

## INVESTIGANDO A TRADUÇÃO DE ACRÔNIMOS E SIGLAS DA ÁREA DE SENSORIAMENTO REMOTO: UMA PROPOSTA DO PONTO DE VISTA FRASEOLÓGICO

*Dalila dos Santos Hasmann  
Diva Cardoso de Camargo  
Adriane Orenha-Ottaiano*

### RESUMO

Com base nos Estudos da Tradução Baseados em *Corpus* (BAKER, 1995), na Linguística de *Corpus* (BERBER SARDINHA, 2004) e na Fraseologia (PAVEL, 2003), e com o auxílio do programa WordSmith Tools 6.0, visamos analisar um *corpus* de sensoriamento remoto, a fim de identificar siglas e acrônimos em inglês e buscar seus equivalentes em português. Resultados apontam que não há uma padronização nas formas de traduzir tais siglas e acrônimos.

**PALAVRAS-CHAVE:** siglas; estudos da tradução baseados em *corpus*; sensoriamento remoto.

### Introdução

**S**ensoriamento remoto (em português brasileiro), *deteção remota* ou *tele-deteção* (em português europeu), é o conjunto de técnicas que possibilita a obtenção de informações sobre alvos na superfície terrestre. Esses alvos podem ser objetos, áreas ou fenômenos, e sua captação se dá através do registro da interação da radiação eletromagnética com a superfície, realizado por sensores distantes ou remotos. Geralmente esses sensores estão presentes em plataformas orbitais ou satélites, aviões e em nível de campo (como os sensores portáteis,

por exemplo). A NASA é uma das maiores captadoras de imagens recebidas por seus satélites. No Brasil, o principal órgão que atua nesta área é o Instituto Nacional de Pesquisas Espaciais – INPE. A definição científica mais usual para o sensoriamento remoto é dada por Jensen:

Um instrumento de sensoriamento remoto coleta informação sobre um objeto ou fenômeno dentro do campo de visada instantâneo do sistema sensor sem estar em contato físico direto com ele. O instrumento pode estar localizado a apenas poucos metros acima do solo e/ou a bordo de uma aeronave ou satélite (JENSEN, 2009, p. 3).

Os dados resultantes dessa técnica são úteis para o estudo de alguns dos temas que se tornaram comuns nestes últimos anos, como aquecimento global, desastres ambientais, desmatamento da Amazônia, efeito estufa, geração de energia sustentável, mudanças climáticas, previsão do tempo entre muitos outros, que utilizam e se beneficiam diretamente da tecnologia de sensoriamento remoto. Quando se trata de estudar áreas imensas e de condições extremas como a Amazônia, por exemplo, o avanço deste campo de pesquisa contribui concretamente para a obtenção de conhecimento dos ecossistemas dessa região.

Tais estudos e obtenção de dados sobre nosso território são de interesse não só do Brasil, mas do mundo. Segundo os próprios especialistas da área “há uma urgente necessidade de uniformização de linguagem e protocolos no âmbito da comunidade de sensoriamento remoto, para que a comunicação, difusão e troca de informações e ideias se tornem mais profícuas” (NOVO *et al.*, 2005, p. 270).

Compartilhando desse ponto de vista, já no ano de 1989, pesquisadores da SELPER – *Sociedad de Especialistas Latinoamericanos en Percepción Remota* (Associação de Especialistas Latino-americanos em Sensoriamento Remoto) – juntaram-se e produziram um dicionário trilingue de terminologia em inglês, espanhol e português. Essa empreitada teve a participação de diferentes países da América Latina e Caribe, envolvendo o INPE e a Agência Canadense de Desenvolvimento Internacional (*Canadian International Development Agency* – CIDA). A publicação desse dicionário foi uma contribuição da CIDA na

transferência de tecnologia de sensoriamento remoto para países de línguas espanhola e portuguesa. Essa iniciativa é um dos componentes do Teledetecção II, um projeto de transferência de tecnologia em sensoriamento remoto do Canadá ao Peru.

Durante vários anos de experiência da CIDA na transferência de tecnologia para países em desenvolvimento, tornou-se evidente que uma das barreiras mais difíceis de transpor era a linguagem e, mais particularmente, a terminologia.

Em sensoriamento remoto, a terminologia se divide em dois grandes grupos. O primeiro grupo compreende os termos que já existiam no inglês antes do desenvolvimento de sensores remotos (advindos de disciplinas-mãe), mas para alguns dos quais o uso corrente tem dado um significado não convencional e muitas vezes confuso (SELPER, 1989). O segundo grupo consiste de novas palavras criadas pela tecnologia. Em muitos casos, não existem definições de qualquer tipo para essas palavras. Os usuários que não estão familiarizados com o idioma inglês, e que não têm acesso à extensa literatura técnica da área, muitas vezes têm dificuldade em inferir uma definição de trabalho. Neste grupo de termos, podemos incluir as siglas, amplamente utilizadas em sensoriamento remoto, mas que são, em sua grande maioria, ininteligíveis para o recém-chegado. Para os autores do glossário, “[o] lamentável resultado dos problemas de linguagem e terminologia é que somente os profissionais com sólido conhecimento da língua inglesa podem participar ativa e efetivamente em sensoriamento remoto<sup>1</sup>” (SELPER, 1989). Considerando-se esses obstáculos, a CIDA concordou em incluir um dicionário de termos de sensoriamento remoto no Projeto Teledetecção II.

Passados dezoito anos dessa publicação, no ano de 2007 houve uma segunda iniciativa dos pesquisadores que compõem a SELPER para atualizar o dicionário e disponibilizá-lo *on-line*. Sem dúvida, foi um enorme avanço poder tornar esse trabalho acessível por meio da internet. Além da atualização dos termos, também foi adicionado um quarto idioma, o francês. Na interface virtual, o usuário, ao iniciar sua consulta, escolhe em que língua

---

<sup>1</sup> The unfortunate result of the problems of language and terminology is that only professionals with a solid knowledge of English can participate actively and effectively in remote sensing (SELPER, 1989) [Todas as traduções são de nossa responsabilidade].

fará as buscas. A divulgação é promovida e assistida pela SELPER Internacional e pelos membros que participaram de sua revisão e atualização. Para os autores:

A “área dos sensores remotos” é um campo de estudo multidisciplinar a qual convergem ramos das ciências e da tecnologia. Como tal, requer, para sua atividade, o uso de termos científicos que ocorrem transversalmente. As ciências e a tecnologia, por sua vez, geram uma nova terminologia produto de sua relação com o campo dos sensores remotos<sup>2</sup>.

Assim, justifica-se a escolha do sensoriamento remoto como foco principal deste estudo. Por meio de sua multidisciplinaridade, promover o conhecimento de sensoriamento remoto é disseminar informação e contribuir para diversas outras áreas. É necessário haver mais uma atualização deste grande trabalho iniciado há mais de duas décadas, dessa vez não só sob a ótica da Terminologia, em busca de termos, mas também de expressões e combinações de termos (colocações especializadas) que carregam sentidos próprios. Essa atualização é, acima de tudo, necessária, pois, como em qualquer outra tecnologia, a linguagem em sensoriamento remoto avança dia a dia com o desenvolvimento de novas técnicas, com o lançamento de novos foguetes, satélites, plataformas orbitais, estações espaciais, *softwares* etc.

Desta forma, os objetivos deste artigo são: 1) analisar um *subcorpus* paralelo (original) em inglês da área de sensoriamento remoto em busca das palavras-chave mais frequentes; 2) partir das palavras-chave selecionadas para extrair as siglas e acrônimos originados por elas; 3) proceder à busca de equivalentes no *subcorpus* paralelo (traduzido) em língua portuguesa; e 4) buscar validação nos *corpora* comparáveis em inglês e em português, para verificar situações reais de uso dessas siglas e acrônimos pela comunidade do sensoriamento remoto. Por meio destes objetivos, tem-se a intenção de apresentar a importância do acesso a *corpora* comparáveis na decisão pelo melhor equivalente para algumas siglas e colocações, bem como alguns acrônimos, que podem vir a constituir desafios para os tradutores.

---

<sup>2</sup> SELPER - Dicionário Latino-americano de Sensoriamento Remoto. Disponível em: <<http://www.inpe.br/unidades/cep/atividadesccep/dicionarioselper/portugues/home.htm>> 06/05/2013.

## Fundamentação teórica

Por questões de economia linguística, o uso de unidades terminológicas sintagmáticas acaba resultando no uso recorrente de siglas e acrônimos. Siglas e acrônimos são processos linguísticos ao qual o falante recorre para dar mais fluidez ao seu discurso. As siglas se formam pela redução de um grupo de palavras às suas iniciais e são lidas letra por letra (ex.: IBGE, INSS) e os acrônimos se formam através da junção dessas mesmas letras iniciais de um grupo de palavras, mas com a diferença de serem pronunciados como uma palavra só (ex.: Inpe, Ibama)<sup>3</sup>. Para Nadin (2013, p. 180), “[a] variação denominativa provocada pelo uso de siglas é, pois, um fenômeno relevante na descrição do uso especializado de uma língua. Dita variação pode provocar variação conceitual ou interferir na comunicação especializada”.

Para atingir os objetivos deste trabalho, elegemos a abordagem da Linguística de *Corpus*. Na visão de Baker (1995, p. 226), entende-se por *corpus* “um conjunto de textos naturais” em oposição a exemplos/sentenças criados com o propósito específico de mostrar um fenômeno linguístico, “organizados em formato eletrônico, passíveis de serem analisados, preferencialmente, de forma automática ou semi-automática (em vez de manualmente)”<sup>4</sup>. Um dos tipos de *corpora* utilizados neste trabalho é o paralelo. Para a autora, o *corpus* paralelo é composto de textos originais em determinada língua de origem (neste caso, inglês) e suas respectivas traduções para outra língua (língua da tradução, sendo aqui o português). Esse tipo de *corpus* permite pesquisar traduções consagradas de certos itens lexicais ou estruturas sintáticas, peculiaridades de determinado(s) tradutor(es), diferenças entre traduções de um mesmo texto, produzidas em períodos diversos, normas tradutórias, etc.

Outro *corpus* utilizado, o comparável, “como sugestão da própria palavra pode ser comparado de acordo com critérios externos e dar ideias sobre

<sup>3</sup> DT - Dicionário Terminológico para consulta em linha do Ministério da Educação de Portugal. Disponível em: <dt.dgic.min-edu.pt> “procurar por *sigla e acrônimo*”. Acesso em: 20 mar. 2013.

<sup>4</sup> *Corpus* mean[s] any collection of running texts (as opposed to examples/sentences), held in electronic form and analysable automatically or semi-automatically (rather than manually) (BAKER, 1995, p. 226).

dois sistemas linguísticos independentes”<sup>5</sup> (TOGNINI-BONELLI, 2001, p. 133). Para a autora, o ideal é que o *corpus* comparável seja composto de: 1) um *corpus* comparável de textos originais na língua 1 (neste caso, inglês); 2) um *corpus* de estudo com os textos traduzidos em questão (o qual pode ser ou não um *corpus* paralelo); e 3) um *corpus* comparável de textos originais na língua 2 (neste caso, português). Segundo Camargo (2007, p. 49), “esse tipo de *corpus* permite uma investigação mais completa de textos técnicos ou especializados”; possibilita identificar padrões que são ou restritos ao texto traduzido ou que ocorrem em frequências significativamente mais altas ou mais baixas no texto traduzido do que em relação aos textos originais. Outra contribuição desse tipo de estudo é a elaboração de glossários de termos especializados acompanhados de seu contexto, os quais aparecem com maior frequência na língua de especialidade representada nos textos traduzidos e textos originais selecionados para análise. Por língua de especialidade, entende-se:

A língua de especialidade é um subconjunto da língua geral que serve para transmitir um saber atinente a um campo de experiência particular. Ela tem em comum com a língua geral a gramática e uma parte de seu inventário léxico-semântico (morfemas, palavras, sintagmas e regras combinatórias), mas faz deles um uso seletivo e criativo que reflete as particularidades dos conceitos em jogo e que apresenta variações sociais, geográficas e históricas (PAVEL, 2003, p. 100).

Com o avanço da pesquisa em diversas áreas do conhecimento, houve um aumento no número de estudos relacionados aos discursos especializados. A tradução tem sido uma forte aliada e tem contribuído significativamente para a ampla divulgação do que está sendo produzido atualmente. Como recurso auxiliar fundamental para o exercício do tradutor e dos próprios pesquisadores, pode-se mencionar os glossários bi- ou multilíngues e os dicionários terminológicos das mais diversas especialidades. No entanto, muitos

---

<sup>5</sup> Comparable *corpora*, as the word suggests, can be compared according to external criteria and give insights into two independent linguistic systems (TOGNINI-BONELLI, 2001, p.133).

deles apenas enfocam termos e, ao fazê-lo, deixam de fora as combinatórias de palavras, tais como, as *colocações* em contextos especializados, denominadas por Orenha-Ottaiano (2009, p. 32) de *colocações especializadas*. Um trabalho voltado para o estudo dessas colocações mostra-se importante, uma vez que a busca por combinações de palavras (colocações) equivalentes e correspondentes torna-se, em geral, mais difícil para o tradutor do que a busca por equivalentes de termos simples. Por termo, ou unidade terminológica, entende-se: “uma palavra (simples, derivada, composta) ou um grupo compacto de palavras (sintagma lexical, nominal, verbal ou adjetival) que designa um conceito de tipo objeto, ação ou propriedade, pertencendo ao sistema conceitual de uma especialidade” (PAVEL, 2003, p. 105).

Ao adentrarmos o ambiente da Fraseologia, rapidamente percebemos haver uma série de perspectivas para seu tratamento, uma rica diversidade denominativa e conceitual, bem como uma grande dificuldade de delimitar e classificar seu objeto de estudo. Por essa razão, com o intuito de uniformizar a nomenclatura utilizada neste trabalho, optamos por nos basear na pesquisa de Orenha-Ottaiano (2009). Conforme mencionamos, a autora refere-se a essas combinatórias sintagmáticas de unidades terminológicas, presentes nas línguas de especialidade, por meio do termo *colocações especializadas*, assim, diferenciando-as das *colocações* da língua geral.

De acordo com Anderson (2009, p. 53) “[o] termo *colocação* refere-se à tendência das palavras de ocorrer no ambiente próximo de outras determinadas palavras”<sup>6</sup>. A esse respeito, Firth (1957), em seu artigo *Modes of Meaning*, também criou a frase: “*you shall know a word by the company it keeps*”, evidenciando que uma palavra é conhecida por aquelas que a acompanham.

Orenha-Ottaiano (2009, p. 33-4) destaca que as ligações entre os elementos que formam essas colocações não são semânticas, mas sintagmáticas e convencionais: “são regidas pela convenção que lhes foi imposta para que, desse modo, seus elementos possam se combinar. Há, dessa forma, uma restrição combinatória entre seus elementos”. Para a autora, assim como as colocações, que pertencem à língua geral, as colocações especializadas, pertencentes às línguas de especialidade, também podem ser definidas basi-

<sup>6</sup> The term collocation refers to the tendency of words to occur in the close environment of particular other words (ANDERSON, 2009, p. 53).

camente da mesma forma. Ambas são definidas pelo uso convencional; no entanto, a convenção, no caso das colocações especializadas, é estabelecida dentro de uma comunidade linguística específica, dentro de um grupo de especialistas.

Segundo Pavel (2003, p. 106), por fraseologia dentro de uma língua de especialidade entende-se a “combinatória sintagmática das unidades terminológicas [UT] decorrentes de uma estrutura conceitual coerente. As UT são tomadas como núcleo de coocorrências usuais ou privilegiadas nos textos de uma especialidade”.

Pelo fato de as combinações lexicais especializadas estarem inseridas em uma comunidade linguística específica, como um subconjunto da língua geral, e pertencerem ao “seletivo e criativo inventário léxico-semântico desse subconjunto” (ORENHA-OTTAIANO, 2009, p. 44), os tradutores precisam adquirir conhecimento a respeito do comportamento de tais unidades terminológicas para, assim, poderem utilizá-las em um contexto adequado. A este respeito, Cowie (1998, p. 14) justifica que a falta de domínio de expressões idiomáticas e colocações pode ter um efeito negativo sobre a eficácia da escrita [ou tradução], capaz de desviar a atenção do leitor do conteúdo da mensagem para a forma linguística e, como consequência, esta deficiência pode acabar não atingindo as expectativas estilísticas de uma comunidade acadêmica. O autor defende a noção de que a proficiência em uma língua estrangeira depende consideravelmente de um estoque de unidades pré-fabricadas - ou *'prefabs'* - que variam em complexidade e estabilidade interna. À luz de Carneado Moré (1983, p. 39), nota-se que:

Há muito tempo lexicógrafos observaram que o significado de uma palavra isolada nunca se perde. Isto frequentemente acontece quando uma palavra, ao juntar-se com outra, forma uma combinação indivisível. Daí a necessidade de os dicionários incluírem combinações deste tipo<sup>7</sup>.

<sup>7</sup> Hace tiempo los lexicógrafos notaron que la significación de una palabra aislada nunca se pierde. Esto sí sucede, frecuentemente, cuando una palabra, al unirse con otra, forma una combinación indivisible. De ahí la necesidad de incluir en los diccionarios combinaciones de este tipo (CARNEADO MORÉ, 1983, p. 39).

Para Croitoru (2006, p. 298), o domínio no uso de colocações parece ser o ponto mais fraco de um falante não nativo. Esse aspecto possibilita que um falante nativo da língua-alvo perceba ou identifique que o tradutor não é um falante nativo da língua para qual está traduzindo. Além disso, as colocações estranhas e não usuais que o tradutor acaba “criando”, sob influência do texto-fonte, serão interpretadas como colocações incomuns em sua própria língua materna. Quanto a uma estratégia de tradução que se baseia mais na forma (ao invés do conteúdo) da mensagem, deve-se considerar que pode ser uma abordagem enganosa e até resultar em uma “falha pragmática”, ou seja, pode demonstrar uma incapacidade do tradutor de compreender o significado do que é dito pela maneira que é dito. Além disso, a transcodificação palavra por palavra pode nem sempre ser apropriada e, às vezes, até mesmo tornar-se totalmente inapropriada. Em razão de considerarmos o automatismo na produção fraseológica, também compartilhamos da visão de que a tradução de colocações especializadas costuma ser uma tarefa mais árdua para o tradutor, pois podem conter vários aspectos e transmitir significados que não podem ser inferidos a partir da soma dos elementos que os constitui (ORENHA-OTTAIANO, 2009, p. 52).

A fraseologia representa um grande desafio para a prática lexicográfica. A própria natureza de dicionários, com a sua organização linear, obriga as palavras a serem consideradas isoladamente. No entanto, a evidência de textos, incluindo *corpora*, mostra claramente que as palavras não são isoladas; pelo contrário, conectam-se umas com as outras através de uma padronização fraseológica e têm uma relação de significação a partir do contexto de uso. Além disso, os próprios modelos de produção de linguagem apontam para sua natureza fragmentada e fraseológica (MOON, 2008).

## Metodologia

Neste trabalho, usamos *corpora* paralelo e comparáveis, além de dois *corpora* de referência da língua geral. O *corpus* paralelo consiste de dois *subcorpora*: 1) *subcorpus* original, composto por um livro originalmente escrito em inglês, doravante ScPI; e 2) *subcorpus* traduzido, composto pela respectiva tradução para português (ScPP). Ambos os livros foram digitalizados por meio de escaneamento, utilizando o Reconhecimento Óptico de Caracteres. A obra original

digitalizada foi *Remote Sensing of the Environment: An Earth Resource Perspective*, de John R. Jensen, 2ª ed., lançado pela Editora Pearson Prentice Hall, em 2007, contendo 592 páginas, e a versão traduzida *Sensoriamento Remoto do Ambiente: Uma Perspectiva em Recursos Terrestres*, lançada pela Editora Parêntese, 2009, que conta com 672 páginas. Sua tradução foi feita por pesquisadores do INPE, o que confere a esta edição do livro uma confiabilidade técnico-científica.

Já os *corpora* comparáveis, ambos são constituídos por artigos científicos de sensoriamento remoto, publicados em revistas nacionais e internacionais renomadas na área. Para a coleta do *corpus* comparável de língua inglesa (doravante CCI) encontramos a base de dados *Remote Sensing Magazine*, que oferecia importante contribuição para as necessidades da pesquisa. Ao final da recolha dos textos, obtivemos um total de 114 artigos que contêm 697.320 palavras, representando mais que o triplo do tamanho do *corpus* de estudo. O *corpus* comparável de língua portuguesa (CCP) foi coletado por meio da base de periódicos da *SciELO*. Ao final, atingimos um total de 201 artigos coletados contendo 598.390 palavras corridas. Esta comparação entre o *corpus* de estudo e os comparáveis, tem o propósito de examinar o uso das colocações especializadas (acrônimos e/ou siglas) mais frequentes encontradas no *corpus* paralelo e identificar seus padrões formais de uso.

No tocante ao *corpus* de referência, é formatado como uma lista de frequência de palavras e funciona como termo de comparação para a análise. “A sua função é fornecer uma norma com a qual se fará a comparação das frequências do *corpus* de estudo” (BERBER SARDINHA, 2004, p. 97). Um *corpus* da língua geral normalmente é composto por milhões de palavras e é utilizado para contrastar a frequência das palavras encontradas no *corpus* de estudo. Desse modo, se uma palavra tiver alta frequência no *corpus* de estudo, porém baixa frequência no *corpus* de referência, ou seja, alta frequência na área de sensoriamento remoto e baixa frequência na língua geral, esta palavra poderá ser considerada *chave* e passa a ser uma candidata à análise. O *corpus* de referência de língua inglesa utilizado é o *BNC*, com 100 milhões de palavras do inglês britânico escrito e falado e, para o *corpus* de referência de língua portuguesa, foi utilizado o Lácio-Ref, que possui em torno de 8 milhões de palavras do português brasileiro contemporâneo escrito.

O manuseio dos *corpora* com rapidez e precisão é possível graças ao *software* WordSmith Tools (SCOTT, 2012), versão 6.0, que possui três

ferramentas principais: WordList, KeyWord e Concord, as quais fornecem, respectivamente, a) produção de lista de palavras contendo todas as palavras do arquivo selecionado, b) extração de palavras-chave e c) extração de linhas de concordância. Esse programa tem contribuído para a divulgação da Linguística de *Corpus* no Brasil.

## Análise dos dados

Quanto aos primeiros resultados obtidos da compilação dos *corpora*, utilizamos a ferramenta WordList para criar as listas de palavras dos *subcorpora*, bem como dos comparáveis, e apresentamos, na Tabela 1, alguns resultados estatísticos:

	ScPI (original)	ScPP (traduzido)	CCI	CCP
Formas	12.797	15.987	17.929	24.120
Itens	226.285	262.877	697.320	598.390

Tabela 1. Estatística simples a partir dos *corpora* utilizados.

Ao analisarmos os primeiros resultados estatísticos, podemos nos ater, inicialmente, aos valores dos *itens*. São considerados *itens* todas as palavras de um *corpus*, portanto, todas são contabilizadas, não importando quantas vezes apareçam. O número de *itens* de ambos os *corpora* comparáveis (CCI e CCP) é, em média, 3 vezes maior do que os *subcorpora* paralelos (ScPI e ScPP). Por exemplo, o CCI possui 697.320 *itens*, dividido pela quantidade de *itens* do ScPI (226.285 *itens*), o resultado é aproximadamente 3,08. Isto quer dizer que o CCI é em torno de três vezes maior do que o ScPI em número de palavras totais utilizadas nos *corpora*. Se calcularmos essa diferença em forma de porcentagem, teremos a informação de que o CCI é 208% maior do que o ScPI e o CCP é 127% maior que o ScPP.

No entanto, ao analisarmos a coluna das *formas* (onde são contabilizados apenas os vocábulos do *corpus*, ou seja, cada vocábulo é contabilizado apenas uma vez) percebemos que a porcentagem de *formas* que os *corpora* comparáveis têm a mais do que os *subcorpora* paralelos é relativamente pequena se comparada aos números de *itens*. Em números percentuais, o CCI é apenas

40% maior do que o ScPI, enquanto o CCP é 50% maior que o ScPP. Os dados estatísticos estão nos mostrando que, embora o *corpus* comparável de inglês possa ser mais de duzentos por cento maior que o *subcorpus* paralelo de inglês em número total de palavras, ao se tratar do número de *formas*, o número foi superior em apenas quarenta por cento.

Esse fenômeno indica que, em termos de *itens*, o *corpus* poderia continuar crescendo exponencialmente, conforme fôssemos alimentando-o, mas, a partir de certo ponto, esse crescimento deixaria de representar uma elevação significativa no que diz respeito ao número de *formas*. Os vocábulos tenderiam a se repetir cada vez mais, devido ao grau de especificidade da área escolhida. Isto remete a uma passagem onde Alves e Tagnin (2011, p. 194) afirmam que, quando se trata de *corpus* especializado, o tamanho do *corpus* não é o aspecto mais importante de todos. Segundo os autores, o número estipulado de palavras num *corpus* dependerá do objetivo e do tema a ser investigado. “É de se supor que, quanto mais específica a área, menor o número de textos para cobrir todo o vocabulário”. Dessa forma, acreditamos que os *corpora* utilizados nesta pesquisa podem ser considerados satisfatórios para análise com *corpus*.

Dando prosseguimento ao trabalho, as listas de palavras dos *subcorpora* paralelo geradas pelo WordList foram confrontadas com os *corpora* de referência (de língua geral inglês e português) e geraram uma nova listagem, dessa vez de palavras-chave (*Keywords*). No Quadro 1, trazemos uma amostra das 50 primeiras palavras-chave retiradas do ScPI (original), para podermos proceder a uma análise:

REMOTE	SPATIAL	ACCURACY	PIXEL	INFRARED
SENSING	NDVI	NM	IMAGES	PIXELS
DATA	MODIS	FOREST	SENSOR	HYPERSPECTRAL
M	FIGURE	VALUES	ET	BAND
VEGETATION	USING	SOIL	TREE	MEASUREMENTS
SPECTRAL	RESOLUTION	SATELLITE	WATER	CANOPY
LIDAR	IMAGERY	MAPPING	AL	BIOMASS
REFLECTANCE	LANDSAT	COVER	MODEL	TM
IMAGE	BASED	SURFACE	LAND	MODELING
CLASSIFICATION	IRRIGATED	USA	ANALYSIS	URBAN

Quadro 1. Lista das 50 primeiras palavras-chave extraídas do ScPI (original).

Nota-se que tais palavras-chave não são tão comuns (ou não possuem frequência alta) na língua geral. Palavras como *spectral*, *spatial*, *satellite*, *sensor*, *infrared*, *hyperspectral*, por exemplo, não nos parecem corriqueiras no dia a dia de um falante que não esteja inserido na área de especialidade do sensoriamento remoto e é justamente isso que as torna chave. No entanto, é necessário um refinamento manual desses dados. Numa primeira análise desses resultados, descartamos palavras como “*m*”, abreviatura de ‘metro’; “*figure*”, pois se trata da demonstração de imagens, não necessariamente de terminologia própria da área; *nm*, abreviatura para ‘nanômetro’; as palavras “*et*” e “*al*”, pois se referem às citações teóricas utilizadas no decorrer da escrita acadêmica; bem como “*USA*”. Feito esse refinamento e removidas as palavras consideradas pouco interessantes para esta análise, podemos proceder a um estudo das correlações de cada uma delas em contexto.

O próximo passo foi, então, elaborarmos as linhas de concordâncias utilizando a ferramenta Concord, em busca de colocações especializadas originadas pelas palavras-chave. Cada uma das palavras-chave foi inserida individualmente no campo *Search Word* (palavra de busca) e, com isso, foi gerada uma lista de concordância para cada. Utilizou-se para observação não apenas as linhas de concordância em si, mas também a aba dos *clusters* (agrupamentos lexicais) e dos *collocates* (colocados). Essas abas nos permitem visualizar claramente quais são as palavras que coocorrem com a palavra de busca. Conforme observou Berber Sardinha (2004), essas concordâncias possibilitam ao pesquisador ver a palavra que está sendo buscada, também chamada de nóculo (*node*), acompanhada de seu cotexto – palavras que ocorrem junto ao nóculo no *corpus*. Em razão dessa aplicação, essa ferramenta é indispensável para pesquisas em fraseologia, como é o caso da nossa. Apresentamos no Quadro 2, a seguir, a lista de acrônimos e siglas que selecionamos a partir deste levantamento:

Palavras-chave	Siglas e acrônimos
INFRARED	LWIR - Long-Wavelength Infrared
	NIR - Near Infrared
	SWIR - Short Wavelength Infrared
AERIAL	UAV – Unmanned Aerial Vehicle

REFLECTANCE	BRDF - Bidirectional Reflectance Distribution Function
	BRF - Bidirectional Reflectance Factors
	CARI - Chlorophyll Absorption in Reflectance Index
	TCARI - Transformed Absorption in Reflectance Index
THERMAL	TIR - Thermal Infrared
	TIRS - Thermal Infrared Sensor
VEGETATION	EVI - Enhanced Vegetation Index
	NDVI - Normalized Difference Vegetation Index
	NVI - New Vegetation Index
	OSAVI - Optimized Soil-Adjusted Vegetation Index
	SAVI - Soil-Adjusted Vegetation Index
	TVI - Transformed Vegetation Index
RADAR	SAR - Synthetic Aperture Radar
DIGITAL	DEM - Digital Elevation Model
	DOQQ - Digital Orthophoto Quarter Quad
	DSM - Digital Surface Model
	DTM - Digital Terrain Model
SURFACE	SST - Sea-Surface Temperature
VISIBLE	VNIR - Visible and Near Infrared
INFORMATION	GIS - Geographic Information System
	GPS – Global Positioning System

Quadro 2. Acrônimos e siglas selecionados a partir das palavras-chave do ScPI.

Como podemos ver, nem todas as palavras-chave originaram acrônimos ou siglas. Por outro lado, a palavra *vegetation*, por exemplo, originou um total de seis siglas e acrônimos, apenas no ScPI. Outras siglas foram encontradas nos *corpora* comparáveis no decorrer de nossa análise, mas não pertenciam aos objetivos estabelecidos para este trabalho e foram, portanto, excluídas. Nosso próximo passo foi buscar pelos equivalentes das siglas no ScPP (traduzido) e, simultaneamente, buscar nos *corpora* comparáveis (CCI e CCP) usos que possam confirmar, refutar ou, talvez, apontar para fenômenos adversos.

## BUSCANDO VALIDAÇÃO NOS *CORPORA* COMPARÁVEIS

Juntamente aos equivalentes encontrados no ScPP (traduzido), procedemos a uma busca desses resultados em ambos os *corpora* comparáveis, na tentativa de compreender como se dá a tradução de acrônimos e siglas da área de especialidade do sensoriamento remoto. O trabalho com *corpus* nos dá a possibilidade de identificar as opções de tradução mais utilizadas pelos membros dessa comunidade e/ou descobrir quais expressões são mais usuais na comunidade falante de língua portuguesa.

As siglas originadas a partir do termo *infrared* são, aparentemente, utilizadas em larga escala, pois foram fáceis de encontrar em todos os *corpora* (paralelos e comparáveis). No entanto, notamos uma não linearidade no momento de escolher entre traduzir ou não as siglas para língua portuguesa. As siglas LWIR e SWIR puderam ser encontradas na tradução para o português por estas mesmas iniciais, no entanto, eram frequentemente acompanhadas da tradução de suas formas plenas, ou seja, apareciam as siglas seguidas das colocações especializadas, as quais eram empregadas para explicá-las. Por exemplo, encontramos para LWIR a colocação traduzida da seguinte maneira “infravermelho de ondas longas (LWIR)”. O mesmo ocorreu com a sigla SWIR, a qual foi traduzida pela própria sigla em inglês, seguida da colocação especializada “infravermelho de ondas curtas (SWIR)”. Porém, no caso de NIR (*near infrared*), que tem como pivô o mesmo termo (*infrared*), tanto a colocação quanto a sigla foram completamente traduzidas para língua portuguesa e têm grande utilização na literatura. Em língua portuguesa, a tradução do conceito é “infravermelho próximo” e a sigla se tornou IVP. O tradutor, ou mesmo aluno, interessado em lidar com textos de sensoriamento remoto precisa estar ciente do quanto os termos estão consolidados ou não na área para não acabar “criando” colocações inexistentes ou não aceitas pela maioria dos usuários.

Para nossa surpresa, não encontramos no CCP nenhuma colocação que remeta à sigla UAV (*Unmanned Aerial Vehicle*). No ScPP, encontramos o conceito traduzido por “veículo aéreo não-tripulado”, mas sem menção à sigla nem em português nem em inglês. Isto causou certo estranhamento, pois em uma rápida busca na internet, identificamos que, em português, usa-se o acrônimo VANT para se referir a esta colocação. O buscador do Google trouxe 154.000.000 resultados para a busca por VANT, o que nos faz refletir sobre o

porquê de os tradutores não terem optado pela utilização deste acrônimo se, aparentemente, sua aceitação já está cristalizada em português.

Para as colocações originadas do termo *reflectance*, obtivemos BRDF e BRF que tiveram seus conceitos e siglas igualmente traduzidos para português. BRDF em português é “Função Distribuição da Reflectância Bidirecional” e a sigla tornou-se FDRB, bem como BRF em português é “Fator de Reflectância Bidirecional” e a sigla FRB. As siglas CARI e TCARI não foram encontradas nos *corpora* comparáveis. No ScPP, estas tiveram seus conceitos traduzidos, mas as siglas foram mantidas em inglês.

As colocações originadas do nódulo *thermal*, TIR e TIRS, embora muito semelhantes, tiveram diferenças nos *corpora* comparáveis. A sigla TIR teve grande recorrência em ambos os *corpora* comparáveis. Em português, seu conceito foi traduzido para “infravermelho termal”, mas a sigla remanesceu em inglês. Já TIRS teve poucas ocorrências em formato de sigla, mesmo no ScPI (original). A preferência era mais pela forma plena da colocação; por extenso. Em português, tanto no ScPP (traduzido) quanto no CCP, foi encontrada apenas a utilização do conceito traduzido, mas não a sigla. Aparentemente, em língua portuguesa TIRS perdeu seu status de sigla.

O termo que mais originou colocações, *vegetation*, também apresentou variações na tradução. Três das seis colocações, a saber NDVI, EVI e TVI, tiveram seus conceitos traduzidos para português, mas as siglas foram mantidas em inglês. Contrariamente, SAVI e OSAVI tiveram conceitos e siglas traduzidos e se tornaram, respectivamente, IVAS e IAVAS com alta frequência no CCP. Já a última colocação originada por *vegetation*, NVI, teve variação entre os próprios *corpora* de língua inglesa, sendo encontrada como “*New Vegetation Index*” e “*Normalized Vegetation Index*”. Acreditamos que seria necessário requerer a participação de um profissional envolvido com sensoriamento remoto para nos ajudar, pois apenas o trabalho com *corpus* não elucidaria questões como esta. Um caso que se assemelha a este é o da colocação originada pelo termo *digital*, DOQQ. No ScPP (traduzido), encontramos sua tradução, mais de uma vez, como “ortofotos digitais”; no entanto, o uso desta colocação no *subcorpus* traduzido é muito mais expressivo do que o uso de DOQQ no *subcorpus* original, pois também temos em inglês a colocação “*digital orthophoto*”. Nesse caso, “ortofotos digitais” estaria servindo de tradução para duas colocações aparentemente distintas. Ao percorrermos CCP em busca de “ortofotos

digitais”, encontramos várias ocorrências, mas não podíamos esclarecer se estariam representando “*digital orthophoto*” ou DOQQ. Apenas um especialista poderia fazer tal afirmação.

As demais colocações originadas do termo *digital*, DEM, DSM e DTM, tiveram seus conceitos e siglas traduzidos para português e deram origem, respectivamente, às siglas MDE, MDS e MDT que têm grande utilização nos textos em língua portuguesa. Contrariamente a estas, estão as colocações SAR, originada pelo termo *radar*, SST, originada pelo termo *surface* e VNIR, do termo *visible*, que tiveram seus conceitos traduzidos, mas as siglas foram mantidas em inglês.

A última dupla de colocações analisadas neste trabalho foi originada do termo *information*. As colocações são GIS e GPS que possuem igualmente alta frequência nos *corpora*, mas suas traduções estão em extremos opostos. A colocação GIS (*Geographic Information System*) foi traduzida para português e a sigla que a representa (também traduzida) é SIG (Sistema de Informação Geográfica). No CCP, obtivemos apenas 9 vezes a sigla GIS contra 59 utilizações da versão traduzida SIG, o que pode indicar que a colocação traduzida está plenamente cristalizada na comunidade brasileira. Do outro lado está GPS, que teve seu conceito traduzido, mas a sigla foi mantida em inglês. No CCP, obtivemos 118 ocorrências de GPS, sendo 15 delas acompanhadas do conceito em inglês (*Global Positioning System*) e 5 com o conceito traduzido para português (Sistema de Posicionamento Global), mas não houve nenhuma ocorrência de tentativa de utilização desta sigla em português, como ocorreu com GIS > SIG. Ou seja, ambos os conceitos estão plenamente consolidados na literatura de sensoriamento remoto em língua portuguesa, mas GIS foi traduzida para SIG enquanto GPS se manteve na forma original. Vejamos a seguir, no Quadro 3, as siglas e acrônimos acompanhadas das traduções que selecionamos com a ajuda dos *corpora* comparáveis:

LWIR	Long-Wavelength Infrared	LWIR	Infravermelho de Ondas Longas
NIR	Near Infrared	IVP	Infravermelho Próximo
SWIR	Short-Wavelength Infrared	SWIR	Infravermelho de Ondas Curtas
UAV	Unmanned Aerial Vehicle	VANT	Veículo Aéreo Não-Tripulado

BRDF	Bidirectional Reflectance Distribution Function	FDRB	Função Distribuição de Reflectância Bidirecional
BRF	Bidirectional Reflectance Factors	FRB	Fator de Reflectância Bidirecional
CARI	Chlorophyll Absorption in Reflectance Index	CARI	Índice de Reflectância para Absorção pela Clorofila
TCARI	Transformed Absorption in Reflectance Index	TCARI	Índice de Reflectância para Absorção pela Clorofila Modificado
TIR	Thermal Infrared	TIR	Infravermelho Termal
TIRS	Thermal Infrared Sensor	-	Sensores de Infravermelho Termal
EVI	Enhanced Vegetation Index	EVI	Índice de Vegetação Realçado
NDVI	Normalized Difference Vegetation Index	NDVI	Índice de Vegetação da Diferença Normalizada
NVI	New Vegetation Index	NVI	Novo Índice de Vegetação
OSAVI	Optimized Soil-Adjusted Vegetation Index	IAVAS	Índice de Vegetação Ajustado a Solos Otimizado
SAVI	Soil-Adjusted Vegetation Index	IVAS	Índice de Vegetação Ajustado ao Solo
TVI	Transformed Vegetation Index	TVI	Índice de Vegetação Transformado
SAR	Synthetic Aperture Radar	SAR	Radar de Abertura Sintética
DEM	Digital Elevation Model	MDE	Modelo Digital de Elevação
DOQQ	Digital Orthophoto Quarter Quad	-	ortofotos digitais
DSM	Digital Surface Model	MDS	Modelo Digital de Superfície
DTM	Digital Terrain Model	MDT	Modelo Digital de Terreno
SST	Sea-Surface Temperature	SST	Temperatura da Superfície do Mar
VNIR	Visible and Near Infrared	VNIR	Visível e Infravermelho Próximo
GIS	Geographic Information System	SIG	Sistemas de Informação Geográfica
GPS	Global Positioning System	GPS	Sistema de Posicionamento Global

Quadro 3. Acrônimos e siglas acompanhados dos equivalentes em língua portuguesa.

Para Tagnin (2007), a pesquisa com *corpus* não se restringe a confirmar ou não uma hipótese levantada, mas, no decorrer desta busca pela confirmação, outros problemas e caminhos são descobertos, revistos, repensados. Um *corpus* não responde a todas as perguntas nem tampouco contém todas as informações sobre a língua, mas, por meio dele, ao pesquisarmos um termo, acabamos por detectar fenômenos que não poderiam ser percebidos de outra maneira.

Elaboramos um gráfico a partir das informações analisadas acima para colocar de maneira mais visual a situação da tradução dessas colocações. No decorrer da análise, percebemos que grande parte das colocações tinha seus conceitos traduzidos, mas mantinha as siglas em inglês; enquanto a outra parte, um pouco menor, tinha não só o conceito, mas também as siglas, traduzidos. Foram apenas dois os casos que não se encaixaram em nenhuma dessas duas categorias. Vejamos o Gráfico 1, abaixo:

### Tradução de siglas e acrônimos em sensoriamento remoto

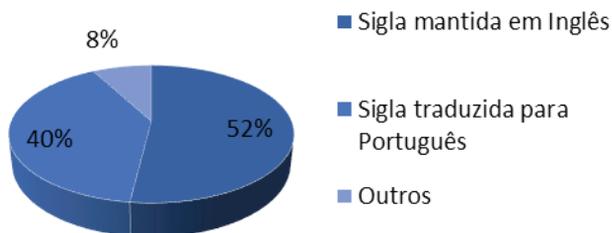


Gráfico 1. Percentual de tradução de acrônimos e siglas em sensoriamento remoto.

Com base no gráfico, vemos a predominância de colocações que tiveram seus conceitos traduzidos, mas mantiveram suas siglas em inglês, como é o caso de treze delas: LWIR, SWIR, CARI, TCARI, TIR, EVI, NDVI, NVI, TVI, SAR, SST, VNIR e GPS. E, do outro lado, temos o caso das colocações que foram plenamente traduzidas (conceito e sigla) para português, como: NIR, UAV, BRDF, BRF, OSAVI, SAVI, DEM, DSM, DTM e GIS, num total de dez. De posse desses resultados, perguntamos: como o profissional que trabalha com a área de sensoriamento remoto deve lidar com as siglas e acrô-

nimos da área? Como saber quais colocações são “traduzíveis” por completo e quais devem ser mantidas em partes como no original? Como evitar ruídos na comunicação entre os pares e garantir que a troca de informações seja proficiente? Por que o dicionário técnico muitas vezes não auxilia nessas traduções?

## Conclusões

Esta análise levantou mais questões do que ofereceu respostas, todavia, acreditamos que os estudos com base em *corpus* podem contribuir para facilitar a atuação do profissional que lida de alguma maneira com a linguagem do sensoriamento remoto. Alves e Tagnin (2011, p. 192) afirmam que grande parte das áreas técnicas carece de obras de referência bilíngue confiáveis. Muito do material que há disponível atualmente está ou desatualizado ou foi compilado sem critérios científicos. Um material para consulta impresso está sempre desatualizado, porque é praticamente impossível manter um dicionário ou glossário impresso atualizado com a velocidade em que se dão as inovações nos campos técnicos e tecnológicos. O tempo gasto para confecção somado ao tempo de revisão, impressão, publicação já bastam para que novos conceitos, técnicas, produtos tenham sido criados. Já no caso do material compilado sem critérios científicos, pode-se observar que “a maioria dessas obras em áreas de especialidade foi compilada visando principalmente a vocábulos monossêmicos, ignorando, muitas vezes, as chamadas colocações e, principalmente, as fraseologias mais longas” (ALVES & TAGNIN, 2011, p. 192). Quando se trata de traduzir uma colocação como uma sigla ou acrônimo, por exemplo, um glossário monossêmico pouco colabora nessa tarefa. Além de o usuário precisar saber como traduzir, seria também necessário descobrir se é mais comum traduzir ou manter tal colocação de acordo com o uso daquela comunidade. Um trabalho com *corpus* pode resultar num glossário bilíngue bem mais confiável e de melhor qualidade por extrair suas informações de textos autênticos.

O sensoriamento remoto é uma área predominantemente desenvolvida no exterior. Países como Estados Unidos, China, Alemanha e Rússia, por exemplo, têm avançados centros de pesquisa e desenvolvimento dessas técnicas para serem aplicadas em prol do bem-estar e defesa do ser humano. No Brasil, por ser um país onde esta disciplina é considerada emergente, a fixidez

lexical ainda se encontra no início. Para Pavel (2003), a maioria dos grupos fixos de palavras deriva das disciplinas-mãe, além disso, notamos não haver uma padronização em suas utilizações.

Em razão dessa não padronização, este artigo teve como objetivo mostrar as dificuldades que tradutores ou alunos podem encontrar ao ter de lidar com acrônimos e siglas na área de sensoriamento remoto. Ao mesmo tempo, tentamos mostrar como um *corpus* comparável pode ser útil nessa tarefa. O acesso aos *corpora* comparáveis nos indicou siglas que usualmente não são traduzidas pela comunidade, bem como nos apontou para as possíveis traduções já consagradas. Um dos maiores benefícios de se trabalhar com *corpora* é que estes são uma fonte rica de dados que podem ser constantemente atualizados, além de fornecer exemplos autênticos de uso, o que confere segurança ao trabalho do tradutor na hora da escolha de qual termo empregar.

## Referências

- ALVES, F.; TAGNIN, S.E.O. Corpora e ensino de tradução: o papel do auto-monitoramento e da conscientização cognitivo-discursiva no processo de aprendizagem de tradutores novatos. In: VIANA, V. & TAGNIN, S.E.O. (Orgs.). *Corpora no ensino de línguas estrangeiras*. São Paulo: HUB Editorial, 2011.
- ANDERSON, W.; CORBETT, J. *Exploring English with online corpora: an introduction*. London: Palgrave Macmillan, 2009, p. 45-66.
- BAKER, M. Corpora in translation studies: an overview and some suggestions for future research. *Target*, Amsterdam, v. 7. n. 2. 1995.
- BERBER SARDINHA, T. *Linguística de Corpus*. Barueri, SP: Manole, 2004.
- British National Corpus* (BNC). Disponível em <<http://www.natcorp.ox.ac.uk/>>. Acesso em 25 jul. 2012.
- CAMARGO, D.C. *Metodologia de pesquisa em tradução e linguística de corpus*. V. 1. 65 p. São Paulo/ São José do Rio Preto: Cultura Acadêmica/ Laboratório Editorial do IBILCE, UNESP, 2007.
- CARNEADO MORÉ, Z. Consideraciones sobre la fraseografía. In: CARNEADO MORÉ, Z.; TRISTÁ PEREZ, A.M. *Estudios de Fraseología*. Havana: Academia de Ciências de Cuba/Instituto de Literatura y Lingüística, 1983, p. 39-46.

COWIE, A.P. Introduction. In: \_\_\_\_\_. (Ed.). *Phraseology: theory, analysis, and applications*. Oxford University Press, 1998, p. 1-19.

CROITORU, E.; DUMITRAȘCU, A.M. Collocations and Colligations in Translation and Interpreting. In: *Annales Universitatis Apulensis*, tom 2, Alba Iulia, 2006.

FIRTH, J.R. Modes of Meaning. In: \_\_\_\_\_. (Ed.). *Papers in Linguistics – 1934-1951*. Oxford: Oxford University Press, 1957.

JENSEN, J.R. *Sensoriamento remoto do ambiente: uma perspectiva em recursos terrestres*. Tradução: José Carlos Neves Epiphânio *et al.* São José dos Campos, SP: Parêntese, 2009.

MOON, R. Dictionaries and collocations. In: GRANGER, S.; MEUNIER, F. (Eds.). *Phraseology: an interdisciplinary perspective*. Amsterdam: John Benjamins, 2008, p. 313-336.

NADIN, O.L. A variação denominativa em terminologia: a problemática das siglas. In: MURAKAWA, C.A.A.; NADIN, O.L. (Org.) *Terminologia: uma ciência interdisciplinar*. São Paulo: Cultura Acadêmica, 2013.

NOVO, E.M.L.M. *et al.* Técnicas avançadas de sensoriamento remoto aplicadas ao estudo de mudanças climáticas e ao funcionamento dos ecossistemas amazônicos. *Revista Acta Amazônica* vol. 35, [2] 2005, p. 259-272.

ORENHA-OTTAIANO, A. *Unidades fraseológicas especializadas: colocações e colocações estendidas em contratos sociais e estatutos sociais traduzidos no modo juramentado e não-juramentado*. 2009. 282 f. Tese (Doutorado em Estudos Linguísticos) - IBILCE, Universidade Estadual Paulista, São José do Rio Preto, 2009.

PAVEL, S. A fraseologia na língua de especialidade. Metodologia de registro nos vocabulários terminológicos. In: FAULSTICH, E.; ABREU, S.P. *Linguística Aplicada à Terminologia e à Lexicologia*: Cooperação Internacional: Brasil e Canadá. Porto Alegre: Ed. da UFRGS, 2003, p. 99-132.

*Projeto Lácio-Web* - Compilação de Corpus do Português do Brasil e Implementação de Ferramentas para Análises Linguísticas. Disponível em <<http://www.nilc.icmc.usp.br/lacioweb/index.htm>>. Acesso em 25 jul. 2012.

SCOTT, M. *WordSmith Tools version 6*. Liverpool: Lexical Analysis Software, 2012.

SELPER – Sociedad de Especialistas Latinoamericanos en Percepción Remota. *Dicionário SELPER – Sensoriamento Remoto*. [S. L.] 1989.

TAGNIN, S.E.O. A identificação de equivalentes tradutórios em corpora comparáveis. In: I Congresso Internacional da ABRAPUI, 2007. *Anais do I Congresso Internacional da ABRAPUI*, 2007.

TOGNINI-BONELLI, E. Working with corpora across languages. In: \_\_\_\_\_. *Corpus Linguistics at work*. Amsterdam/Atlanta, John Benjamins, 2001.

## **INVESTIGATING THE TRANSLATION OF ACRONYMS IN THE AREA OF REMOTE SENSING: A PROPOSAL FROM THE STANDPOINT OF PHRASEOLOGY**

### **ABSTRACT**

The aim is to analyze a corpus of remote sensing in order to identify acronyms in English and then search for their equivalents in Portuguese. The research is based on the approach of Corpus-Based Translation Studies (BAKER, 1995), Corpus Linguistics (BERBER SARDINHA, 2004), and Phraseology (PAVEL, 2003). The program WordSmith Tools version 6.0 is used. The results show that there is no standardization in these translations.

**KEYWORDS:** acronyms; corpus-based translation studies; remote sensing.

Recebido em: 05/08/2013

Aprovado em: 15/10/2013