora (2015), é proporcionar uma descrição precisa da terminologia em um determinado campo de

conhecimento, estabelecendo com maior precisão as equivalências linguísticas ou terminológicas nas pesquisas voltadas à tradução (ZAMORA, 2015). No campo da tradução literária, de jogos eletrônicos e de conteúdo audiovisual, tal estudo representa uma possibilidade de avanço na qualidade do resultado de diversos materiais traduzidos. Tratando-se de longos universos e sagas já estabelecidas, é muito comum que exista um vasto campo terminológico que, sem o devido cuidado, pode gerar diferentes traduções de um mesmo termo conforme novos produtos são lançados e traduzidos por pessoas e equipes distintas. O levantamento terminológico através de um *corpus* pode avaliar os padrões linguísticos historicamente estabelecidos por frequência, evitando ou eliminando a possibilidade de várias traduções para um mesmo termo.

Sendo assim, é possível concluir que a Terminologia desempenha um papel fundamental no estudo de linguagens específicas, e ao compreender a importância da área, é possível explorar a sua aplicação em diversos contextos. O estudo de uma determinada terminologia garante a compreensão da natureza de termos e seus usos, o que auxilia na padronização e otimização do trabalho em diversos campos. Em síntese, a Terminologia é, de acordo com Dias (2000, p.90), como prática, “um conjunto de métodos e atividades voltado para coleta, descrição, processamento e apresentação de termos; como produto, é um conjunto de termos, ou vocabulário, de uma determinada especialidade”.

### **3.2 Teoria Comunicativa da Terminologia**

Ao analisar os fundamentos de Maria Teresa Cabré, Sales (2007) percebe que a Teoria Comunicativa da Terminologia se apresenta de forma antagônica à Teoria Geral da Terminologia, que a partir de uma noção consideravelmente rígida em relação à análise dos termos, apresentava pouca flexibilidade quanto ao estudo dos contextos. Dessa forma, o diferencial da Teoria Comunicativa da Terminologia está na valorização do “caráter flexível do processo comunicativo dentro do universo especializado” (SALES, 2007, p.8).

Sendo assim, Cabré (1999) estabelece a Teoria Comunicativa da Terminologia de forma a reconhecer o caráter interdisciplinar da Terminologia, a relação entre conhecimentos gerais e especializados e a interdisciplinaridade e pluralidade dos termos. Fora isso, a teoria também valoriza a polissemia das unidades lexicais, a existência das sinonímias e a diversificação do discurso perante o tema.

### 3.3 Etnoterminologia

Segundo Latorre, a Etnoterminologia consiste na “zona de intersecção entre os estudos dos universos de discursos literários e os das linguagens de especialidades ou terminologias” (2013, p.73). Sendo assim, este trabalho parte do conceito de Etnoterminologia, muito aplicado no campo literário, para analisar o discurso técnico em um universo multimídia de fantasia científica.

O universo de *Star Wars* é repleto de elementos existentes na realidade, mas que são reimaginados na obra de forma a estabelecer uma rica composição de mundo ancorada em bases conhecidas pelo público. Um exemplo é o próprio conceito de “Força”, que ganha valores religiosos quando aplicado na narrativa (como pode ser visto na definição presente no VoTec). Esse movimento é refletido na análise terminológica quando são encontrados termos que, dentro do universo ficcional estabelecido, admitem um novo sentido e/ou uma mais ampla definição. Sobre essa característica entre a língua comum e as linguagens de especialidade, Barbosa alerta que “[...] as unidades lexicais são plurifuncionais. O estabelecimento preciso de sua função depende de sua inserção em uma norma discursiva, que determina, então, o estatuto de vocábulo ou de termo”. (BARBOSA, 2006).

De acordo com Fromm (2011), a busca pela verossimilhança nas narrativas faz com que a terminologia seja encontrada em obras ficcionais. Para tal conclusão, o autor partiu da análise de *Star Trek*, uma série relativamente próxima aos nossos estudos (também uma saga espacial), e notou a presença massiva de terminologia própria advinda de campos reais, como a Física, a Astronáutica e a Astronomia, por exemplo. Em *Star Wars* é possível encontrar movimentos semelhantes, como nos campos da Física (*Hyperspace*, *Lightspeed*) e do Militarismo (*Star Destroyer*), por exemplo

Sendo assim, este trabalho parte do princípio de que o mesmo pode ser aplicado ao analisar o universo *Star Wars*, considerando o rico estabelecimento de mundo criado por diversos autores desde os anos 70, entre quadrinhos, livros, séries, animações e a consagrada série de filmes.

### 3.4 Terminografia

A Terminografia é uma disciplina prática profundamente ligada à Terminologia. Finatto (2014) analisa o objetivo desse campo de estudos:



Descrição das propriedades linguísticas, conceituais e pragmáticas das unidades terminológicas de uma ou mais línguas, a fim de produzir obras de referência, tais como dicionários, glossários, vocabulários em formato papel ou eletrônico, bases de dados terminológicos e bases de conhecimento especializado. (FINATTO, 2014, p. 439)

Assim, a Terminografia pode envolver a produção de vocabulários técnicos ou glossários de uma área específica, organizando termos e suas definições de forma sistemática. Trata-se de uma área que envolve a busca e sistematização de termos relevantes de um domínio específico, de forma que outros profissionais possam ser beneficiados com as obras de referência possivelmente abastecidas com esse trabalho.

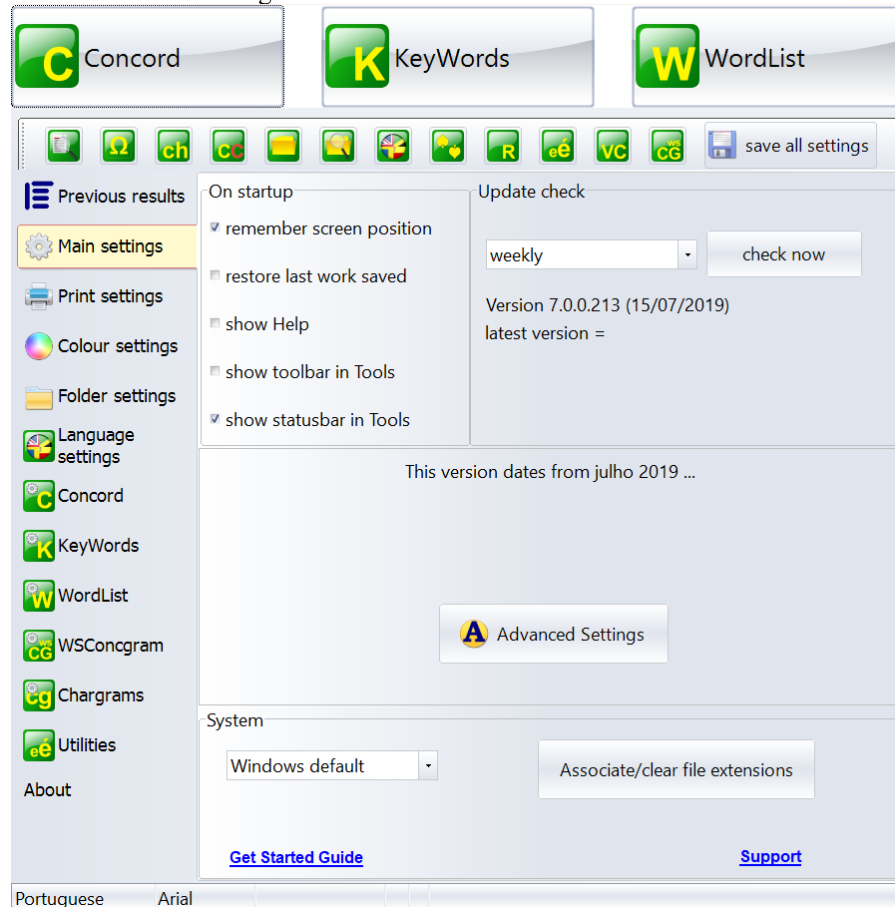
### 3.5 *Wordsmith Tools*

O *Wordsmith Tools* é um *software* desenvolvido pelo linguista Mike Scott em 1996 na Universidade de Liverpool. Mesmo depois de tantos anos, ele continua sendo atualizado, estando, no momento desta escrita, em sua versão 8.0.0.62, lançada em dezembro de 2021. Comercializado pela *Oxford University Press*, o *Wordsmith Tools* integra ferramentas para a análise textual e levantamento de *corpora*, permitindo, por exemplo, o estudo da frequência de palavras, listagem de palavras-chave e identificação de colocações.

Em sua interface, o *Wordsmith Tools* apresenta três ferramentas principais (conforme podemos visualizar na parte superior da figura 2): *Concord*, *Keywords* e *WordList*. Berber Sardinha (2009, p. 9) as resume bem no seguinte parágrafo:

*WordList*: produz listas de palavra contendo todas as palavras do arquivo ou arquivos selecionados, elencadas em conjunto com suas frequências absolutas e percentuais. Também compara listas, criando listas de consistência, onde é informado em quantas listas cada palavra aparece. *Concord*: realiza concordâncias, ou listagens de uma palavra específica (o ‘nódulo’, *node word* ou *search word*) juntamente com parte do texto onde ocorreu. Oferece também listas de colocados, isto é, palavras que ocorrem perto do nódulo. *KeyWords*: extrai palavras de uma lista cujas frequências são estatisticamente diferentes (maiores ou menores) do que as frequências das mesmas palavras num outro corpus (de referência). Calcula também palavras-chave chave, que são chave em vários textos.

A figura abaixo mostra a interface do software.

Figura 2 - Interface do *WordsmithTools 7*.

Fonte: Captura de tela do autor

O software também proporciona ferramentas para tratamento de *corpora*, como o *Text Converter* e o *Corpus Checker*<sup>4</sup>. Segundo Ribeiro, a utilização de ferramentas como o *Wordsmith Tools* ainda pode apresentar certas limitações para os tradutores devido à complexidade de elaboração do corpus (2004). Para isso, se mostra importante a existência de um recurso como o Vocabulário Técnico Online, ou VoTec, de Fromm (2007).

### 3.6 VoTec

O VoTec (Vocabulário Técnico Online) surge como um ambiente *web* de gestão terminológica que busca auxiliar tradutores iniciantes e profissionais no uso de obras terminográficas (FROMM, 2007). O trabalho parte do princípio de que existe uma relevante necessidade de integração entre os campos da Tradução e Terminologia, uma vez que, nas

<sup>4</sup> Ferramentas para conversão de arquivos e checagem de material corrompido, respectivamente.

palavras do autor, “cada vez mais, o tradutor também desempenha o papel de terminólogo e, ocasionalmente, de terminógrafo” (p. 13).

A plataforma é alimentada por pesquisadores a partir de termos levantados previamente por meio de *corpora*, processado por um *software*, como o *Wordsmith Tools*, por exemplo. Juntamente com os candidatos a termos (enumerados a partir da ferramenta *Keywords*), o *software* permite que o pesquisador tenha contato com todos os contextos em que a palavra em análise foi utilizada (ferramenta *Concordance*), permitindo que sejam cadastrados, no VoTec, o termo, os contextos que melhor expliquem sua definição, e a definição final, criada pelo pesquisador a partir unicamente dos contextos selecionados. Os termos cadastrados na plataforma passam por um processo de aprovação pelo administrador, que deve analisar a qualidade dos contextos cadastrados e suas fontes, os traços distintivos, as informações enciclopédicas, o conceito final e a definição de cada termo.

Para a elaboração deste trabalho, foi utilizada a versão do ambiente VoTec que pode ser vista na figura abaixo.

Figura 3 - Interface de administrador do VoTec.

The screenshot shows the VoTec administrator interface. At the top, it says 'Vocabulário Técnico Online' and 'Tela Cheia | English'. Below that, it greets the user: 'Bem Vindo, Guilherme Rodrigues Ferreira'. It also shows the last access time: 'Seu último acesso foi em: 01/06/2023 12:43'. There are four buttons: 'Trocar Senha', 'Novo Termo', 'Cadastro de Fontes', and 'Sair'.

The main content is divided into two sections:

**Fichas ainda não aprovadas**

Termo	Língua	Coletado em	Cadastrada por	Ações
Mandalore		11/05/2023	Guilherme Rodrigues Ferreira	editar
Mandalore		11/05/2023	Guilherme Rodrigues Ferreira	editar
Starfighter		11/04/2023	Guilherme Rodrigues Ferreira	editar
Caça Estelar		11/04/2023	Guilherme Rodrigues Ferreira	editar

Total de Termos: 4

**Fichas aguardando revisão**

Termo	Língua	Coletado em	Cadastrada por	Ações
Nenhum termo para revisar				

Total de Termos: 0

At the bottom right, it says: '01/06/2023 13:29 © 2007 FFLCH - ICMC Jr.'

Fonte: captura de tela do autor.

Em linhas gerais, a abordagem da Linguística de *Corpus* é base, principalmente, como visto, para questões de Terminologia, Terminografia, Etnoterminologia, Teoria Comunicativa da Terminologia, e para as plataformas e *softwares* que utilizamos. No próximo capítulo, analisaremos como foram os processos de levantamento, limpeza, sistematização e exploração do *corpus*.

#### 4 METODOLOGIA

O *corpus* deste trabalho foi levantado no decorrer de três meses, sendo primeiro compilados os arquivos em inglês e, posteriormente, a versão traduzida de cada um. Em ambos os idiomas, o *corpus* foi dividido em cinco partes, quais sejam: *Animated Series/Animações*, *Books/Livros*, *Films/Filmes*, *TV Series/Séries* e *Everything/Tudo*, sendo a última a junção de todas as outras – importante para a criação da *WordList*, posteriormente.

Ao todo, 760 arquivos foram baixados, sendo 380 no idioma original e 380 traduzidos em português. Para a composição do *corpus*, foi selecionado apenas o conteúdo considerado canônico<sup>5</sup> lançado até o momento da coleta e que atendesse facilidade de levantamento, o que exclui, por exemplo, as revistas em quadrinhos, que mesmo canônicas, apresentam considerável dificuldade em transposição para documento de texto.

Os livros correspondem à parcela mais expressiva do *corpus*, sendo quase três vezes maior que o segundo colocado, ou seja, as animações. Este material corresponde à um total de 38 romances canônicos que já foram traduzidos, oficialmente ou não, para a língua portuguesa. As obras selecionadas são constituintes, dentro da linha do tempo de Star Wars, da Era da Alta República, a Era da República, a Era Imperial e a Era da Nova República.

Foi desconsiderada toda a série *Legends*, anteriormente conhecida como Universo Expandido, assim como contos, séries infantis, livros de arte e *behind-the-scenes*, dicionários e enciclopédias. Claramente, os livros representam a parcela mais rica da construção de universo da saga *Star Wars*, tendo sua primeira obra sido publicada antes mesmo do lançamento do universo cinematográfico<sup>6</sup>. Logo, caso todo esse material fosse considerado para o trabalho, o resultado facilmente ultrapassaria mil arquivos, tornando a complexidade de levantamento consideravelmente maior.

Em inglês, os livros foram encontrados em repositórios online<sup>7</sup>. Todos foram baixados em formato EPUB e posteriormente convertidos para TXT, o que apresentou certa dificuldade, uma vez que a conversão, inicialmente, não conseguiu preservar corretamente a codificação de caracteres do texto original – um problema comum entre conversores, principalmente ao lidar com a acentuação da língua portuguesa.

---

<sup>5</sup> Em *Star Wars*, são consideradas canônicas as histórias que continuam a saga original de George Lucas e são reconhecidas como parte da história principal e definitiva, ou seja, não são consideradas “universo expandido”.

<sup>6</sup> *Star Wars: From the Adventures of Luke Skywalker*, 1976.

<sup>7</sup> Z-Library, principalmente.

Para os livros traduzidos em português, encontrar as traduções também foi um processo que demandou tempo. A solução encontrada para a conclusão da pesquisa foi contar com o contínuo trabalho do grupo amador *Tradutores dos Whills*<sup>8</sup>, que traduz para o português brasileiro toda a literatura de *Star Wars* ainda não publicada oficialmente no Brasil. Foi particularmente interessante analisar a forma como a equipe manteve a identidade visual e o projeto gráfico original, tornando mais acessível a leitura de obras que talvez nunca cheguem às livrarias brasileiras. Sendo assim, mesmo que muitas obras tenham sido traduzidas oficialmente e utilizadas no levantamento do *corpus*, é importante mencionar a presença da tradução amadora neste projeto, integrante não apenas dos livros, mas de todas as outras categorias levantadas.

Os livros que compõem o *corpus* são:

Quadro 1 – *Corpus* de Livros (Títulos originais em inglês).

A New Dawn
Aftermath
Aftermath: Empire's End
Aftermath: Life Debt
Ahsoka
Alphabet Squadron
Battlefront - Twilight Company
Battlefront II - Inferno Squad
Bloodline
Brotherhood
Catalyst
Dark Disciple
Heir to the Jedi
Leia, Princess of Alderaan
Lords of the Sith
Lost Stars
Master & Apprentice
Phasma
Queen's Peril
Queen's Shadow
Rebel Rising
Resistance Reborn
Rogue One
Tarkin
The Force Awakens
The High Republic - Into the Dark

<sup>8</sup> “Somos uma Comunidade que se dedica a fomentar e divulgar ainda mais a Saga de Star Wars. Somos fãs que estão colocando o seu tempo e o seu amor por Star Wars a disposição através de materiais como Livros, HQs e etc. [...] Todo o trabalho de tradução, revisão e layout das histórias traduzidas aqui, são feitas por fãs de Star Wars e com o único propósito de compartilhá-lo com outras que falam a língua portuguesa, em especial no Brasil.” Disponível em: <https://tradutoresdoswhills.wordpress.com/sobre/>. Acessado em: 1 jun. 2023.

The High Republic - Light of the Jedi
The High Republic - Midnight Horizon
The High Republic - Out of the Shadows
The High Republic - The Rising Storm
The High Republic – The Fallen Star
The Last Jedi
The Rise of Skywalker
Thrawn - Alliances
Thrawn - Treason
Thrawn Ascendancy - Chaos Rising
Thrawn Ascendancy - Greater Good
Thrawn Ascendancy - Lesser Evil

Fonte: elaborado pelo autor.

Após o levantamento dos livros, foi iniciado o processo de busca dos elementos constituintes da segunda categoria mais expressiva do *corpus*: as animações.

As animações que compõe o *corpus* são:

Quadro 2 – *Corpus* de animações (Títulos originais em inglês).

The Bad Batch
Droids
Ewoks
Star Wars Rebels
Star Wars Resistance
The Clone Wars
Visions

Fonte: elaborado pelo autor.

A expressividade desse *corpus* se dá pela longevidade<sup>9</sup> de algumas dessas séries, como *The Clone Wars*, que chegou no significativo número de 133 episódios divididos entre sete temporadas. *Rebels*, por sua vez, foi encerrada após 75 episódios em quatro temporadas. Em número de arquivos, o *corpus* é consequentemente o de maior valor, com 304 arquivos de legenda para cada idioma.

Além das animações, legendas também foram utilizadas como itens para a composição dos *corpora* de filmes e séries de tv.

<sup>9</sup> Novos lançamentos foram realizados no período de escrita desta pesquisa, como a série *Andor*, *Tales of the Jedi* e as segundas temporadas de *The Bad Batch* e *Visions*.

Compõem o *corpus* dos filmes:

Quadro 3 – *Corpus* de filmes (Títulos originais em inglês).

Star Wars: Episode IV - A New Hope (1977)
The Star Wars Holiday Special (1978)
Star Wars: Episode V - The Empire Strikes Back (1980)
Star Wars: Episode VI - Return of the Jedi (1983)
Caravan of Courage: An Ewok Adventure (1984)
Ewoks: The Battle for Endor (1985)
Star Wars: Episode I - The Phantom Menace (1999)
Star Wars: Episode II - Attack of the Clones (2002)
Star Wars: Episode III - Revenge of the Sith (2005)
Star Wars: The Clone Wars (2008)
Star Wars: Episode VII - The Force Awakens (2015)
Rogue One: A Star Wars Story (2016)
Star Wars: Episode VIII - The Last Jedi (2017)
Solo: A Star Wars Story (2018)
Star Wars: Episode IX - The Rise of Skywalker (2019)

Fonte: elaborado pelo autor.

E das séries:

Quadro 4 – *Corpus* de séries (Títulos originais em inglês).

The Mandalorian
Obi-Wan Kenobi
The Book of Boba Fett

Fonte: elaborado pelo autor.

Para as séries *live-action* e em animação, foram considerados todos os episódios lançados até o momento da coleta do *corpus*: *The Mandalorian* está em sua segunda temporada, enquanto *Obi-Wan Kenobi* e *The Book of Boba Fett* em suas primeiras. Da mesma forma, quanto às animações, *The Bad Batch* e *Visions* estão em suas primeiras temporadas. *The Clone Wars* em sua sétima temporada, *Rebels* em sua quarta e *Resistance* em sua segunda. *Ewoks* e *Droids* foram finalizadas nos anos 80.

#### 4.1 Tratamento do *Corpus*

Os *corpora* de ambos os idiomas foram divididos em cinco partes. Em português: “Desenhos”, “Filmes”, “Livros”, “Séries” e “Tudo”. Em inglês: “*Animated Series*”, “*Films*”, “*Books*”, “*TV Series*” e “*Everything*”. Foram salvos todos os arquivos de cada *corpus* e suas respectivas *Wordlists*, separados por idiomas em pastas distintas.

Para o tratamento do *corpus*, o primeiro passo foi a conversão de todas as legendas e livros para o formato TXT. Para as legendas, bastou a utilização do *SubtitleEdit*, que é capaz de converter e limpar os arquivos em massa. Para os livros, foram utilizadas ferramentas online de conversão de arquivos, como o *Convertio*. Uma segunda conversão, desta vez de codificação, foi necessária para a utilização do *corpus* no *WordSmith Tools 7*. Para isso, dentro do próprio *software*, utilizou-se o recurso *Text Converter*, que faz com que os arquivos operem em *Unicode*.

A falha no processo de tratamento do *corpus* pode gerar erros de codificação como exemplificado abaixo, na figura 4.

Figura 4 - Erro de codificação no corpus.

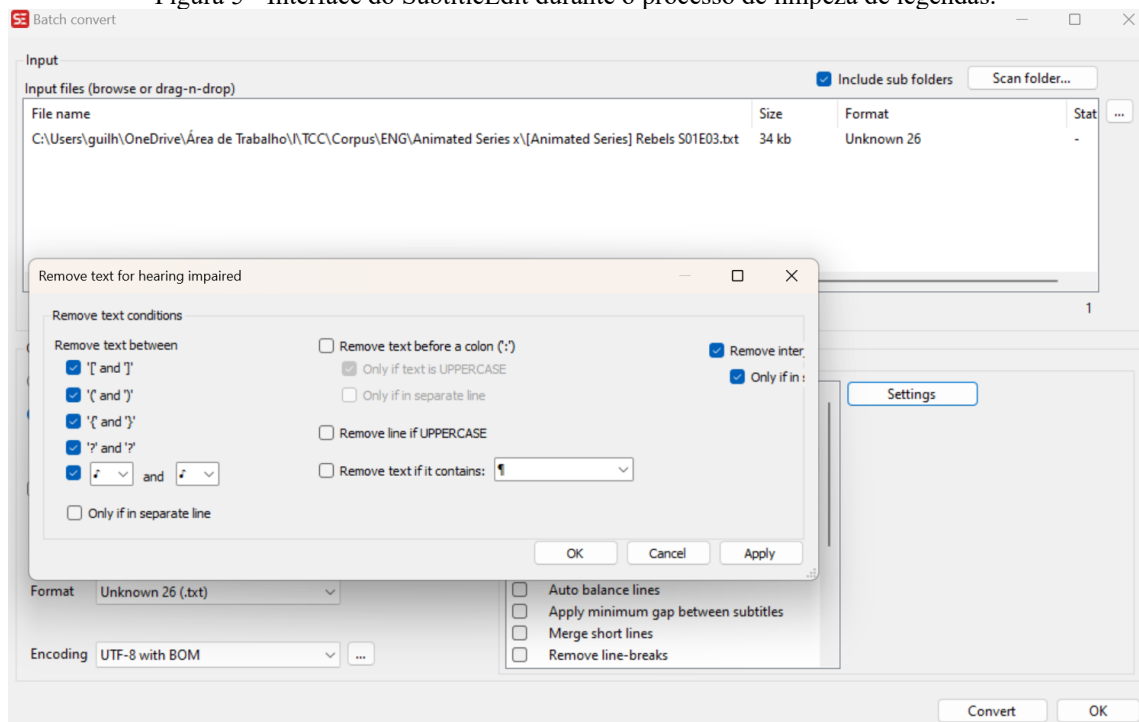
67	EVEN
68	KNOW
69	DIDNÆ™T
70	DOWN
71	THROUGH

Fonte: captura de tela do autor.

O *corpus* das animações, assim como das séries e dos filmes, é constituído de legendas, que foram encontradas gratuitamente na internet. As legendas disponíveis em inglês são frequentemente as consideradas “HI”, ou seja, *Hearing Impaired* (para aqueles com perda de audição total ou significativa). Isso significa que essas legendas apresentam descrições do áudio que não correspondem às falas, como músicas, gritos, choros, risadas, sons e diferentes entonações. Considerando que tais descrições não constituem o foco de nossas análises e, por conseguinte, influenciariam os dados obtidos, as legendas em inglês passaram por um processo adicional para a remoção de todas essas informações. Inicialmente, isso foi realizado de forma manual, e posteriormente com o auxílio do *software SubtitleEdit*, que realiza esse processo automaticamente (conforme podemos verificar nas figuras 5 e 6).



Figura 5 - Interface do SubtitleEdit durante o processo de limpeza de legendas.



Fonte: captura de tela do autor.

Figura 6 - Diferença do material antes e após a limpeza.

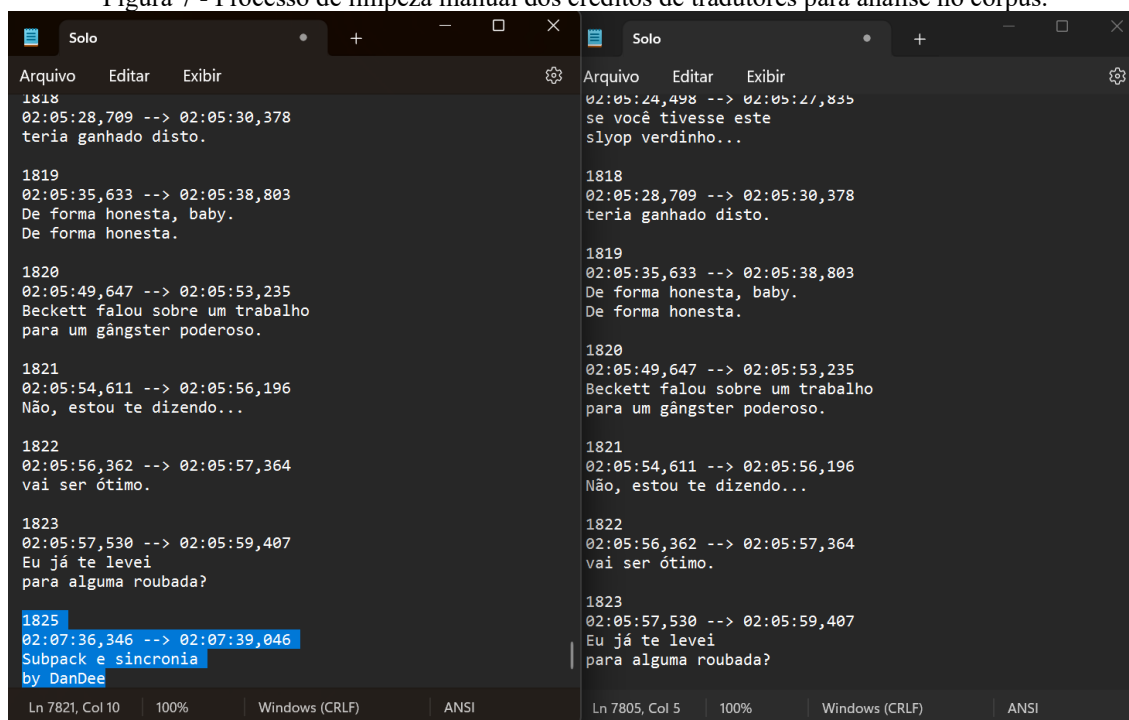
00:00:47,448 --> 00:00:49,448 [Yawning]	always confused me too.
19 00:00:49,450 --> 00:00:53,118 I really thought this Jedi stuff would be more interesting.	10: 00:00:35.08 00:00:37.13 But Master Yoda sure used to say it a lot.
20 00:00:53,169 --> 00:00:55,204 No wonder the old religion died.	11: 00:00:39.04 00:00:42.11 I really thought this Jedi stuff would be more interesting.
21 00:00:55,255 --> 00:00:57,122 [Droid Language]	12: 00:00:42.13 00:00:44.07 No wonder the old religion died.
22 00:00:57,173 --> 00:01:00,292 Come on, kid. Do something. Amuse me.	13: 00:00:46.07 00:00:49.06 Come on, kid. Do something. Amuse me.
23 00:01:00,294 --> 00:01:02,261 "Use the Force!"	14: 00:00:49.08 00:00:51.07 Use the Force!
24 00:01:03,346 --> 00:01:06,298 [Yelping, grunts] Whoa!	15: 00:00:52.00 00:00:54.09 Whoa...
25 00:01:06,300 --> 00:01:08,300 [Laughs]	16: 00:00:59.00 00:01:00.10 Does he have to be here?
	17: 00:01:00.12 00:01:02.13 He's annoying, but there will always be distractions.
	18: 00:01:02.13 00:01:05.09 You need to learn to focus through them.

Fonte: Captura de tela do autor.

No caso das legendas traduzidas, apesar dos elementos HI serem relativamente raros, os itens precisaram passar por um processo de limpeza por um motivo distinto. Tais legendas, principalmente no campo das animações e das séries mais antigas, foram produzidas por grupos amadores, que creditavam o trabalho em todos os episódios em dois ou três momentos. Mesmo entendendo que o trabalho dessas equipes foi crucial para a conclusão desta pesquisa, os créditos precisaram ser removidos para que o *corpus* não fosse poluído.

Diferente dos elementos HI, que são sistematicamente expostos na legenda entre colchetes, esses créditos não apresentam um padrão, tornando impossível para um *software* como o *SubtitleEdit* realizar a remoção automática. Sendo assim, a limpeza das legendas traduzidas foi um trabalho manual (que pode ser visto abaixo na figura 6).

Figura 7 - Processo de limpeza manual dos créditos de tradutores para análise no corpus.



Fonte: Captura de tela do autor.

## 4.2 Seleção de candidatos a termos

Ao realizar a análise do *corpus* com o *Wordsmith Tools*, as palavras-chave são identificadas com base em sua frequência em comparação com um *corpus* de referência. Essas palavras-chave destacam-se por sua ocorrência significativa em relação ao *corpus* geral, fornecendo uma relevante referência sobre os principais termos dos textos analisados – para esta pesquisa, o universo canônico de *Star Wars*.

Os termos escolhidos para cadastro na plataforma VoTec foram selecionados a partir das primeiras 500 primeiras palavras da lista de *KeyWords* gerada pelo *WordSmith Tools*. Para gerar esta lista de palavras-chave, foi utilizado como *corpus* de referência o COCA (*Corpus of Contemporary American English*) restrito em suas primeiras 100 mil palavras. O resultado pode ser conferido na figura 8, abaixo:

Figura 8 - *KeyWord List*.

N	Key word	Freq.	%	Texts	RC. Freq.	%	BIC	Log_L	Log_R	P
1	BACK	9.445	0,20	360	6.366		64.181,05	64.200,92	7,07	0,0000000000
2	TIME	7.762	0,17	364	4.065		55.012,21	55.032,07	7,44	0,0000000000
3	DIDN	5.533	0,12	31	0		49.994,90	50.014,76	139,30	0,0000000000
4	THRAWN	5.392	0,12	22	0		48.720,35	48.740,21	139,26	0,0000000000
5	JEDI	5.215	0,11	242	0		47.120,38	47.140,25	139,21	0,0000000000
6	GOOD	4.646	0,10	362	113		40.910,84	40.930,70	11,87	0,0000000000
7	LL	4.171	0,09	31	0		37.683,29	37.703,16	138,89	0,0000000000
8	VE	3.843	0,08	33	0		34.718,39	34.738,25	138,77	0,0000000000
9	EYES	3.796	0,08	123	461		31.383,93	31.403,79	9,55	0,0000000000
10	LEIA	3.260	0,07	33	0		29.448,44	29.468,30	138,53	0,0000000000
11	DON	4.104	0,09	31	2.176		29.020,97	29.040,83	7,42	0,0000000000
12	WASN	3.189	0,07	31	0		28.806,65	28.826,51	138,50	0,0000000000
13	FIND	3.167	0,07	347	975		24.108,46	24.128,32	8,20	0,0000000000
14	CAPTAIN	2.625	0,06	178	18		23.493,35	23.513,21	13,69	0,0000000000
15	SHIP	5.839	0,13	285	26.683		22.729,52	22.749,38	4,31	0,0000000000
16	VOICE	3.371	0,07	71	2.580		22.363,94	22.383,80	6,89	0,0000000000
17	VADER	2.297	0,05	39	0		20.743,54	20.763,40	138,03	0,0000000000
18	SIDE	2.475	0,05	182	253		20.673,08	20.692,94	9,79	0,0000000000
19	DROID	2.452	0,05	201	256		20.455,55	20.475,41	9,76	0,0000000000
20	PADMÉ	2.255	0,05	22	0		20.363,89	20.383,75	138,00	0,0000000000
21	ANAKIN	2.237	0,05	115	0		20.201,18	20.221,04	137,99	0,0000000000
22	FELT	2.531	0,05	64	848		19.069,94	19.089,80	8,08	0,0000000000
23	THINGS	2.148	0,05	260	45		18.958,82	18.978,68	12,08	0,0000000000
24	COULDN	2.091	0,05	31	0		18.881,43	18.901,29	137,89	0,0000000000
25	SHIPS	2.237	0,05	142	397		17.976,46	17.996,33	9,00	0,0000000000
26	OBI	1.988	0,04	92	0		17.950,38	17.970,24	137,82	0,0000000000
27	ROOM	1.872	0,04	113	157		15.800,18	15.820,05	10,08	0,0000000000
28	MAN	1.927	0,04	160	288		15.693,40	15.713,27	9,25	0,0000000000
29	HADN	1.711	0,04	31	0		15.446,48	15.466,34	137,60	0,0000000000
30	DROIDS	1.735	0,04	178	78		15.021,75	15.041,62	10,98	0,0000000000
31	POE	1.583	0,03	24	0		14.289,44	14.309,30	137,49	0,0000000000
32	POWER	1.813	0,04	220	459		14.092,02	14.111,88	8,49	0,0000000000
33	TARKIN	1.506	0,03	22	0		13.593,41	13.613,27	137,42	0,0000000000

Fonte: captura de tela do autor.

Os termos extraídos das 500 primeiras *KeyWords* foram classificados em nove categorias semânticas específicas, que posteriormente se tornaram subáreas no banco de dados do VoTec.

As categorias e os termos pré-selecionados são indicados no quadro abaixo.

Quadro 5 – Termos pré-selecionados.

Veículos	Espécies	Lugares	Grupos	Armas	Droides	Tecnologia	Trabalho	Religião
Starfighter	Wookie	Coruscant	Rebels	Lightsaber	Astromech	Droid	Stormtrooper	Force
Star Destroyer	Ewok	Alderaan	Clones	Blaster	Protocol Droid	Hyperspace	Twilight Soldier	Sith
Imperial Star Destroyer	Hutts	Naboo	Separatists	Laser Cannons	Medical Droid	Hyperdrive	First Senator	Jedi
Spaceport		Mandalore	Youngling		Server Droid	Datapad		
		Endor	Padawan		Pill Droid	HoloNet		
					Battle Droid	Lightspeed		
					Tactical Droid			
					Droid Tri Fighter			
					Moderator Droid			

Fonte: elaborado pelo autor.

No quadro acima estão todos os termos encontrados, no recorte das primeiras palavras-chave, de cada categoria. Para sequencial cadastro no VoTec, foram escolhidos dois exemplares de cada categoria que melhor exemplificassem o uso da plataforma e a utilidade da proposta.

### 4.3 Exploração das linhas de concordância

A exploração dos termos foi possível graças à função *Concord*, presente no *WordSmith Tools*, que permite aos pesquisadores identificar os contextos em que uma palavra ou expressão específica ocorre no *corpus*. Por exemplo, se o pesquisador está trabalhando com um *corpus* de *Star Wars* em inglês, é possível encontrar todos os contextos em que a palavra “*Force*” foi utilizada, sejam eles contextos culturais, científicos, religiosos etc. Fora isso, também é possível analisar as palavras que rodeiam o termo, seu contexto, sendo possível descobrir até mesmo novos fraseologismos a partir disso – como foi o caso de *Imperial Star Destroyer*, que mostrou relevância durante a exploração dos contextos de *Star Destroyer*.

Para o cadastro no VoTec, é preciso levantar exemplos contextuais que demonstrem diferentes conceitos sobre o termo estudado, de forma que seja possível produzir uma definição terminológica a partir da palavra em uso.

Abaixo, exemplos de contextos para definir o termo “Ewok”:

Figura 9 - Contextos encontrados para "Ewok".

N	Concordance	Set	Tag	Word	# Sent	Sent	Para	Para	H.	H.	Sect	Sect	File	Date	%	
					# Pos.	# Pos.	# Pos.	# Pos.	# Pos.	# Pos.	# Pos.	# Pos.				
49	Ewok. 99: 00:07:43.04 00:07:45.15 Quer dizer que o Ewok que vimos... 100: 00:07:45.17 00:07:48.03			Ewok	817	94	9	0	816			0	816	[Desenho] Ewoks S01E13.txt	2022/nov/14 00:	35%
50	242: 00:18:21.05 00:18:23.22 Esse não pode ser o Ewok que me derrotará. 243: 00:18:25.22 00:18:			Ewok	1.821	252	10	01.	0			01.	0	[Desenho] Ewoks S02E06.txt	2022/nov/14 00:	82%
51	gorjeta. 319: 00:22:20.13 00:22:24.11 Acho que os Ewok fazem mal para nossos negócios. 320: 00:22:			Ewok	2.564	324	8	02.	3			02.	3	[Desenho] Ewoks S01E05.txt	2022/nov/14 00:	95%
52	02:48.20 00:02:52.09 Pronto. Kneesaa, você é outra Ewok. 41: 00:02:52.11 00:02:54.14 Sim. Mas qual?			Ewok	368	30	6	0	367			0	367	[Desenho] Ewoks S02E13.txt	2022/nov/14 00:	15%
53	Chirpa. 43: 00:04:13.08 00:04:15.18 Nem pensar. Ewok. 44: 00:04:16.11 00:04:19.19 Venha e se junte			Ewok	353	33	7	0	352			0	352	[Desenho] Ewoks S02E04.txt	2022/nov/14 00:	15%
54	:26.05 00:17:28.07 O que faremos com a pequena Ewok? 216: 00:17:28.20 00:17:30.00 E a Kneesaa!			Ewok	1.669	221	11	01.	8			01.	8	[Desenho] Ewoks S02E08.txt	2022/nov/14 00:	74%
55	: 00:08:41.09 00:08:46.16 Então, minha pequena Ewok, não é divertido ser um Jinda? 137: 00:08:			Ewok	1.146	116	8	01.	5			01.	5	[Desenho] Ewoks S01E05.txt	2022/nov/14 00:	42%
56	- uma pequena bolota dada a ela pelo pequeno Ewok conhecido como Wickel. Ela cultivou a planta			Ewok	60.2703	3	19	03.	9			03.	9	[Livro] Aftermath - Divida de Honra.txt	2022/out/14 00:	91%
57	:47. 00:04:33.13 00:04:37.07 Agora, pequeno Ewok, junte-se ao meu totem. 48: 00:04:39.19 00:04:			Ewok	379	37	7	0	378			0	378	[Desenho] Ewoks S02E04.txt	2022/nov/14 00:	17%
58	:69. 00:05:51.18 00:05:55.07 - Teebo? - Pequeno Ewok, junte-se ao meu totem. 70: 00:05:57.20 00:05:			Ewok	553	65	4	0	552			0	552	[Desenho] Ewoks S02E04.txt	2022/nov/14 00:	24%
59	:134. 00:10:07.19 00:10:12.08 Nada mau, pequeno Ewok, mas tenho mais azar para você. 135: 00:10:			Ewok	1.139	139	8	01.	8			01.	8	[Desenho] Ewoks S02E11.txt	2022/nov/14 00:	44%
60	a árvore-santuário que lhe fora dada pelo pequeno Ewok Wickel. Ela nunca foi capaz de sentir a árvore			Ewok	77.8177	6	14	07.	6			07.	6	[Livro] Aftermath - Fim do Império.txt	2022/out/14 00:	117%
61	:14.06.05 Eles devem ter lhe dado aquele pirralho Ewok 174: 00:14:06.07 00:14:08.03 para acalmá-la.			Ewok	1.439	166	12	01.	8			01.	8	[Desenho] Ewoks S01E12.txt	2022/nov/14 00:	64%
62	:217. 00:17:57.08 00:18:02.03 Que linda princesa Ewok. 218: 00:18:02.05 00:18:03.10 Não. 219: 00:			Ewok	1.974	263	8	01.	3			01.	3	[Desenho] Ewoks S02E02.txt	2022/nov/14 00:	79%
63	:214. 00:17:43.13 00:17:46.13 Temos uma princesa Ewok para a Rainha Lesmona. 215: 00:17:47.23 00:			Ewok	1.955	250	8	01.	4			01.	4	[Desenho] Ewoks S02E02.txt	2022/nov/14 00:	78%
64	:200. 00:16:37.01 00:16:40.21 Nossa, uma princesa Ewok para a Rainha Lesmona. 201: 00:16:45.04 00:			Ewok	1.836	243	8	01.	5			01.	5	[Desenho] Ewoks S02E02.txt	2022/nov/14 00:	73%
65	:25.03 00:13:30.00 - Vão, tragam-me uma princesa Ewok. - Sim, Rainha Lesmona. 153: 00:13:35.16 00:			Ewok	1.504	198	10	01.	3			01.	3	[Desenho] Ewoks S02E02.txt	2022/nov/14 00:	60%
66	:101. 00:06:58.09 00:07:01.06 uma figura de proa Ewok. 102: 00:07:01.13 00:07:04.22 - Ewok? - Da			Ewok	762	93	14	0	761			0	761	[Desenho] Ewoks S02E09.txt	2022/nov/14 00:	34%
67	:202. 00:13:49.04 00:13:51.12 Levania se pudesse. Ewok. 203: 00:13:51.14 00:13:53.09 Mas,			Ewok	1.709	194	8	01.	8			01.	8	[Desenho] Ewoks S01E05.txt	2022/nov/14 00:	62%
68	:214. 00:14:02.20 00:14:04.15 Você disse qualquer Ewok. 215: 00:14:05.02 00:14:07.01 Tudo bem, se			Ewok	1.692	231	8	01.	1			01.	1	[Desenho] Ewoks S02E07.txt	2022/nov/14 00:	66%
69	:58.02 00:13:01.23 Apoto que transformo qualquer Ewok em um guerreiro em um dia. 197: 00:13:02.13			Ewok	1.555	211	9	01.	4			01.	4	[Desenho] Ewoks S02E07.txt	2022/nov/14 00:	61%
70	: 00:10:54.15 00:10:56.09 - Sabão Ewok! - Sabão Ewok! 150: 00:10:56.19 00:10:59.10 Vamos, rápido.			Ewok	1.218	145	4	01.	7			01.	7	[Desenho] Ewoks S01E02.txt	2022/nov/14 00:	53%
71	: Pequeni! 252: 00:19:16.06 00:19:18.05 O sabão, Ewok! 253: 00:19:20.17 00:19:22.17 Entregue,			Ewok	2.046	258	7	02.	5			02.	5	[Desenho] Ewoks S01E02.txt	2022/nov/14 00:	89%
72	se limpando. 95: 00:06:47.02 00:06:48.09 Sabão Ewok. 96: 00:06:48.22 00:06:50.19 Não há nada			Ewok	800	86	6	0	799			0	799	[Desenho] Ewoks S01E02.txt	2022/nov/14 00:	35%
73	temos aqui? 149: 00:10:54.15 00:10:56.09 - Sabão Ewok! - Sabão Ewok! 150: 00:10:56.19 00:10:59.10			Ewok	1.215	144	7	01.	4			01.	4	[Desenho] Ewoks S01E02.txt	2022/nov/14 00:	53%
74	:51.28.08 O cristal pertence a Kainik, a sacerdotisa Ewok. 361: 00:51:31.18 00:51:34.12 Mas antes que			Ewok	2.805	462	12	02.	4			02.	4	[Filme] Caravana da Coragem.txt	2022/nov/14 00:	65%
75	:59. 00:05:12.05 00:05:17.11 Um monstro selvagem Ewok com seis braços 60: 00:05:17.18 00:05:21.02			Ewok	481	51	8	0	480			0	480	[Desenho] Ewoks S01E13.txt	2022/nov/14 00:	21%
76	: fto. 45: 00:03:25.07 00:03:25.17 Eu também, Ewok. 46: 00:03:31.02 00:03:34.00 Mas alguns			Ewok	370	33	7	0	369			0	369	[Desenho] Ewoks S01E08.txt	2022/nov/14 00:	17%
77	: seja. Arasad som! - Eu poderia colocá-lo na terapia Ewok, em vez disso. Algumas das criaturas nativas			Ewok	99.3833	2	8	03.	2			03.	2	[Livro] Aftermath - Divida de Honra.txt	2022/out/14 00:	150%
78	:35.09 00:16:38.07 E melhor fazer um bom trabalho. Ewok. 247: 00:16:38.09 00:16:40.01 Senão. 248: 00:			Ewok	2.049	240	11	02.	8			02.	8	[Desenho] Ewoks S01E05.txt	2022/nov/14 00:	75%
79	se reúne para uma cerimônia tradicional Ewok. 301: 00:42:19.05 00:42:22.22 Antes de			Ewok	2.330	398	13	02.	9			02.	9	[Filme] Caravana da Coragem.txt	2022/nov/14 00:	54%
80	: 00:12:06.21 00:12:09.06 E melhor olhar para trás, Ewok. 191: 00:12:17.16 00:12:20.16 Kvarik. Corram.			Ewok	1.591	211	10	01.	0			01.	0	[Desenho] Ewoks S01E07.txt	2022/nov/14 00:	57%
81	:27. 00:01:59.21 00:02:04.03 Saudações da tribo Ewok. Rainha das Wisties. 28: 00:02:16.15 00:02:			Ewok	244	13	8	0	243			0	243	[Desenho] Ewoks S01E01.txt	2022/nov/14 00:	11%
82	: 00:18:18.16 00:18:23.01 Lutaremos até o último Ewok para proteger nossas árvores de alma. 275: 00:			Ewok	2.177	262	9	02.	6			02.	6	[Desenho] Ewoks S01E10.txt	2022/nov/14 00:	84%

Fonte: captura de tela do autor.

Para este termo, cinco contextos foram selecionados, cada um formando um conceito diferente. As fontes variam, ou seja, para Ewok, os excertos são provenientes do filme Caravana da Coragem - Uma Aventura Ewok, de 1984, e da série animada *Ewoks*, de 1985. A partir dos contextos escolhidos, foi possível descobrir que Ewoks são uma tribo de amigos vivendo em árvores nas florestas de Endor, com cerimônias próprias, guerreiros e magos.

Na figura abaixo, é possível analisar os contextos escolhidos para criação de definição.

Figura 10 - Contextos cadastrados no VoTec.

Exemplo	Conceito	Fonte	Ações
"Somos Ewoks, somos peludos, vivendo nas árvores altas, vivendo na espiral. Dançando na floresta na lua de Endor. Todos os Ewoks juntos e se divertindo. Amigos sempre juntos, amigos para sempre. Ewoks. Somos incautos, pequenos Ewoks. Gostamos de aventura. Ajudamos amigos em perigo na floresta. Compartilhando a magia na lua de Endor. Todos os Ewoks juntos e se divertindo Amigos sempre juntos. Amigos para sempre. Ewoks. Somos os Ewoks."	Amigos vivendo em árvores em uma floresta de Endor.	Ewoks 21/04/2023	<a href="#">editar</a> - <a href="#">excluir</a>
Os aventureiros se reúnem para uma cerimônia tradicional Ewok.	Cerimônia	Caravana da Coragem 21/04/2023	<a href="#">editar</a> - <a href="#">excluir</a>
"Eu diria que ele é o melhor e mais corajoso guerreiro Ewok de Endor."	Guerreiro	Ewoks 21/04/2023	<a href="#">editar</a> - <a href="#">excluir</a>
"Latara, nenhum mago Ewok consegue fazer isso. Não sem o Manto de Cristal."	Magos	Ewoks 21/04/2023	<a href="#">editar</a> - <a href="#">excluir</a>
"É a Rainha Izrina!" "Saudações da tribo Ewok, Rainha das Wisties."	Tribo	Ewoks 21/04/2023	<a href="#">editar</a> - <a href="#">excluir</a>

Contextos Cadastrados: 5

Fonte: captura de tela do autor.

A utilização do *Concord* não é apenas útil para a definição de termos, mas também, entre diversos usos, para explorar aspectos de tradução – um dos temas deste trabalho. Voltemos ao exemplo do *Star Destroyer*. 385 diferentes contextos de *Star Destroyer* foram encontrados no *corpus* em inglês, contra apenas 21 no *corpus* em português. A discrepância claramente demonstra que o termo foi traduzido, cabendo ao pesquisador o exercício de verificar um dos contextos do termo em inglês e localizar o mesmo trecho na versão nacional, de modo a encontrar o termo nacionalizado. Com isso, foi possível descobrir que o termo, historicamente, foi mais frequentemente traduzido como “Destróier Estelar”. Fora isso, também foi possível localizar 21 contextos em que o termo não foi traduzido em materiais brasileiros, o que enfatiza a necessidade de um guia terminológico como o VoTec.

Após todas estas etapas, os termos estavam prontos para cadastro no VoTec. No próximo capítulo, analisaremos como essa etapa foi realizada.

## 5 CRIAÇÃO DOS VERBETES

Como previamente explicitado, este trabalho tem como objetivo geral a criação de uma proposta de vocabulário de *Star Wars* na plataforma VoTec. O Vocabulário Técnico Online pode ser acessado gratuitamente a partir de qualquer computador ou dispositivo móvel, o que torna a plataforma um ambiente simplificado de consulta para tradutores.

### 5.1 Criação do projeto no VoTec

O projeto foi criado dentro do campo de Terminologia em Ficção, onde, atualmente, já se encontram estudos sobre O Senhor dos Anéis, *Doctor Who* e Jornada nas Estrelas, por exemplo. O projeto deste trabalho, no entanto, foi o primeiro a contar com subáreas, o que refina ainda mais o processo de pesquisa. Isso significa que, ao buscar pelos termos, o usuário pode navegar pelas categorias vistas no Quadro 5.

### 5.2 Elaboração dos verbetes

O pesquisador usuário do VoTec tem acesso à uma página em que é possível fazer o cadastro de novos termos, alimentando a plataforma. Na página inicial, como possível analisar na figura abaixo, o pesquisador tem acesso à listagem de todos os termos cadastrados ainda não aprovados ou aguardando revisão, além das possibilidades de criar um termo e cadastrar uma nova fonte.

Figura 11 - Interface do VoTec.



Trocar Senha   Novo Termo   Cadastro de Fontes   Sair

**Fichas ainda não aprovadas**

Termo	Língua	Coletado em	Cadastrada por	Ações
Mandalore		11/05/2023	<a href="#">Guilherme Rodrigues Ferreira</a>	<a href="#">editar</a>
Mandalore		11/05/2023	<a href="#">Guilherme Rodrigues Ferreira</a>	<a href="#">editar</a>
Wookiee		11/05/2023	<a href="#">Guilherme Rodrigues Ferreira</a>	<a href="#">editar</a>
Wookiee		11/05/2023	<a href="#">Guilherme Rodrigues Ferreira</a>	<a href="#">editar</a>
Ewoks		21/04/2023	<a href="#">Guilherme Rodrigues Ferreira</a>	<a href="#">editar</a>
Ewoks		13/04/2023	<a href="#">Guilherme Rodrigues Ferreira</a>	<a href="#">editar</a>
Destróier Estelar		13/04/2023	<a href="#">Guilherme Rodrigues Ferreira</a>	<a href="#">editar</a>
Star Destroyer		12/04/2023	<a href="#">Guilherme Rodrigues Ferreira</a>	<a href="#">editar</a>
Caça Estelar		11/04/2023	<a href="#">Guilherme Rodrigues Ferreira</a>	<a href="#">editar</a>
Starfighter		11/04/2023	<a href="#">Guilherme Rodrigues Ferreira</a>	<a href="#">editar</a>

Fonte: captura de tela do autor.

O primeiro passo para criar um termo é, evidentemente, nomeá-lo. Algumas palavras apresentam as possibilidades de singular ou plural, como é caso de *Ewoks*, na figura acima, que

não foi cadastrado como *Ewok*, simplesmente. Para este trabalho, consideramos para cadastro no VoTec a forma mais frequente no *corpus* – o que significa que, em todo o universo canônico de Star Wars, o termo *Ewoks* se mostra mais frequente que *Ewok*.

Apesar disso, ambas as formas (no singular e plural) foram posteriormente lematizadas, ou seja, agrupadas no *corpus* como um único termo, explicando a possibilidade de se encontrar, por exemplo, exemplificações contextuais no singular de um termo que foi cadastrado no plural.

Após nomear o termo, é preciso escolher o idioma, a grande área, a subárea 1 e a subárea 2. A primeira diz respeito ao material a ser trabalhado, ou seja, no nosso caso, *Star Wars*. A segunda consiste nas categorias que foram cadastradas. A figura abaixo mostra, a partir do cadastro do termo *Wookiee*, como é realizado o processo.

Figura 12- Processo de cadastro de um novo termo.

Vocabulário Técnico Online Tela Cheia | English

## Novo Termo

[Voltar ao painel](#)

**Passo 1**

Wookiee

**Língua**  
Escolha uma língua: Inglês

**Ontologia**  
Grande Área: Terminologia em Ficção  
Sub-área 1: Star Wars  
Sub-área 2: Escolha uma Sub-área

- Escolha uma Sub-área (s)
- Armas
- Droids
- Espécies
- Lugares
- Organizações
- Profissões
- Religião
- Tecnologia
- Veículos

31/05/2023 05:46 © 2007 FFLCH - ICMC Jr.

Fonte: captura de tela do autor.

O segundo passo é cadastrar os contextos. A definição do termo será criada a partir unicamente de fragmentos contextuais reais. Sendo assim, como já explicado anteriormente, é com o auxílio da ferramenta *Concord* que o pesquisador consegue buscar um termo específico no *corpus* e analisar todas as situações em que ele foi utilizado. O objetivo é buscar os mais variados conceitos.



A tela de cadastro de contextos no VoTec pode ser vista na figura abaixo.

Figura 13 - Cadastrando contextos.

### Novo Contexto

#### Passo 2

Dados do Contexto

Exemplo\*:

Conceito\*:

Fonte\*:  [Cadastrar](#)

[Nova Fonte](#)

Data de Coleta\*:  (dia/mês/ano ex.: 18/03/2007)

#### Contextos Cadastrados

Exemplo	Conceito	Fonte	Ações
"Let him have it. It's not wise to upset a Wookiee."	Dangerous	ANewHope 10/05/2023	<a href="#">editar</a> - <a href="#">excluir</a>
General Hux led the way. They passed officers and stormtroopers, droids and maintenance crew, and although a giant, hairy Wookiee occasionally made someone do a double take, Hux's presence gave them unhindered passage through the ship's corridors.	Giant and hairy	RiseSkywalkerBook 10/05/2023	<a href="#">editar</a> - <a href="#">excluir</a>
"Thanks for not breaking my neck." The Wookiee replied with a guttural, modulated rumble. Finn chose to interpret it as an apology of sorts.	Guttural voice	TheForceAwakens 10/05/2023	<a href="#">editar</a> - <a href="#">excluir</a>
"Even a wookiee can't crush First Order armor."	Strong	TheForceAwakens 10/05/2023	<a href="#">editar</a> - <a href="#">excluir</a>

Contextos Cadastrados: 4

Fonte: captura de tela do autor.

Assim que for selecionada no *corpus* uma passagem que explique o termo de forma satisfatória, esse trecho é inserido no campo "Exemplo". O segundo contexto cadastrado na Figura 13, por exemplo, serve para mostrar as características físicas de um *Wookiee* – este é um **Conceito**, quanto mais tivermos, mais rica será a definição.

Fora isso, também são inseridas a fonte e a data da coleta. As fontes precisam ser cadastradas, como exemplificado na figura abaixo.

Figura 14 - Cadastrando uma fonte de contexto.

## Nova Fonte

Informações sobre a fonte

Nome\*:

Abreviação\*:

Link:

Fonte: captura de tela do autor.

O próximo passo, exemplificado na Figura 15, consiste em detalhar o termo e elaborar conceito e definição. Inserimos a categoria gramatical, o número (singular ou plural), o gênero, a posição na ordem de frequência e o número de ocorrências do termo. Também é inserido o termo equivalente em outro idioma, caso ele já tenha sido cadastrado no VoTec, além de informações enciclopédicas.

Figura 15 - Detalhando o termo.

**Contextos** **Conceito Final / Definição**

Contextos

Exemplo	Conceito	Fonte
1 "Let him have it. It's not wise to upset a Wookiee."	Dangerous	ANewHope 10/05/2023
2 "Thanks for not breaking my neck." The Wookiee replied with a guttural, modulated rumble. Finn chose to interpret it as an apology of sorts.	Guttural voice	TheForceAwakens 10/05/2023
3 "Even a wookiee can't crush First Order armor."	Strong	TheForceAwakens 10/05/2023
4 General Hux led the way. They passed officers and stormtroopers, droids and maintenance	Giant and hairy	RiseSkywalkerBook 10/05/2023

**Dados** **Traços Distintivos** **Semântica** **Termo Equivalente** **Termos Remissivos** **Informações Enciclopédicas**

Dados

Ontologia: Terminologia em Ficção > Star Wars > Espécies

Categoria Gramatical:  Número:

Gênero:  Sigla/Acrônimo:

Entrada por Extenso:

Var. Morfossintáticas:

Acepção Nº:

Córpus

Posição na Ordem de Frequência:  Nº de Ocorrências do termo:

Fonte: captura de tela do autor.

A partir dos excertos selecionados, elaboramos um conceito final que explique, de fato, o que é o termo. A definição, por sua vez, traz em uma única sentença uma explicação polida e sucinta desenvolvida a partir do conceito final. A diferença entre o conceito final e a definição pode ser analisada na figura abaixo, ainda no mesmo exemplo.

Figura 16 - Criando conceito final e definição.

Contextos	Conceito Final / Definição
Conceito Final / Definição	
Conceito Final:	Giant, hairy creature that is capable of physically overpowering and potentially harming others if provoked or upset. Their communication is primarily through guttural, modulated rumbles, and they have a reputation for being strong and difficult to defeat in combat.
Definição:	Giant, hairy and strong creature that communicates through guttural rumbles, known for their ability to physically overpower others and their reputation as skilled combatants.

Fonte: captura de tela do autor.

Feito todo este processo, o termo está pronto para ser enviado ao orientador, que por sua vez deve avaliar, analisar e aprovar (ou não) o trabalho. Para este projeto, este processo foi realizado 36 vezes, já que para cada uma das nove categorias, dois termos foram selecionados e cadastrados duas vezes (em português e inglês). Abaixo, na Figura 17, é possível conhecer a interface do VoTec para o usuário final.

Figura 17 - Interface do VoTec para o usuário final.

Terminologia em Ficção	Star Wars	Espécies
Wookiee	Buscar	
<b>Tipos de Exibição</b>	▼ Português	
Normal	<a href="#">Voltar ao resultado da busca</a>	
Descritiva	Wookiee. <i>Espécies</i> . s.m.s. Criatura gigante, peluda e forte que se comunica por meio de estrondos guturais, conhecida por sua capacidade de dominar fisicamente os outros e sua reputação como combatentes habilidosos. Ex.: "Não é bom contrariar um Wookiee.". <i>Corpus: Posição na Ordem de Frequência</i> : (1608); <i>Nº de Ocorrências do termo</i> : (286). <i>Informações Enciclopédicas</i> : Os Wookiees eram uma espécie alta e peluda que eram nativos do planeta Kashyyyk Em: <a href="#">Wookiee - Star Wars Wiki em Português</a>	
<b>Tipos de Consulta</b>	▼ English	
Total	<a href="#">Go back to search results</a>	
Tradutor	Wookiee. <i>Species</i> . n.m/f.s. Giant, hairy and strong creature that communicates through guttural rumbles, known for their ability to physically overpower others and their reputation as skilled combatants. Ex.: "Let him have it. It's not wise to upset a Wookiee.". <i>Corpus: Frequency order position</i> : (1478); <i>Term number of occurrences</i> : (204). <i>Encyclopedic Information</i> : Wookiees were a species of tall, hairy humanoids that were native to the planet Kashyyyk. em: <a href="#">Wookiee - Wookieepedia</a>	
Modular		
<b>Consultas Externas</b>		
Corpus NILC		
Google		
Answers.com		
Wikipedia		
CORTEC		

01/06/2023 13:12 © 2007 Guilherme Fromm - ICMC Jr.  
Termo elaborado por [Guilherme Rodrigues Ferreira](#) (pt)

Fonte: captura de tela do autor.

O resultado pode ser encontrado, enfim, na página do VoTec (<http://ic.votec.ileel.ufu.br/>), onde todos os termos aprovados são disponibilizados para consulta pela comunidade. Ao usuário final, como pode ser visto na Figura 17, cabe inserir as áreas que gostaria de explorar, buscar por um termo e analisar o resultado nos idiomas fornecidos.

## 6 CONSIDERAÇÕES FINAIS

O estudo da Linguística de *Corpus* tem proporcionado, a partir da análise de *corpora*, uma mais facilitada análise de padrões linguísticos, podendo resultar no desenvolvimento de ferramentas e aplicações ou ainda no estudo da sociolinguística. A relação entre a Linguística de *Corpus* e a Tradução é antiga, como explicitado por Kruger, Wallmach e Munday (2011), uma vez que a possibilidade de contrastar *corpora* em diferentes idiomas se mostra bastante e cada vez mais atrativa para pesquisadores. Partindo deste princípio, fizemos o levantamento de dois *corpora* do universo canônico de *Star Wars*, em português e inglês, com o intuito de auxiliar tradutores e estudantes de tradução. Consideramos também o impacto e a relevância cultural da franquia, que não apenas tem seus elementos traduzidos nos diversos lançamentos todos os anos, mas também em todas as outras mídias que de alguma forma referenciam e vão referenciar esse material, como os episódios de ou sobre *Star Wars* em Uma Família da Pesada (*Family Guy*), Os Simpsons (*The Simpsons*) e South Park (*South Park*).

Como um protótipo de um vocabulário bilíngüe inglês/português de *Star Wars* baseado em *corpus*, os objetivos da pesquisa foram cumpridos uma vez que a proposta está disponível no VoTec de forma bastante detalhada, possibilitando a consulta de tradutores que estão, de alguma maneira, lidando com o universo *Star Wars* em seus trabalhos e pesquisas. Este trabalho, no entanto, pode ser continuado, com o prosseguimento da análise do *corpus* e seus cadastros no VoTec. Para trabalhar com toda a terminologia de *Star Wars*, seria necessária uma maior disponibilidade de tempo para cadastrar todos os termos e atualizar os *corpora*, uma vez que a franquia é ativa e novos conteúdos continuam sendo lançados. Este trabalho também desconsiderou mídias canônicas de difícil transposição para *corpus*, como quadrinhos e jogos eletrônicos, que seriam pertinentes em um trabalho mais aprofundado com este mesmo material.

No fim, conclui-se que a franquia *Star Wars* tem um universo consideravelmente rico para exploração em *corpus*, podendo ser feitas, inclusive, análises no campo da literatura, da tecnologia e até mesmo da geografia que vem sendo criada desde o final dos anos 70. Neste trabalho, seguimos um caminho mais técnico, focando na construção de um protótipo de vocabulário com o intuito de auxiliar as atividades profissionais e amadoras de tradutores que podem se beneficiar do VoTec.

## REFERÊNCIAS

BARBOSA, Maria Aparecida. Para uma etno-terminologia: recortes epistemológicos. **Ciência & Cultura**, v. 58, n. abr./maio/ju 2006, p. 48-51, 2006.

Cartaz de Cinema: "STAR WARS - O IMPÉRIO CONTRA ATACA" - "APERTEM OS CINTOS ... O PILOTO SUMIU!" - "FAMA" - "MEU AMIGO, O DRAGÃO". Anos Dourados: Imagens & Fatos, 2022. Disponível em: <http://www.anosdourados.blog.br/2022/05/decada-de-80-imagens-cartaz-de-cinema.html>. Acesso em: 02 jun. 2023.

CABRÉ, Maria Teresa. **La Terminología: representación y comunicación: elementos para una teoría de base comunicativa y otros artículos**. Barcelona: Institut Universitari de Lingüística Aplicada, 1999.

DIAS, Cláudia Augusto. Terminologia: conceitos e aplicações. **Ciência da Informação**, Brasília, v. 29, n. 1, p. 90-92, abr. 2000. IBICT. DOI <http://dx.doi.org/10.1590/s0100-19652000000100009>.

FINATTO, Maria José Bocorny. Orientações para a terminografia: das teorias às práticas em busca de amplitude da informação terminológica. In: ISQUERDO, A. N.; CORNO, G. O. M. dal. (org.) **As Ciências do Léxico: lexicologia, lexicografia e terminologia**, Campo Grande. MS: Ed.UFMS, vol. VII, 2014, p. 439-459.

FROMM, Guilherme. Ficção, Tradução, Terminografia e Linguística de Corpus: confluências. In: XIII Simpósio Nacional de Letras e Linguística e III Simpósio Internacional de Letras e Linguística, 2011, Uberlândia. **Anais do SILEL**. Volume 2, Número 2. Uberlândia: EDUFU, 2011. Uberlândia: EDUFU, 2011. v. 2.

FROMM, Guilherme. **VoTec: a construção de vocabulários técnicos eletrônicos para aprendizes de tradução**. 2007. 215f. Tese (Doutorado em Estudos Linguísticos e Literários em Inglês). Departamento de Letras Modernas, Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2007.

KRUGER, Alet; WALLMACH, Kim; MUNDAY, Jeremy. **Corpus-Based Translation Studies: research and applications**. Londres: Continuum, 2011.

LARA, Marilda Lopes Ginez de. Linguagem documentária e terminologia. **Transinformação**, Campinas, v. 16, n. 3, p. 231-240, dez. 2004. FapUNIFESP (SciELO). DOI <http://dx.doi.org/10.1590/s0103-37862004000300003>.

LARA, Marilda Lopes Ginez de. **Elementos da Terminologia**: (apostila para uso didático). São Paulo, 2005.

LATORRE, Vanice Ribeiro Dias. A dialética entre os extremos: da terminologia à etnoterminologia. **Caderno Seminal**, [S.L.], v. 19, n. 19, p. 70-94, 29 ago. 2013. Universidade de Estado do Rio de Janeiro. DOI <http://dx.doi.org/10.12957/cadsem.2013.12062>.

RIBEIRO, Gabriela Castelo Branco. Tradução técnica, terminologia e lingüística de corpus: ferramenta WordSmith Tools. **Cadernos de Tradução**, Rio de Janeiro, v. 2, n. 14, p. 160-174, jan. 2004.

SALES, Rodrigo de. **Teoria comunicativa da terminologia (TCT) como aporte teórico para a representação do conhecimento especializado**. Apresentação de pôster no VIII ENANCIB – Encontro Nacional de Pesquisa em Ciência da Informação, Salvador, out. 2007.

SARDINHA, Tony Berber. **Pesquisa em Lingüística de Corpus com WordSmith Tools**. Campinas: Mercado de Letras, 2009.

**STAR Wars I – A Ameaça Fantasma**. VHS Mania, 2016. Disponível em: <https://vhs-mania.blogspot.com/2016/03/star-wars-i-ameaca-fantasma.html>. Acesso em: 02 jun. 2023.

**Star Wars: Episódio IV – Uma Nova Esperança** (Comentários em áudio de George Lucas, Carrie Fisher, Ben Burt e Dennis Muren). Direção de George Lucas. Estados Unidos: Twentieth Century-Fox Film Corporation, 2011. Blu-ray (121 min).

ZAMORA, Rita Elena Melian. Sobre tradução e terminologia das ciências sociais e humanas: quando a cultura encontra a “cultura”. **Mutatis Mutandis**: Revista Latinoamericana de Traducción, [S. L.], v. 8, n. 2, p. 548-566, 2015.