

ABSTRACTS E RESUMOS: UMA ANÁLISE COMPARATIVA PELA LINGÜÍSTICA DE CORPUS

Angelo Augusto Peixoto Coutinho^{1*}

Resumo: O objetivo deste trabalho é mostrar o resultado da análise de um *corpus* de 100 resumos e seus respectivos *abstracts* de dissertações e teses em Linguística da UFMG no período que correspondente de 2012 a 2015. Teoricamente, respaldei-me nos pressupostos teóricos de Baker (1993), destacando a proposta de que a língua traduzida está sujeita aos chamados universais da tradução. Para tanto, foi montado um *corpus* comparável permitindo o estudo das linhas de concordância. Em conclusão, aliado a ferramentas estatísticas e às comparações com o *COCA* foi possível confirmar que os universais estão presentes nos resumos traduzidos em seus *abstracts*.

Palavras-chave: corpora; abstracts; tradução.

Abstract: The purpose of this work is to show the results of the analysis of a *corpus* of 100 abstracts written in Portuguese and their English translation. The texts were retrieved from UFMG's Linguistics graduate program, in the period from 2012 to 2015. The paper was endorsed with the theoretical assumptions by Baker (1993), highlighting the issues of translation's universals. To study the texts, a comparable *corpus* was compiled, so that the concordance lines could be analyzed. To conclude, using the statistical tools and comparisons with *COCA*, it was possible to confirm the translation universals are present in the translated abstracts.

Key-words: corpora; abstracts; translation.

1 ^{*} Artigo apresentado à disciplina Introdução à Linguística de Corpus, 2º semestre de 2015. Universidade Federal de Minas Gerais. Orientadora: Heliana Mello.

1. Introdução

Textos traduzidos têm um papel essencial em nosso cotidiano. Quando analisamos nosso entendimento de outras culturas, arte e muitas outras áreas do conhecimento, vemos que uma grande parte do que consumimos está intrinsecamente ligado à tradução. No ambiente acadêmico não é diferente: sempre que se escreve uma dissertação, tese de mestrado ou doutorado, deve-se introduzir o trabalho com resumo e sua respectiva tradução, o *abstract*. Tendo isso em vista, a análise do texto traduzido se faz cada vez mais importante para poder buscar uma língua traduzida mais perto da língua que é realmente falada. Muitos teóricos da tradução afirmam que o que se conhece como linguagem da tradução poderia ser definido como “traducionês”, haja vista que muitas vezes as traduções se afastam da língua-alvo sendo consideradas artificiais demais. No geral, muitos dos textos que lemos são mediados pela tradução (BAKER, 1993). Com o desenvolvimento e popularização da área da Linguística Aplicada e mais especificamente da área da Linguística de Corpus nos anos 1980, a atividade tradutória pode, agora, usar das ferramentas presentes nos *corpora*. O crescente interesse por estudos empíricos junto ao desenvolvimento da informática e suas tecnologias resultou na ascensão da Linguística de Corpus. Como consequência, a presença de *corpus* eletrônico e acessível junto com as ferramentas de processamento presentes nele, permitem ao tradutor acessar informações sobre diversos aspectos da língua.

Portanto, o desenvolvimento de *corpora* foi de suma importância para a atividade tradutória, graças à possibilidade de análise da tradução junto com as ferramentas presentes nesses *corpora*. Esse riquíssimo leque de ferramentas nos permite buscar por padrões e aproximar mais a língua traduzida da língua natural.

O objetivo geral desse trabalho é estudar, sob a ótica da Linguística de Corpus e da Teoria da Tradução, resumos de dissertações, de mestrado e teses de doutorado e suas respectivas traduções para o inglês – os *abstracts*, elaborados por alunos do Programa de Pós-Graduação em Estudos Linguísticos da UFMG (doravante Pos-Lin) (mestrado e doutorado) entre os anos 2012 e 2015. Para tanto, foram selecionados 100 textos aleatoriamente e com eles foi elaborado um *corpus* comparado para entender as áreas mais problemáticas da tradução. Este *corpus* montado será referido como *minicorpus*.

Especificamente, este trabalho pretende analisar os usos do pronome *it* em inglês quando vem no início de sentença. Ademais, estudaremos o uso dos substantivos *objetivo*, *propósito* e *alvo* contrastando com suas respectivas traduções *objective*, *purpose* e *aim*. Além disso, pretendemos compreender o uso dos advérbios em português e suas traduções, mais particularmente a classe de advérbios terminados em *-mente* e sua possível correspondência tradutória para advérbios do inglês terminados em *-ly*. Propõe-se também o estudo da complexidade textual em relação a densidade de lexemas e tamanho das sentenças. E por último, problematizar os usos de determinantes: *other* e *another*.

2. Metodologia

Para realizar o estudo, compilou-se um *corpus* de 100 resumos e *abstracts* de teses e dissertações Pos-Lin. Para tanto, foram selecionados manualmente e aleatoriamente os textos diretamente do *site* do programa, onde estão disponibilizados. Logo após a seleção, montou-se um *corpus* comparado dos resumos e dos *abstracts* com o programa *AntConc*, que também está disponível na internet, gratuito para download. Desenvolvido pela Universidade de Waseda, o *AntConc* permite explorar di-

versas ferramentas que auxiliam na compilação de um *corpus*, análise das linhas de concordância e colocados, por exemplo.

Após montado o *corpus* comparado, foi analisada a frequência das palavras e os colocados. Usamos também o programa *UAM CorpusTool*, desenvolvido na Universidade de Madri, que permite retirar dados estatísticos dos grupos de textos e obter complexidade lexical, densidade de lexemas e tamanho das sentenças. A ferramenta, apesar de ser focada na anotação, foi usada nesse artigo puramente para estatística.

Além disso, exploramos a ferramenta acadêmica do corpus compilado pela Universidade Young Birmingham, o *COCA (Corpus of Contemporary American English)*. Portanto, uma análise quantitativa e qualitativa foi feita dos textos a fim de buscar padrões linguísticos que revelassem normas e convenções na tradução de resumos do português para inglês. A seguir será feita uma análise e discussão dos dados apresentados.

3. Referencial teórico

Nesse trabalho nos embasaremos no fato de que a tradução é um evento comunicativo mediado (BAKER, 1993). Como consequência, pesquisadores começaram a considerá-la um terceiro código, levando-se em conta que o texto traduzido traz uma série de desvios das normas, convenções tradutórias, influência da língua fonte e dos universais da tradução. Este último é um conceito introduzido por Baker (1993) depois de ter estudado alguns padrões da língua traduzida. Apesar de alguns cientistas acreditarem que esses universais sejam hipóteses, Baker (1993) claramente assume que eles são as características dos textos traduzidos às quais se deve dar mais atenção quando estudados através de *corpora*. Essas características incluem: tendência a

simplificação, aumento do nível de explicitação, preferência por normalização, tendência a gravitar no centro de um contínuo (nívelamento).

Primeiramente, a tendência pela simplificação indica que existe uma inclinação do tradutor por simplificar a língua traduzida. Isso se manifesta em uma divisão de sentenças longas por curtas, evitar repetições, desambiguação, deixar o uso de pontuação mais claro e diminuir a densidade lexical, ou seja, diminuir a frequência de palavras lexicais e aumentar a de gramaticais (BAKER, 1993).

Ainda dentro da simplificação, uma das maneiras de observar esse universal que merece destaque é a explicitação. No tocante ao aumento do nível de explicitação, isso implica em não deixar nada implícito. “Parece que a explicitação é usada não somente para tampar espaços vazios no conhecimento em caso de textos incongruentes, como também, para aumentar a compreensão já que o código explícito é mais fácil de processar do que informação implícita.” (BIEL, 2010, p. 9). Muito comumente, nota-se a explicitação principalmente no aumento de estruturas coesivas e da adição e especificação de itens lexicais e gramaticais na língua-alvo. Em nosso *corpus*, percebemos que houve um aumento geral de 30% nos advérbios dos *abstracts* quando comparados com os resumos. Junto a isso, se soma uma queda de quase 40% dos lexemas por sentença nos *abstracts*, mas sem alteração do tamanho do texto, ou seja, houve um aumento de palavras gramaticais nos *abstracts* em relação aos seus respectivos resumos.

A preferência por normatização consiste em normatizar estruturas não-gramaticais. Na realidade, trata-se de uma tendência a conformar a tradução às regras típicas da língua-alvo, às vezes ao ponto do exagero. Em textos traduzidos, existe uma prefe-

rência pelo “normal” ao invés de explorar estruturas criativas e inovadoras que também seriam possíveis na língua-alvo. Além disso, os tradutores também evitam metáforas, neutralizando-as na língua-alvo. Com relação a esse universal, observamos o contrário do esperado, quando analisados os inícios de sentença com *it*: houve um uso de “*it were* + verbo no particípio passado” muito pouco comum no inglês. Na realidade, percebe-se uma tendência dos tradutores a iniciar muitas sentenças com *it* e separando o sujeito em duas partes; esse tipo de estrutura, apesar de correta, deixa o texto pouco fluido e mais complicado de processar, afastando-o da língua-alvo. Percebemos também que os usos de *other* e *another* se apresentam um pouco confusos para os tradutores. Apesar de fazermos uma análise qualitativa dos poucos casos em que o uso fugiu ao prescrito pela gramática da língua inglesa, essas ocorrências já demonstram que o universal da normatização nem sempre é válido.

Como mencionado anteriormente, todo esse grupo de universais da tradução está conectado com uma característica em comum: a nivelção. Essa característica se relaciona a tendência dos textos traduzidos de gravitar no centro de um contínuo. Mantendo-se em curso mediano entre dois extremos, convergindo para o centro (BAKER, 1993). Em outras palavras, textos traduzidos tendem a parecer mais entre eles do que os originais. As características da normatização estão presentes no *minicorpus*, já que percebemos que os textos dos *abstracts* contêm um excesso de sentenças iniciadas com *it*, e há uma frequência muito alta de *purpose*, *objective* e *aim*. A repetição frequente dessas palavras indica que os textos traduzidos se parecem muito mais entre si quando comparados com a maior variedade lexical presente nos resumos.

Considerando essa problemática dos universais da tradução,

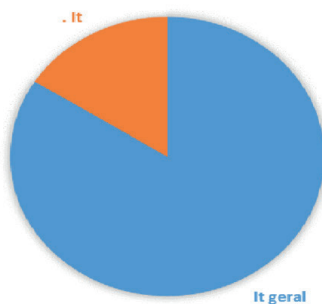
um *corpus* comparável vai nos permitir analisar como esses universais se manifestam, já que com o uso de ferramentas quantitativas estatísticas podemos entender, por exemplo, a frequência, pontuação, *tokens* e comprimento de sentença.

4. Discussão

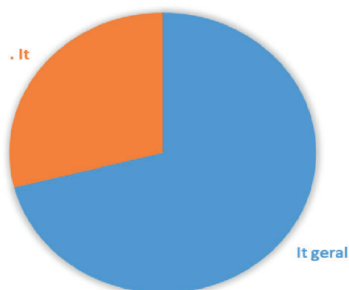
Nesta seção serão contempladas algumas considerações feitas em relação à análise dos textos dos *corpora*. Vale salientar que a pesquisa busca por padrões, convenções, normas e preferências tradutórias.

Notamos que existe um padrão muito comum de tradução de iniciar sentenças com *it* devido a um arquétipo extremamente habitual de início de sentença no português: *É....* Existem diversos problemas que surgem a partir de um texto que tem excesso de sentenças iniciando com *it*. Em uma frase como “*It is necessary to plan concrete objectives for the writing teaching*” nota-se que tanto *it* quanto “*to plan concrete objectives for the writing teaching*” podem ser sujeitos da oração. No caso, o *it* é o sujeito gramatical da oração, mas o sujeito retórico não está presente no *it*. Essa sentença foi retirada de um dos *abstracts* e demonstra como torna-se confuso um texto que explora demais essa construção. Outro problema que surge é que esse padrão de sentença está perfeito em inglês; o problema é semântico e de facilidade de leitura e clareza do texto, com isso tradutores tendem a abusar um pouco desse tipo de estrutura. Observamos que no *COCA*, de todos os usos de *it* em textos acadêmicos em inglês, somente 16% iniciam uma sentença. Já no *minicorpus*, como já esperado, o uso desse tipo de padrão sobe para 25%. Observe os gráficos:

COCA



MINICORPUS



GRAF. 1 e 2 - Porcentagem de *it* no início de sentença em relação ao uso de *it* no geral

Fonte: elaborada pelo próprio autor

O estudo dos colocados de *it* também revelou padrões tradutórios típicos de traduções literais. No *minicorpus* os colocados verbais mais comuns foram os verbos *is* e *was*, o que revela que em relação

ao *COCA*, as ocorrências estão invertidas. Assim, no *COCA*, 37% dos colocados são o verbo *is*, no *minicorpus* a porcentagem cai para 25%; já *was* tem 36% de ocorrências no *minicorpus*, enquanto no *COCA*, ele só aparece 12%. Observe:

Was e were como colocados de *it* no *minicorpus*

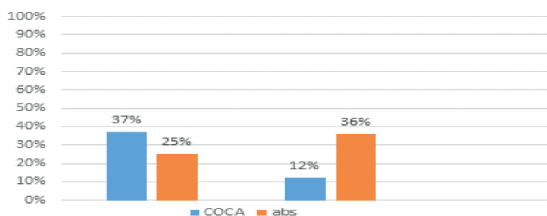


GRÁFICO 3 - *Was* e *were* como colocados de *it* no *corpus*

Fonte: elaborada pelo próprio autor

Houve um uso no *minicorpus* de passiva no plural com *it*: “*It were elaborated tables divided by levels (6^o a 9^o) to quantify texts and genres*”. De acordo com *COCA*, os únicos usos da construção com verbo *to be* no plural acompanhado de particípio passado se aplicam ao subjuntivo. Considerando isso, ao pesquisar pela expressão “*it were *ed_v*” (*it were* + verbo no particípio passado) no *COCA*, tem-se como retorno 121 *hits*, no entanto, vale ressaltar que em todas as sentenças existe um *if* antes da frase, o que caracteriza subjuntivo. A seguir temos a imagem das 20 últimas linhas de concordância da seção acadêmica do *COCA*, com destaque para o total de linhas de concordância:

80	☐	IT WERE CONFINED	1	
81	☐	IT WERE CONCEDED	1	
82	☐	IT WERE COMMUNICATED	1	
83	☐	IT WERE COMMITTED	1	
84	☐	IT WERE COOED	1	
85	☐	IT WERE CLOSED	1	
86	☐	IT WERE CLEARED	1	
87	☐	IT WERE CHRONICLED	1	
88	☐	IT WERE CAUSED	1	
89	☐	IT WERE CALLED	1	
90	☐	IT WERE B LOCKED	1	
91	☐	IT WERE B LACKTOPPED	1	
92	☐	IT WERE ASSIGNED	1	
93	☐	IT WERE APPROVED	1	
94	☐	IT WERE APPORTIONED	1	
95	☐	IT WERE APPLIED	1	
96	☐	IT WERE ADVISED	1	
97	☐	IT WERE ADOPTED	1	
98	☐	IT WERE ACTED	1	
99	☐	IT WERE ACHIEVED	1	
100	☐	IT WERE ACCOMPANIED	1	
		TOTAL	121	

FIGURA 3 - Uso no COCA de “*it were *ed_y*”.

Fonte: elaborada pelo próprio autor

Veja abaixo as primeiras 25 linhas de concordância feitas como AntConc. Notamos que para realizar a pesquisa, a expressão usada é “. *It*” o que não corresponde sempre ao pronome *it* como no caso de *Italy*, então foi feita uma limpeza dos resultados e só o que interessava foi selecionado:

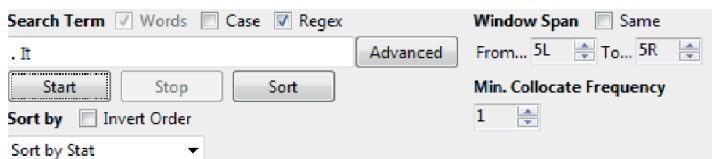


FIGURA 4 - Interface de pesquisa do AntConc

Fonte: elaborada pelo próprio autor

ter *rancid butter* coocorrendo, mas nunca teríamos *rancid bread*, mesmo *rancid* sendo um dos sinônimos para *podre*. Isso significa que se quiséssemos traduzir *pão podre* para inglês, teríamos uma tradução muito mais natural se escolhêssemos outro item lexical diferente de *rancid*, já que ele não coloca com *bread*. Partindo agora para os colocados do *it* inicial, observa-se que os mais comuns no *minicorpus* foram: *analyses, aims, necessary, expected, elaborated, focuses, observed, assumed, noted, prepared, verified*. Desses verbos, nenhum se encontra nem mesmo entre as cinquenta ocorrências mais comuns desse tipo de estrutura no inglês acadêmico, de acordo com o *COCA*. Isso revela que o universal do nivelamento poderia ser comprovado através dessa análise. Os textos dos *abstracts* se parecem muito mais entre si e tem um estilo muito mais próprio do que os originais. Como mencionado acima, os verbos usados pelos tradutores são poucos usados em inglês acadêmico, o que faz com que os textos traduzidos convirjam para o centro desse contínuo de *abstracts* que tem muito mais características em comum do que os resumos. Para mais, ao contrário do que se esperava de acordo com o universal da normatização (BAKER, 1993), o uso de uma construção de passiva pluralizada deixa o texto pouco natural e pouco fluido para o inglês. Mais uma vez devemos lembrar que “*it were* + verbo no particípio passado” corresponde ao subjuntivo e que essa modalização é rara no inglês moderno. Não só em relação ao subjuntivo, mas a alta frequência de *it* no início de sentença demonstra um texto fragmentado e de difícil leitura devido à quebra do sujeito gramatical em duas partes na sentença. A investigação dos colocados também revela um texto pouco natural, dado que observamos que os colocados comuns de *it* nos *abstracts* não correspondem às melhores escolhas no inglês acadêmico. Por isso, vemos aqui que ao contrário do esperado, o texto poderia estar melhor normatizado e aproximar-se mais da língua-alvo.

Partirei agora para o uso das palavras: *aim*, *objective* e *purpose*. Essas palavras aparecem muito comumente na introdução dos trabalhos e vêm como tradução de *propósito* ou *objetivo*. É bom salientar que em nenhum momento os pesquisadores ao escreverem suas pesquisas em português colocaram *alvo* (*aim*) na introdução do trabalho. No *minicorpus* observou-se que *aim* apareceu em 21% dos textos, *objective* em 14% e *purpose* em 17%. Já no COCA o número de *hits* revela que, dos três, *purpose* é o mais comum, seguido de *objective*, sendo *aim* o menos comum. Em relação a tradução, *aim* sempre era tradução de *objetivo* ou *propósito*, levando-se em conta o fato citado acima. Já a palavra *objetivo* foi sempre traduzida literalmente.

Mais uma vez os resultados confirmam os universais da tradução, principalmente o do nivelamento (BAKER, 1993). Os textos dos *abstracts* em questão se parecem muito mais entre si do que os originais. O uso extremamente comum de *aim* só serve para corroborar o fato de que os *abstracts* se caracterizam por ter um estilo próprio de padronizar esse tipo de texto traduzido do português. Como observado anteriormente, Baker (1993) propõe que textos traduzidos tendem muito mais a diminuir na variação lexical e apresentar vocabulário mais parecido quando comparados. A observação do universal da nivelção se faz clara ao vermos a repetição frequente de *objective*, *purpose* e especialmente de *aim* em *abstracts*, produzindo léxico menos variado.

Nossa próxima análise será a de advérbios terminados em *-mente* e suas traduções em *-ly*. É relevante perceber que nos resumos só temos 210 advérbios, já nos *abstracts* esse número sobe para 256. Nos *abstracts*, encontrou-se 30% a mais de advérbios, havendo um caso extremo em que o número mais que dobrou, representando um aumento de 125% no número de advérbios desse *abstract* em relação ao seu respectivo resumo. Houve dois casos em que os ad-

vérbios foram adicionados à tradução e não existiam no resumo. No caso, *detalhadamente* foi traduzido para *in detail* e *não apenas* foi traduzido para *only*. Associo esse caso à problemática de explicitação e da simplificação postulada por Baker (1993), considerando que notamos que com advérbios, os tradutores sempre procuram explicitar ao máximo a ideia da sentença que estavam traduzindo, por isso existe esse número elevado no aumento de advérbios nos textos. Houve casos em que os tradutores optaram por uma expressão com valor adverbial, como *atualmente*, que foi traduzido como *in the present day*, e *somente*, que foi traduzido para *by restricting*. Esses dois últimos casos demonstram claramente o universal da explicitação (BAKER, 1993). Provavelmente, na visão dos tradutores, uma sentença construída com apenas um advérbio não passaria tão bem a ideia que essas duas expressões adverbiais passam. Portanto, a expansão de uma sentença de forma sintática ao invés do uso do morfema adverbial demonstra a tendência ao nível maior de explicitação, já que esse tipo de construção esgota todas as possibilidades interpretativas, não deixando espaço para qualquer erro de processamento ou interpretação. A imagem abaixo ilustra uma comparação das primeiras 25 linhas de concordância dos resumos e dos seus *abstracts*, respectivamente:

Results for both experiments show that only those participants who classified as H regarding to typography and colours (respectively). In addition, it is worthy to discussions on the potential meanings, mainly, produced by images and through which readers are projected by the textbook. Finally, the comparative analysis showed that. These characteristics could be inferred, mostly, by the concepts of text each basic assumptions: a prosodic word gets only one primary stress; the relationship between constituents is not isomorphic. Morphologically we analyze the notion of phonologically we analyze the notion of freely and factoring are criteria for the that identify them as prosodic words. Prosodically, we look as the primary stress marked words and phonological rules that apply to them, in order to identify aims to investigate a singular and historically situated reality, concerning the. The nature of this study is, essentially, qualitative, centered on descriptive in their potential to writing, because only by this way they will be way they will be able to properly take on the authorship in their that discussions of this work can positively impact teacher continuing education and, exact teacher continuing education and, consequently, students formation concerning rad (1902), considering, still, features probably related to the style of the style of translation in literary texts. Initially, pairs of contrast, which indicate. However, the Brazilian public school only offers space for literature in the ng of foreign languages is concerned only with the linguistic aspect of the opinion articles published in Boletim, a weekly newspaper from Universidade Federal relations, a situation that is materialized discursively. Thus, we assume that doxic e relations, this paper will be focused mainly on the works of Munanga (2008) and (in press), it was believed that, proportionally, the GET would tend to produce from a statistical point of view. Finally, it was noted that the first

para ambos os experimentos mostram que **somente** os participantes que foram classificados em suas características visuais, que interferem **diretamente** na interpretação e no significado de multimodal de LD de língua inglesa. **Adicionalmente**, esse estudo se fundamenta na discussão dos possíveis sentidos produzidos, **principalmente**, pelo layout e pela imagem nos e constituintes prosódicos não é isomórfica. **Morfologicamente**, verificamos se a noção de os identificam como palavras prosódicas. **Prosodicamente**, demonstramos se o acento pri compreensão de uma realidade singular, **historicamente** situada, quanto ao ensino-a questão. A natureza do estudo é, **essencialmente**, qualitativa, centrada na descrição escrita, pois só assim poderão assumir **devidamente** a autoria dos textos que produzem. As discussões deste trabalho possam impactar **positivamente** para a formação continuada de a formação continuada de professores e, **consequentemente**, para a formação de alunos o estilo da tradução em textos literários. **Inicialmente**, verificaram-se pares de contraste, os sujeitos, situação que se materializa **discursivamente**. Dessa forma, presumimos qu ciais, esta dissertação irá se basear, **principalmente**, nos trabalhos de Munanga (2008) as cotas no ensino superior, mas **principalmente** sobre as relações raciais de Tenuta (no prelo), acreditava-se que, **proporcionalmente**, o GET tenderia a produzir m, analisando como cada sujeito se inscreve **discursivamente** como sujeito bilíngue, visto q metodológicos seguidos nessa fase foram **critériosamente** escolhidos a fim de se criar s termos prosódicos, distribucionais e funcionais. **Funcionalmente**, o tópico é definido como a possíveis realizações prosódicas do tópico. **Finalmente**, o AE minicorpus é oferecido com associam recursos verbais e não verbais, **especificamente** tirinhas da Mafalda, as quais ser oretto (2010). Esses dois últimos autores, **respectivamente**, auxiliaram nas análises das tip que o trabalho com os textos **somente** verbais que pertencem ao tipo textual motivam questões de cunho gramatical, mas **somente** questões de verificação de leitura de do português brasileiro e inglês americano, **respectivamente**. Durante a análise ficou dar

FIG. 6 e 7 - Linhas de concordância geradas no *AntConc* dos *abstracts* e dos resumos do sufixo *-ly* e *-mente* respectivamente

Fonte: elaborada pelo próprio autor

Analisaremos, nesta sessão, a densidade lexical e complexidade textual dos resumos e *abstracts*. Em relação aos lexemas, os resumos apresentam uma média de 29,52 lexemas por sentença, enquanto nos *abstracts* essa média cai para 16,75. A porcentagem de lexemas por sentença, ou seja, de palavras lexicais nos resumos é de 94,20% e de 55,09% nos *abstracts*: uma queda de 39,11%. Já levando em

conta a complexidade textual, os resumos têm uma média de tamanho de palavra de 5,63 caracteres, média semelhante nos *abstracts*: 5,44. O tamanho médio de uma sentença dos resumos é de 32,02 palavras e de 30,96 nos *abstracts*, uma queda de menos de 2%. No todo, considerando o comprimento dos resumos, a média de número de palavras por texto é de 316,25 e de 314,59 nos *abstracts*. Observando-se o número de sentenças dos textos, os resumos apresentam uma média de 10,45 sentenças, enquanto os *abstracts* apresentam 10,37. Observe o gráfico 4 a seguir:

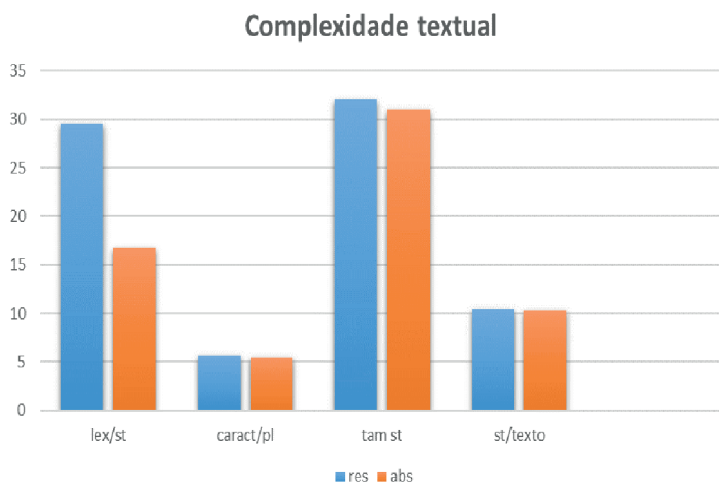


GRÁFICO 4 – Complexidade textual

Fonte: elaborada pelo próprio autor

File Information

File:	Abs/abs_1.txt
Length:	
- Words in text:	356
- Sentences in text:	9
Text Complexity:	
- Av. Word Length:	5.89
- Av. Sentence Length:	39.5
Lexical Density:	
- Lexemes per sentence:	24.8
- Lexemes % of text:	62.92
Reference Density: (% of tokens)	
- 1p Reference:	1.6853
- 2p Reference:	0.0
- 3p Reference:	1.1235

File Information

File:	Res e abs/res_1.txt
Length:	
- Words in text:	400
- Sentences in text:	11
Text Complexity:	
- Av. Word Length:	5.60
- Av. Sentence Length:	36.3
Lexical Density:	
- Lexemes per sentence:	34.7
- Lexemes % of text:	95.5
Reference Density: (% of tokens)	
- 1p Reference:	0.0
- 2p Reference:	0.0
- 3p Reference:	0.0

FIG. 8 e 9 - Resultados da análise estatística feitos pelo UAM para um *abstract* e seu respectivo resumo.

Fonte: elaborada pelo próprio autor

Ao lado podemos ver um exemplo de como a ferramenta *UAM CorpusTool* retorna dados estatísticos para um resumo e seu respectivo abstract. Para analisar os dados demonstrados acima foram analisados todos os dados presentes nas tabelas exceto “*Reference Density*”. Após a coleta dos dados foram feitos simples cálculos matemáticos de média estatística e depois os resultados foram transformados em gráfico.

Como previsto nos universais (BAKER, 1993), houve uma queda muito grande dos lexemas por sentença sem grandes diferenças nos outros indicadores. Isso significa que os tradutores, como já esperado, aumentaram o número de palavras gramaticais do texto e diminuíram o número de palavras lexicais. Mais uma vez os tradutores preferiram explicitar mais o texto traduzido de maneira a simplificar sua leitura, assim possíveis dificuldades de interpretação são eliminadas e o texto se torna mais fácil de processar (BIEL, 2010). É interessante ressaltar que o inglês é uma língua muito mais direta e de padrão de sentença muito mais curto que o do português, então o mais esperado seria que o número de sentenças no inglês aumentasse ao invés de permanecer quase o mesmo.

Por último, vamos estudar os determinantes/pronomes: *other* e *another*. Para entender como os determinantes funcionam na língua inglesa, foi necessário a consulta de uma boa gramática de referência. Para tanto, consultei a gramática do *Cambridge Dictionary Online* para obter uma explicação clara de como *other* e *another* funcionam no inglês. Vale ressaltar que nessa sessão analisei uso certo ou errado de acordo com a gramática tradicional da língua inglesa. Um outro aspecto importante é que analisaremos os usos de *other* e *another* de maneira qualitativa, como observaremos todos os usos fora do prescrito pela gramática da língua inglesa, fazer um estudo estatístico de uma estrutura com apenas quatro ocorrências fora da norma não se faz necessário. A tradução de determinantes

e pronomes é sempre problemática devido ao fato de que em inglês temos duas opções (*other* e *another*) para o que em português corresponde a apenas um item lexical: outro. Como resultado, o estudo qualitativo já serve para demonstrar como mesmo em textos acadêmicos traduzidos, lidar com as diferenças sintáticas e semânticas do uso de pronomes e determinantes pode afastar os textos traduzidos da língua-alvo.

Other pode ter duas funções na língua: determinante e pronome. Como determinantes *other* obedece às características dos determinantes em inglês, ou seja, nunca tem flexão de plural, gênero ou número. Já quando é pronome, ele sempre tem que se referir a algo já dito antes, podendo assim assumir forma de plural. O uso do artigo determinado antes de *other* indica que se faz referência ao segundo item de uma série de duas coisas ou pessoas. Em relação a *another* ele também assume as mesmas funções de *other* de determinante ou pronome. Vejamos essas quatro sentenças problemáticas:

1- *...one based on production and other on comprehension.*

2- *...that the fostering of this parameter is related to others, such as historical, ideological...*

3- *In others words, this research...*

4- *...such tasks articulate and progress within a textbook and in relation to others in the...*

A primeira sentença traz *other* como pronome, nesse caso já que o autor está se referindo ao que foi dito antes falta o artigo *the*. Em relação à sentença dois, *other* também funciona como pronome,

porém sem referente. A terceira sentença apresenta o único caso de erro com *other* no lugar de determinante, nesse caso, ele não admite flexão de plural. Já a quarta sentença, nos mostra *other* de novo sendo usado como pronome, porém o referente está distante demais. A respeito disso, é sempre bom lembrar algumas diferenças entre inglês e português: uma sentença em inglês não pode ter a extensão que essa sentença apresenta, o inglês é uma língua mais direta e de sentenças mais curtas. Portanto, para fazer retomada de *tasks* tão longe na sentença o autor poderia ter dividido a sentença em dois ou colocado o *other* em algum ponto anterior na sentença. A respeito do uso de *another* houve só um problema ortográfico, de acordo com o *Cambridge Online Dictionary* não existe *an other* como foi usado em: “*Aware that an interpretation can always become an other and that...*”

Esses padrões de sentença põem em cheque um dos universais da tradução, o da normatização. De acordo com Baker (1993), os tradutores vão sempre normatizar ao máximo suas traduções a fim de mantê-las o mais simples possível. O problema com essas sentenças é que provavelmente devido à falta de conhecimento do autor da gramática do inglês, alguns erros de normatização ocorreram.

5. Conclusão

Para concluir, podemos observar que um *corpus* comparável tem uma grande utilidade quando se trata de analisar como os tradutores lidam com seus textos.

Ao estudar o universal da explicitação e da simplificação (BAKER, 1993) vimos que o aumento de 30% dos advérbios terminados em *-ly* nos *abstracts* corrobora a teoria. Isso demonstra que ao traduzir um resumo, esse aumento expressivo de advérbios indica que na visão dos tradutores provavelmente o texto se torna-

ria mais processável. Soma-se a isso, o fato de que além de haver um aumento no número de advérbios, em alguns casos houve uma expansão sintática de toda a sentença. Ainda em relação aos casos que corroboram a explicitação, a análise da densidade lexical revela uma diminuição de 39,11% das palavras lexicais por sentença nos *abstracts* enquanto não houve grandes alterações no tamanho das sentenças, nem no número de sentenças por texto. Esses dados indicam um aumento na explicitação e na simplificação, já que temos um resumo que ao ser traduzido se mantém de um mesmo tamanho, mas há diminuição dos lexemas por sentença nos *abstracts*. Portanto, os espaços vazios ocupados pelas palavras lexicais forma ocupados por palavras gramaticais. Mais uma vez, o preenchimento dos *abstracts* com palavras gramaticais tornaria o texto mais explícito e, portanto, mais fácil de ser interpretado.

No tocante ao universal da nivelação (BAKER, 1993), essa característica dos textos traduzidos pode ser observada pela análise de *it* no início de sentença e dos substantivos *purpose*, *objective* e *aim*. O estudo de *it* revelou que há uma alta frequência do uso de *it* no início de sentença, a observação do quão frequente é o uso dessa estrutura, nos permite afirmar que os *abstracts* se parecem muito mais entre si. O mesmo acontece com *propósito*, *objetivo* e *alvo* e suas respectivas traduções *purpose*, *objective* e *aim*, a investigação nos mostra, mais uma vez, que devido à alta frequência das mesmas palavras sendo usadas em inglês, o universal da nivelação pode ser confirmado. Isso significa que os tradutores têm um estilo muito mais próprio de traduzir um resumo que tende a gravitar em torno das mesmas palavras sempre, se comparássemos com o português veríamos uma variedade lexical maior na língua original.

O universal da nivelação não foi confirmado pelo nosso estudo. Primeiramente, o uso frequente de *it* no início de sentença partindo o sujeito gramatical em dois, juntamente com uma ocorrência de

“*it were* + passiva pluralizada” revela que os textos poderiam estar mais naturalmente traduzidos, ou seja, fugir menos da língua-alvo. Um tradutor obviamente se preocupa com a naturalidade de seu texto na língua-alvo, portanto, a produção de um texto de fragmentado e de difícil leitura devido à sintaxe utilizada na construção das sentenças vai de encontro ao proposto por Baker (1993). Outras ocorrências que também vão de encontro ao universal da normatização são as ocorrências de *other* e *another*. Nesse trabalho observamos apenas as ocorrências que fogem à gramática prescritiva, que apesar de serem quatro, já servem para revelar que existem problemas na normatização.

Uma aproximação da problemática dos universais da tradução através de *corpora* comparável é uma técnica promissória e abre novas possibilidades e perspectivas na tradução e nos seus desafios. Exploramos o uso de diversas expressões através de *corpus* comparável e analisamos como os universais da tradução são ou não perceptíveis em estudo de textos traduzidos. Portanto, ter em mãos *corpora* comparáveis nos possibilita ter uma metodologia com um set de ferramentas preparadas para serem usadas.

REFERÊNCIAS BIBLIOGRÁFICAS

BAKER, Mona. *“Corpus Linguistics and Translation Studies: Implications and applications: Text and Technology.* Amsterdam and Philadelphia: John Benjamins, 1993. Disponível em <<http://wenku.baidu.com/view/23fdab40b307e87101f696e4.html>>. Acesso em: 01 jun. 2016.

BERBER SARDINHA, Tony. *Linguística de Corpus.* 1.ed. Barueri: Manole, 2004. 410 p.

BIEL, Łucia. *Corpus-based studies of legal language for translation purposes: Methodological and practical potential.* Aarhus, 2010. Disponível em <<http://bcom.au.dk/fileadmin/www.asb.dk/isek/biel.pdf>>. Acesso em: 01 jun. 2016.

DAVIES, Mark. *The Corpus of Contemporary American English: 520 million words:* software livre. Birgham: Universidade Birgham Young, 1990. Disponível em: < <http://corpus.byu.edu/coca/>>. Acesso em: 08 de ago. 2016.

LAURENCE, Anthony. *Antconc:* software livre. Versão 3.4.4. Tóquio: Universidade de Waseda, 2014. Disponível em: < <http://www.laurenceanthony.net/software/antconc/>>. Acesso em: 02 de jun. de 2016.

O'DONNELL, Mick. *UAM Corpus Tool:* software livre. Versão beta 2.8. Madrid: Universidad Autónoma de Madrid, 2011. Disponível em:< <http://www.corpustool.com/>>. Acesso em: 02 jun. 2016.

OSTER, Urick; Lawick, H. “*Semantic preference and semantic prosody*: A corpus-based analysis of translation-relevant aspects of the meaning of phraseological units.” Espanha: Fundación Caixa Castelló-Bancaixa, 2008. Disponível em <http://www.academia.edu/1065479/Semantic_preference_and_semantic_prosody_A_corpus-based_analysis_of_translation-relevant_aspects_of_the_meaning_of_phraseological_units>. Acesso em: 01 jun. 2016.