

# A TERMINOLOGIA ACADÊMICA DA ODONTOLOGIA EM LÍNGUA INGLESA: ANÁLISE ESTATÍSTICO-LEXICAL DE ARTIGOS ESCRITOS POR NATIVOS E POR BRASILEIROS

## Contextualização da pesquisa e revisão curta da literatura

A expressão *publish or perish*, indiretamente, tem impacto nos trabalhos escritos em língua inglesa por brasileiros, uma vez que os artigos devem demonstrar habilidade linguística na escrita acadêmica. Dentre as habilidades, encontra-se a adequação no que tange ao uso da terminologia acadêmica comumente encontrada em determinada área do conhecimento. Dessa forma, a utilização da terminologia acadêmica nos textos técnico-científicos em língua inglesa configura-se como um suporte na produção e divulgação de conhecimento nas mais diversas áreas de especialidade.

Definimos a terminologia acadêmica como sendo um conjunto de substantivos, verbos, adjetivos e advérbios típicos da comunicação acadêmica. Destarte, a terminologia acadêmica não é aquela entendida como “pertencente” a um domínio de especialidade específico, estanque. Pelo contrário, a terminologia acadêmica, é aquela que está presente em diversos domínios e não acaba sendo identificada *a priori* como tal. Outrossim, segundo Barbosa (2009, p. 39) “podemos detectar termos/vocábulo que se situam na interface entre o discurso científico e o discurso banal, como, por exemplo, o termo/vocábulo câncer”. A autora indica que esse tipo de termo garante a comunicação entre especialistas de uma área, entre leigos, entre os primeiros e os segundos. “Por conseguinte, entre o mais alto grau de cientificidade e o mais alto grau de banalização, existe sempre um subconjunto que tem dupla natureza, a de termo e a de vocábulo” (BARBOSA, 2009, p. 39).

Tal como Pavel e Nolet (2002, p. xvii), entendemos a terminologia como um conjunto de palavras técnicas pertencentes a uma ciência, uma arte, um autor ou um grupo social, como, por exemplo, a terminologia da medicina ou a terminologia usada pelos especialistas em computação.

Conforme Barros (2006, p. 24), “a terminologia pode colaborar na elaboração de estratégias e de instrumentos de aprendizado do vocabulário especializado, contribuindo para a melhoria do ensino e para o sucesso escolar”. Até certo ponto, ao servir como insumo na elaboração de material de divulgação acadêmica e técnico-científica, a terminologia acadêmica proporciona traços específicos ao texto científico, fazendo com que seja mais parecido com outros textos do mesmo gênero.

No que diz respeito à Terminologia, a fundamentação teórica de nossa pesquisa contou com a Teoria Comunicativa da Terminologia (TCT), proposta por Cabré (1993, 2007). Quanto à Linguística de Corpus, os autores utilizados foram Biber, Conrad e Reppen (1998), Sinclair (2004) e Berber Sardinha (2004) – e em relação à Estatística Lexical, foram, principalmente, Muller (1977) e Oakes (1998).

## Objetivo do trabalho

Levando em consideração o exposto acima, o objetivo da pesquisa é identificar as diferenças de utilização da terminologia acadêmica (Coxhead, 2000), em artigos escritos em língua inglesa por brasileiros e por nativos, no domínio da Odontologia.

## Metodologia

A primeira etapa de nossa pesquisa foi a constituição de dois corpora de trabalho, sendo o primeiro formado por 392 artigos acadêmicos escritos em língua inglesa e publicados na revista *Brazilian Dental Journal*, disponíveis na base Scielo, apresentando 1.170.557 *tokens* e 42.374 *types*. Apesar de, na maioria das vezes, os artigos terem sido escritos por brasileiros, é possível que tenham sido revisados e/ou escritos com o auxílio de falantes nativos de inglês. A título de comparação, constituímos um corpus com dimensão semelhante no domínio da Odontologia, apresentando 1.224.807 *tokens* e 32.036 *types*. Esse corpus contém exclusivamente material escrito em língua inglesa por falantes nativos.

No tocante ao auxílio de *software*, a ferramenta linguístico-estatística *WordSmith Tools*, versão 5, foi utilizada para processar o conteúdo textual dos dois corpora. O recurso *WordList* auxiliou na busca pela terminologia acadêmica selecionada para o estudo. Todo o conteúdo textual estava em língua inglesa e foi digitalizado para que pudesse ser posteriormente analisado pelo programa computacional *WordSmith Tools*. Para facilitar a manipulação dos dados pelo *software*, todos os textos tiveram que ser convertidos para o formato “texto simples”, com a extensão “.txt”. Nessa conversão, as fórmulas, as tabelas, as figuras e os gráficos não puderam ser aproveitados. As legendas, no entanto, foram aproveitadas e constam no corpus.

Com relação à terminologia acadêmica, em nosso trabalho, utilizaremos 60 palavras da Academic Word List, AWL.

Em nosso estudo, as ocorrências das palavras nos dois corpora de estudo foram analisadas pelo teste *log-likelihood*, também conhecido como razão de log-verossimilhança (ou  $G^2$ ). Esse teste permite verificar se os valores de frequência observados numa determinada amostra correspondem ao valor teoricamente esperado em outra amostra, partindo do princípio de que as frequências deveriam ser as mesmas. O teste *log-likelihood* pode ser usado para comparar, por exemplo, as ocorrências de determinados itens lexicais pertencentes a corpora de dimensões distintas. Em função dos procedimentos estatísticos adotados pelo teste, os resultados obtidos são facilmente comparáveis.

A fim de agilizar os cálculos estatísticos relativos ao do teste *log-likelihood*, é possível utilizar uma planilha do aplicativo MS Excel previamente configurada, desenvolvida no *University Centre for Computer Corpus Research on Language*, ligado à Universidade de Lancaster, disponível gratuitamente para download no seguinte endereço eletrônico: <http://ucrel.lancs.ac.uk/llwizard.html>.

## Resultados

Os valores esperados das frequências (*expected frequencies*) são aqueles que deveriam ser registrados caso as ocorrências fossem proporcionalmente iguais. Na coluna F, são reportados os valores obtidos por meio do cálculo *log-likelihood*.

Atribuímos valores negativos aos valores da coluna F sempre que a frequência da palavra indicada na coluna B apresentava-se maior que na coluna C. Em nosso trabalho, valores negativos indicam menor utilização da palavra no corpus de artigos escritos em língua inglesa por brasileiros, ao passo que valores positivos representam uma utilização acima do normal.

Os valores obtidos pelo cálculo *log-likelihood* que se encontram no intervalo compreendido entre 0 e  $\pm 3,84$  não apresentam diferenças de uso significativas. A terminologia,

nesse caso, está sendo utilizada de acordo com os parâmetros adequados comumente encontrados em textos acadêmicos originalmente escritos em língua inglesa.

**Tabela 1 – As 14 palavras acadêmicas com frequência normal no corpus brasileiro de Odontologia**

evident	-2.9	formula	-0.38
role	-2.22	export	-0.10
legislate	-1.34	authority	-0.06
policy	-0.81	distribute	-0.05
environment	-0.7	establish	0.58
legal	-0.63	income	0.86
context	-0.4	sector	1.47

Na sequência, na tabela 2, listamos agora as palavras acadêmicas que obtiveram valores negativos:

**Tabela 2 – As 37 palavras acadêmicas com frequência abaixo do esperado no corpus brasileiro de Odontologia**

respond	-978.96	vary	-53.66
require	-443.31	economy	-37.68
structure	-258.72	research	-37.42
function	-232.83	occur	-36.36
indicate	-180.46	define	-36.26
proceed	-168.13	interpret	-28.30
individual	-165.89	benefit	-23.13
create	-143.97	assess	-20.23
available	-139.61	consist	-19.47
principle	-132.8	theory	-14.91
area	-131.09	contract	-13.99
involve	-120.19	derive	-13.42
factor	-113.94	finance	-8.97
approach	-97.58	labour	-7.48
estimate	-76.42	constitute	-7.39
concept	-69.18	issue	-7.02
major	-56.73	assume	-5.26
specific	-56.36	identify	-4.26
section	-55.06		

Com efeito, a observação da Tabela 2 permite constatar que o corpus de Odontologia (corpus 2), ao servir de referência de comparação, revela que houve uma grande disparidade em relação ao emprego de *respond* (-978.96) e *require* (-443.31). Um número negativo tão expressivo revela que a palavra acadêmica em questão não faz parte do leque vocabular

brasileiro. Grande parte dessas palavras possui origem latina e não deveria deixar de ser inserida nos artigos. O valor identificado permite inferir que essa subutilização pode ter efeitos negativos no texto acadêmico escrito por brasileiros.

Na tabela 3, apresentamos agora as palavras que tiveram um uso excessivo:

**Tabela 3 – As 9 palavras acadêmicas com frequência acima do esperado no corpus brasileiro de Odontologia**

similar	4.1	period	61.25
process	10.96	method	258.26
source	14.34	analyse	479.79
percent	16.96	significant	629.03
data	55.08		

Chama a atenção o uso acentuado de *significant* (629,03). Aqui, além de ser utilizada exageradamente, a palavra sugere que pode haver problemas de estilo nos textos. É provável que haja muita repetição de *significant* por desconhecimento de outros itens lexicais equivalentes - *meaningful*, *crucial*, *vital* e *important*. Até certo ponto, os artigos escritos por brasileiros acabam se construindo por palavras não-comumente tão usadas no domínio da Odontologia. Revestem-se, por assim dizer, de uma aparência distinta dos artigos escritos em língua inglesa no exterior.

## Conclusão

Ao servir como insumo na elaboração de material de divulgação acadêmica e técnico-científica, a terminologia acadêmica proporciona traços específicos ao texto científico, fazendo com que seja mais parecido com outros textos do mesmo gênero.

A comparação dos textos escritos por brasileiros com os textos escritos por nativos mostrou diferenças substanciais no uso da terminologia acadêmica, mais precisamente das palavras da AWL constantes na primeira sublista de Coxhead (2000). Eliminar ou reduzir ao mínimo essas diferenças permitiria que os textos científicos produzidos por brasileiros ganhassem uma aceitação melhor perante a comunidade científica internacional.

No que diz respeito às principais diferenças observadas no emprego da terminologia acadêmica na produção textual brasileira, das 60 palavras analisadas, 15% apresentaram uso excessivo, 23,33% foram empregadas normalmente e 61,66% deixaram de ser utilizadas com a frequência esperada em um texto da área de Odontologia. É justamente nesse quesito que a diferença entre o texto do nativo e do brasileiro mostra-se mais pronunciada.

Um caminho indicado por nosso trabalho é de utilizar como recurso para o aprendizado dessa terminologia a AWL, que fornece uma referência linguística pedagógica valiosa para que outros estudos se aprofundem no caráter dos textos acadêmicos. Os resultados por ora apresentados podem servir para que professores de língua inglesa possam direcionar estratégias de ensino específicas. Da mesma forma, pesquisadores das mais variadas áreas podem avaliar a problemática discutida aqui e repensar sua atividade-fim, a redação acadêmica em língua inglesa.

## Referências

- BARBOSA, Maria Aparecida. Terminologia aplicada: percursos interdisciplinares. *Polifonia*. Cuiabá: Edufmt, n. 17, p. 29-44, 2009.
- BARROS. Lídia Almeida. Aspectos epistemológicos e perspectivas científicas da terminologia. *Ciência e Cultura*, v. 58, n. 2, São Paulo, p. 22-26, abr./jun. 2006.
- BERBER SARDINHA, Tony. *Linguística de Corpus*. Barueri: Manole, 2004.
- BIBER, Douglas; CONRAD, Susan.; REPPEN, Randi. *Corpus linguistics: investigating language structure and use*, Cambridge : Cambridge University Press, 1998.
- CABRÉ, M. T. La terminologia: teoria, metodologia, aplicaciones. Barcelona: Editorial Antartida/Empuries, 1993.
- \_\_\_\_\_. Constituir um corpus de especialidad: condiciones y posibilidades. (2007). Disponível em: <[www.upf.edu/pdi/df/teresa.cabre/docums/ca07arra.pdf](http://www.upf.edu/pdi/df/teresa.cabre/docums/ca07arra.pdf)>. Acesso em: 28 mar. 2012.
- CANTOS GÓMEZ, Pascual. Do we need statistics when we have linguistics? *DELTA*, v.18, n.2, p. 233-271, 2002
- COXHEAD, Averil. A new academic word list. *TESOL Quarterly*, v. 34, n. 2, p. 213-238, 2000.
- MULLER, Charles. *Principes et méthodes de statistique lexicale*. Paris: Hachette, 1977.
- NATION, Paul. *Como estruturar o aprendizado de vocabulário*. Tradução de Cristiane Arruda. São Paulo: Special Book Services, 2003.
- OAKES, Michael. P. *Statistics for corpus linguistics*. Edinburgh: Edinburgh University Press, 1998.
- PAVEL, Silvia; NOLET, Diane. *Manual de terminologia*. Tradução Enilde Faulstich. Québec: Travaux Publics et Services Gouvernementaux, 2002.
- SINCLAIR, John. Corpus and text: basic principles. In: WYNNE, Martin. (Ed.). *Developing linguistic corpora: a guide to good practice*. 2004. Disponível em: <<http://ota.ox.ac.uk/documents/creating/dlc/chapter1.htm>>. Acesso em: 7 fev. 2012.