
Nominalizações deverbais na escrita científica e indexação temática: Uma análise léxico-morfológica (1)

Deverbal nominalizations in the scientific writing and the subject indexing: a lexical and morphological analysis.

Vânia Lisbôa da Silveira Guedes

Faculdade de Administração e Ciências Contábeis da Universidade Federal do Rio de Janeiro.
Endereço: Rua Alberto Maranhão, 99, 3º andar, apto. 303 - Rio de Janeiro – RJ Brasil CEP 21940-490.
E-mail: vanialisboa@facc.ufrj.br

Resumen

Este estudo analisa o discurso científico sobre Vinicultura, em português do Brasil, investigando a recorrência de nominalizações deverbais no léxico do gênero artigo de periódico e suas funções de índices temáticos. Objetiva desenvolver uma análise linguística léxico-morfológica do discurso sobre Vinicultura e, assim, contribuir para o desenvolvimento de modelos de sistemas semiautomáticos de indexação. Considera como referencial teórico a análise de discurso, a teoria de gêneros, a teoria lexical e a indexação fundamentada na Bibliometria. A amostra é composta por quatro artigos sobre Vinicultura, publicados na *Scientific Electronic Library Online*. Os artigos foram processados pelo *Software RankWords 2.0* e, para cada artigo, foi produzida uma lista de frequência de palavras. Verifica-se, em cada lista, a aplicação das leis de Zipf e do Ponto de Transição de Goffman, assim como a produtividade e frequência relativa de nominalizações em *-ção*, *-mento*, *-ncia*, *-agem* e *-da*, e os graus distintos de relevância de suas funções de índice temáticos. Para a confrontação dos dados, o método é aplicado a dois artigos sobre Economia. Os resultados obtidos revelam a adequação do método, sobretudo, aos artigos sobre Economia e evidenciam a importância das abordagens teóricas e descritivas à nominalização para a indexação da informação científica.

Palavras-chave: Análise de gêneros discursivos. Leis de Zipf. Ponto de Transição de Goffman. Discurso Científico. Nominalização deverbal. Indexação temática.

1. Introdução

Estudos linguísticos e bibliométricos acerca da função de índice temático da nominalização deverbal na escrita científica vêm revelando a importância dessas abordagens teóricas e metodológicas para a identificação do conteúdo informativo de artigos científicos, na Ciência da Informação (CI).

Abstracts

This study analyses the scientific discourse about the Viniculture in Brazilian Portuguese language, in order to investigate the productivity and recurrence of deverbal nominalizations present in the lexicon of papers and their index role in the subject indexing. The aim is the linguistic and statistical analysis of the discourse about Viniculture and the establishment of patterns to develop subject indexing semiautomatic systems. The theoretical framework used was the discourse analysis, genre theory, lexical theory and indexing based on Bibliometrics. To identify those lexico-morphological features in such terminology, four papers about the wine industry were selected which were published in the *Scientific Electronic Library Online*. These papers were processed by *Software RankWords 2.0*. For each paper, a list of words frequencies was obtained. In each list it was investigated the Zipf's Laws and Goffman Transition Point adequacy, the productivity of deverbal nominalization in *-ção*, *-mento*, *-ncia*, *-agem*, *-da* and the relative frequencies of these words and the different degrees of relevance of their functions as an index term for the subject indexing. Aiming data confrontation, the method is applied into two text about Economics. The results show the importance of studies about nominalization for the subject indexing, in the Information Science.

Keywords: Discourse genre Analysis. Zipf's Laws. Goffman's Transition Point. Scientific Discourse. Deverbal Nominalization. Subject Indexing.

Este estudo propõe o desenvolvimento de uma análise bibliométrica e léxico-morfológica do discurso científico, na área de Vinicultura, com o objetivo de estimar a produtividade e a frequência de uso de nominalizações deverbais na escrita científica e suas funções de índices temático. Nesse sentido, a pesquisa baseia-se em estudos bibliométricos, voltados para a indexação

ão temática na CI, na análise do discurso e na teoria de gêneros na área de Sociolinguística.

Para o desenvolvimento da pesquisa impõem-se os seguintes postulados: (i) o padrão de nominalização $[X]v \rightarrow [[X]v -\text{ção}]N$ é predominante na escrita científica de artigos sobre Vinicultura; (ii) a frequência do item e processos léxico-morfológicos de nominalizações deverbais mantêm relação com o gênero científico; (iii) do ponto de vista bibliométrico, as nominalizações deverbais com função de índices temáticos ocorrem na Região de Transição de Goffman.

2. Reflexões sobre Morfologia, Léxico e Análise de Domínio na CI

A Morfologia, segundo Aronoff & Anshen (2001), trata da estrutura interna das palavras potencialmente compostas de uma língua e encontra as bases sobre as quais opera no léxico. Em Basílio (2011) o léxico é concebido como sistema de formas simbólicas que evocam significados, utilizadas na construção de enunciados, para fins de comunicação. Aronoff & Anshen (2001) entendem a produtividade léxico-morfológica como a extensão na qual determinado afixo é provavelmente utilizado na produção de palavras. Em Aronoff (1976), a noção de produtividade lexical é considerada basilar no âmbito da Morfologia Derivacional. O mesmo autor (ibidem) defende que uma nova palavra é gerada pela aplicação de uma Regra de Formação de Palavra (RFP) a um item pré-existente no léxico e introduz noções de graus de produtividade e de bloqueio de RFP. Bybee (2007) destaca a frequência do item como vetor facilitador para o seu acesso, diretamente relacionado ao grau de força lexical das relações morfológicas do item. Basílio (2007, p.78) define nominalização como “conjunto de processos que formam substantivos a partir de adjetivos e, sobretudo, de verbos”, que desempenham funções textuais e semânticas, sendo que a semântica da nominalização é a de denominação.

Neste estudo, defende-se a importância da aproximação do estudo de nominalizações deverbais à análise bibliométrica, enunciada por Zipf e Goffman, para a identificação da faixa de ocorrência de palavras significativas na escrita de artigos científicos e seleção de termos de indexação. Bybee & Hopper (2001) ao discutirem a relação da frequência com a emergência de estruturas linguística mencionam que Zipf foi pioneiro nos anos 30. “The Psycho-biology of Language” editado em 1935 baseia-se no Princípio do Menor Esforço e antecipa muitos temas recentes de investigação sobre a relação entre frequência e importância de estruturas

linguísticas. Para Basílio (1979, p. 1), os “[...] estudos lexicais constituem um caminho promissor para o desenvolvimento de pesquisas interdisciplinares em Linguística e Ciência da Informação”.

Na Organização do Conhecimento na CI, Hjørland (2002) indica onze abordagens de Análise de Domínios discursivos, dentre as quais são de interesse desse estudo : (1) pesquisas sobre indexação e recuperação da informação em domínios especializados; (2) produção e interpretação de estudos bibliométricos; (3) estudos dos gêneros de documentos em domínios do conhecimento e (4) estudos terminológicos, linguagens para fins especiais e análise do discurso em áreas específicas do conhecimento. Hjørland (2004) menciona que os estudos linguísticos na Análise de Domínios oferecem uma perspectiva teórica e prática relevante para o desenvolvimento do arcabouço teórico e prático da CI.

Na aplicação de abordagens teóricas e metodológicas das metrias à indexação, são analisados *corpora* textuais, utilizando-se conhecimentos linguísticos e bibliométricos na construção de referenciais, com o objetivo de contribuir para o desenvolvimento de sistemas semiautomáticos de indexação. Os pressupostos teóricos das Leis de Zipf e Ponto de Transição (T) de Goffman aproximam a frequência de uso da palavra e seu potencial de representação temática de um *corpus* em análise. Nesse sentido, Hrebicek (2002) alerta que, do ponto de vista estatístico, nenhuma objeção razoável, manifestando-se contra a redução de palavras e de *corpora*, é feita às Leis de Zipf. Do ponto de vista linguístico, entretanto, existe um *gap* entre unidades lexicais e *corpora* que indica a ausência de fenômenos relevantes ligados às estruturas léxico-morfológicas. As propriedades semânticas reais das palavras podem ser observadas quando são testadas como unidades lexicais em textos.

3. Metodologia

A linha teórico-metodológica baseia-se na Sociolinguística e na Bibliometria. Como amostra, optou-se por um *corpus*, composto por quatro artigos sobre Vinicultura e, visando à confrontação dos dados, dois artigos sobre Economia. A seleção dos textos segue os seguintes critérios: (1) quatro artigos científicos sobre Vinicultura e dois sobre Economia, em língua portuguesa, indexados na *Scielo Brazil Library Online* e publicados no período de 1998 a 2008; (2) artigos com cerca de 2.000 a 6.000 palavras. (Quadro 1, em anexo 1).

3.1 Procedimento de análise

- (a) seleção de quatro artigos sobre Vinicultura e dois sobre Economia;
- (b) contagem das palavras pela utilização do *Software RankWords2.0*;
- (c) ordenação das palavras operando o *software* mencionado em (b) que produziu três colunas: as palavras, as frequências em ordem decrescente de ocorrência e o *rank* das palavras;
- (d) cálculos matemáticos e estatísticos a partir do transporte dos resultados em planilhas utilizando-se o *Software Microsoft Office Excel 2003*, para a projeção dos seguintes produtos: total de palavras, total de palavras distintas, palavras com frequência 1 e *ranking* das nominalizações deverbais;
- (e) identificação dos padrões de nominalizações em *-ção, -mento, -ncia, -agem e -da*;
- (f) estimativa dos graus de produtividade e frequência de uso das nominalizações;
- (g) identificação do Ponto T de Goffman e delimitação da Região de Transição de Goffman;
- (h) investigação das nominalizações que se situam na Região de Transição de Goffman;
- (j) confrontação dos dados observados nos textos sobre Vinicultura e sobre Economia.

3.2 Bibliometria: Leis de Zipf e Ponto de Transição de Goffman

O termo Bibliometria, introduzido por Pritchard (1969), designa a área do conhecimento que utiliza métodos matemáticos e estatísticos para análise de processos de comunicação escrita (Pao, 1989). Okubo (1997) afirma que as abordagens bibliométricas são baseadas na noção de que a produção de conhecimento é a essência da pesquisa científica e, nesse sentido, a literatura científica é a manifestação constituinte do conhecimento. Okubo esclarece que os indicadores bibliométricos podem ser usados para a descrição de questões que surgem no mundo da Ciência (Guedes, 2012).

No contexto das principais leis bibliométricas, George Kingsley Zipf observou certas regularidades na estrutura linguística de textos e enunciou duas leis (Zipf 1945, 1949). A primeira lei está relacionada as palavras de alta frequência e a segunda as de baixa frequência. A primeira lei de Zipf enuncia que “o produto da ordem de série (*r*) de uma palavra pela sua frequência (*f*) é aproximadamente constante (*C*):

$$(1) r \times f = C$$

A segunda lei enuncia que, “em um texto, várias palavras de baixa frequência de ocorrência têm a mesma frequência”. A modificação proposta por Booth para a segunda lei, conhecida como lei Zipf-Booth, é representada matematicamente da seguinte forma:

$$(2) I_1 = n(n + 1)/2$$

Goffman, citado por Pao (1978), defende que essas leis possibilitam a indicação do ponto de transição de palavras de baixa frequência para as de alta frequência, onde teoricamente existe uma região que concentra palavras com alta carga semântica. Para o cálculo do Ponto T, a expressão da Segunda Lei de Zipf teria que fornecer o comportamento típico das palavras de alta frequência, onde o número de palavras com frequência *n* tende a 1. Substituindo-se, na expressão da Segunda Lei de Zipf-Booth, I_n por 1, obtém-se:

$$(3) n = \frac{-1 + \sqrt{1 + 8I_1}}{2}$$

Pao (1978), ao testar a hipótese de que a região de transição de Goffman produziria as palavras com alta carga semântica, delimitou a região projetando para baixo o mesmo número de palavras que ocorreu acima da frequência ligada ao Ponto T. Nessa região, foi calculado o total de palavras e excluídas as palavras sem conteúdo semântico, tais como: preposições, artigos, conjunções, pronomes e verbos auxiliares. Pao menciona que os resultados obtidos foram satisfatórios, para os textos analisados, e parecem indicar que a seleção de termos de indexação, a partir de uma lista de frequência de ocorrência de palavras, é uma perspectiva promissora para a indexação automática.

Bybee e Hopper (2001, p.1), ao discutirem a relação da frequência com a emergência de estruturas linguísticas, afirmam que George Kingsley Zipf foi o pioneiro, ao editar a obra “*The Psycho-biology of Language*”, em 1935, baseada no Princípio do Menor Esforço. Segundo os autores, Zipf antecipou muitos temas de pesquisa investigados recentemente sobre a frequência de tipo e de uso de estruturas linguísticas na fala e na escrita (Guedes, 2010).

4. Análise, Síntese e Discussão dos Resultados

No que se refere à análise bibliométrica, a aplicação da Primeira Lei de Zipf apontou desvios

da média nos seis artigos analisados. Similarmente, a aplicação da Segunda Lei de Zipf também apresentou resultados com desvios da média nos seis artigos investigados. Por exemplo, no processamento do artigo 1 (T1), a aplicação da Primeira Lei de Zipf, considerando como palavras de alta frequência as que se situam na ordem de série (r) igual ou menor que 5 ($r \leq 5$), apresentou desvios da média de 216,4, para mais, de 10,9% ($r=3$), de 46,02% ($r = 4$), de 1,66% ($r = 5$) e, para menos, de 34,38% ($r=1$), de 24,21% ($r = 2$). Já a aplicação da Segunda Lei de Zipf ($n \leq 5$), onde $n = f$, apresentou desvios, para menos, de cerca de 29,21% ($n = 2$), 52,62% ($n = 3$), 60,52% ($n = 4$) e 32,88% ($n = 5$).

Quanto à análise linguística léxico-morfológica, a síntese dos resultados obtidos na investigação da produtividade e da frequência de uso dos sufixos nominalizadores em tela, nos artigos 1,2,3,4,5 e 6, é apresentada nas tabelas 1 e 2. (Tabela 1 e 2, em anexos 2 e 3).

Os resultados sintetizados (tabelas 1 e 2) evidenciam alto grau de produtividade do padrão $[X]v \rightarrow [[X]v -\text{ção}]N$. Em contrapartida, evidenciam baixo grau de produtividade do padrão $[X]v \rightarrow [[X]v -\text{da}]N$, tendendo a zero em alguns artigos. Constatou-se que o gênero artigo científico, nas áreas em análise, prevê frequência alta de produtividade na ocorrência dos padrões de nominalizações deverbais em *-ção*, seguido de *-mento*. Em função das chances de emergência, tais formas bloqueiam e enfraquecem o processamento dos demais padrões possíveis, conforme comprovado na pesquisa. Desse modo, atesta o conceito de força lexical, postulado em Bybee (1988), pois a frequência de uso dos itens influencia o grau de produtividade dos padrões de nominalização investigados.

Cotejados os resultados da análise dos artigos sobre Vinicultura com os obtidos pela análise dos artigos sobre Economia, constatam-se diferenças significativas na relação entre os números de bases distintas e de frequências relativas de uso, associados aos sufixos nominalizadores *-ção* e *-mento*. Os textos sobre Economia apresentaram menor diferença entre as frequências relativas de uso dos sufixos nominalizadores *-ção* (162) e *-mento* (151), apesar de terem tido ocorrência relevante em relação às bases conectadas aos sufixos *-ção* (47) e *-mento* (12) na formação de nominalizações deverbais.

A aplicação da fórmula do ponto T e do mecanismo de delimitação da Região de Transição de Goffman revelou-se satisfatória. Os resultados obtidos pela observação dos itens lexicais frequentes na Região de Transição, em 100%

dos textos, indicaram a ocorrência de palavras de alta carga semântica. Ou seja, para os textos analisados, a região concentra frequências associadas a palavras de alto conteúdo semântico. Essa assertiva é reforçada pelo fato dos termos incluírem-se nos títulos e/ou nos resumos e/ou nas palavras-chave dos seis textos analisados. Contudo, os resultados obtidos no processamento dos textos 3, 4, 5 e 6 sobre Vinicultura e Economia verificaram-se ainda mais satisfatórios, principalmente na análise dos textos sobre a Economia.

Como exemplo, no processamento do artigo 1, o Ponto T de Goffman localiza-se aproximadamente na frequência 27 (27,07). Para a delimitação da região, projeta-se para baixo o mesmo número de palavras que ocorre acima da frequência 27, obtendo-se um total de 21 palavras. Nessa região, ocorrem 15 operadores sintáticos, tais como: preposições, artigos, conjunções e pronomes, que foram excluídos por não poderem ser utilizados na indexação. Assim, na Região de Transição de Goffman, foram identificadas as seguintes palavras: *vinhos*, *ácido*, *Sant*, *Ana*, *Livramento*, a sigla *pH* (Potencial Hidrognônico) e o símbolo *k*, que integram o resumo e são de alto teor semântico em relação ao artigo em análise.

Finalmente, em relação à regularidade das nominalizações mais frequentes, analisaram-se duas nominalizações em *-ção* e duas em *-mento*. Os resultados obtidos indicaram que a regularidade da regra se deve à designação de processos, atos, produtos e conceitos. No trecho, a seguir, observa-se a equivalência semântica da nominalização com o verbo correspondente, evidenciando a função semântica denominadora da forma *fermentação* que designa a visão abstrata nominal do processo de fermentar a uva.

[...] após a *fermentação* alcoólica [...],
[...] após o álcool ter fermentado a uva.

5. Conclusão

Conclui-se que, para os textos analisados, a aplicação das Leis de Zipf não se verificou satisfatoriamente. Entretanto, observa-se que os baixos *ranks* (altas frequências) estão associados às palavras de baixo conteúdo semântico como artigos, preposições, conjunções, verbos auxiliares etc. como mencionado em Rouault (1987), Eggins (2004), dentre outros.

A aplicação da fórmula do Ponto T e a delimitação da Região de Transição revelaram-se satisfatória. Verificaram-se frequências de uso de nominalizações deverbais na região, entretanto,

os resultados obtidos no processamento dos textos 3, 4, 5 e 6 verificaram-se ainda mais satisfatórios, principalmente nos textos sobre Economia (5 e 6).

Os achados de pesquisa demonstram a importância da análise léxico-morfológica da produtividade e frequência de uso de nominalizações deverbais no discurso científico, na busca de evidências empíricas, sobretudo para comprovar que as nominalizações em *-ção*, em artigos científicos, apresentam alta densidade de informação.

Constata-se ainda que os processos analisados apresentam função denominadora predominante. A formação de nominalizações resulta na constituição de rótulos que têm representação conceptual nos campos de conhecimento. Assim, os resultados confirmam a teoria de que o gênero científico é caracterizado por alto grau de informatividade, de tal modo que a produtividade de nominalizações deverbais em artigos é esperada.

Numa perspectiva léxico-morfológico, a formação de nominalização formalmente representada por $[X] v \rightarrow [[X] v -ção] N$ é predominante no campo semântico da Vinicultura, apontando caminhos mais apropriados para os processos de indexação temática em sistemas eletrônicos de linguagens controladas, na Ciência da Informação.

Finalmente, o estudo das nominalizações deverbais contribui para a adequação do método bibliométrico de indexação em análise, especialmente no que se refere à identificação de bases e sufixos de alto conteúdo informativo. Sugere-se estender esta investigação para maior número de textos e áreas de assunto, com o intuito de validação do método.

Referencias

- Aronoff, M. (1976). *Word formation in generative grammar*. Cambridge, Massachusetts: the MIT Press. (Linguistic Inquiry Monographs, 1).
- Aronoff, M.; Anshen, F. (2001). Morphology and the Lexicon: lexicalization and productivity. 2001. In: Spencer, A.; Zwicky, A. M. (Ed.) *The Handbook of Morphology*. Oxford: Wiley-Blackwell.
- Basílio, M. M. de P. (2007). *Teoria lexical*. 8. ed. revista e atualizada. São Paulo: Ática. (Série Princípios, 88).
- Basílio, M. M. de P. (1979). Interface Linguística e Ciência da Informação: potencialidade na análise de estruturas

- lexicais. In: REUNIÃO BRASILEIRA DE CIÊNCIA DA INFORMAÇÃO, 2., 1979, Rio de Janeiro. *Trabalhos apresentados...* Rio de Janeiro: IBICT.
- Booth, A. D. (1967). A "Law" of occurrences for words of low frequency. *Information and Control*, [s.l.], v. 10, n.4, p. 386-393.
- Bybee, J. L. (2007). *Frequency of use and the organization of language*. New York: Oxford University Press.
- Bybee, J. L.; Hopper, P. (2001). *Frequency and the emergence of language structure*. Amsterdam: John Benjamins.
- Eggins, S. (2004). *An introduction to systemic functional linguistics*. 2. ed. Nova York: Continuum International Publishing Group.
- Guedes, V. L. da S. (2012). Bibliometria: uma ferramenta estatística para a gestão da informação e do conhecimento, em sistemas de informação, de comunicação e de avaliação científica e tecnológica. *PontodeAcesso*, Salvador, UFBA.
- Guedes, V. L. da S. (2010). *Nominalizações deverbais em artigos científicos: uma contribuição para a análise e a indexação temática da informação*, 2010. ... f. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro.
- Hjørland, B. (2004). Domain analysis: a socio-cognitive orientation for information science research. *Bulletin of American Society of Information Science and Technology*, v. 30, n. 3. Disponível em: <http://www.asis.org/Bulletin/Feb-04/hjorland.html> Acesso em: 08/06/2010.
- Hjørland, B. (2002) Domain analysis in information science: eleven approaches - traditional as well as innovative, *Journal of Documentation*, v. 58, n. 4, p. 422-462. Disponível em <http://www.emeraldinsight.com/Insight/viewPDF.jsp?contentType=Article&Filename=html/Output/Published/EmeraldFullTextArticle/Pdf/2780580404.pdf> Acesso em: 08/06/2009.
- Hřebíček, L. (2002). Zipf's law and text, *Glottometrics*, Czech Republic, v. 3, n. 22, p. 23-38.
- Okubo, Y. (1997). Bibliometric indicators and analysis of research systems: methods and examples. Paris: OECD Publishing. Science, Technology and Industry Working Papers, 1997/01. Disponível em: [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=OCDE/GD\(97\)41&docLanguage=En](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=OCDE/GD(97)41&docLanguage=En). Acesso em: 27/03/2012
- Pao, M. L. (1978). Automatic text analysis based on transition phenomena of word occurrences, *Journal of the American Society for Information Science*, New York, v. 29, n.3, p. 121-124.
- Pritchard, A. (1969). A statistical bibliography or bibliometrics? *Journal of Documentation*, [s.l.], v.25, n.4, p. 348-349.
- Rouault, J. (1987). *Linguistique automatique: applications documentaires*. Berne: Peter Lang.
- Zipf, G.K. (1945) The meaning frequency relationship of words. *The Journal of general Psychology* 33, 251- 266.
- Zipf, G.K. (1949). *Human behavior and the Principle of Least Effort*. Cambridge, MA: Addison-Wesley.

Anexo 1

1. Rizzon, L. A.; Zanuz, M. C.; Miele, A. (1998). Evolução da acidez durante a vinificação de uvas tintas de três regiões vitícolas do Rio Grande do Sul. *Ciência e Tecnologia de Alimentos*, Campinas, v. 18, n. 2, p. 179-183. Disponível em: http://www.scielo.br/scielo.php?pid=S0101-20611998000200007&script=sci_arttext Acesso em: 27/05/2015.
2. Rizzon, L. A.; Miele, A. (2003). Avaliação da cv. Merlot para elaboração de vinho tinto. *Ciência e Tecnologia de Alimentos*, Campinas, v. 23, supl., p. 156-161. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-20612003000400029 Acesso em: 27/05/2015.
3. Rizzon, L. A.; Miele, A. (2002). Avaliação da cv. Cabernet Sauvignon para elaboração de vinho tinto. *Ciência e Tecnologia de Alimentos*, Campinas, v. 22, n. 2, p. 192-198. Disponível em: http://www.scielo.br/scielo.php?pid=S0101-20612002000200015&script=sci_arttext Acesso em: 27/05/2015.
4. Mamede, M. E. de O.; Pastore, G. M. (2004). Avaliação da produção dos compostos majoritários da fermentação de mosto da uva por leveduras isoladas da região da "Serra Gaúcha" (RS). *Ciência e Tecnologia de Alimentos*, Campinas, v. 24, n. 3, p. 453-458. Disponível em: http://www.scielo.br/scielo.php?pid=S0101-20612004000300026&script=sci_arttext Acesso em: 27/05/2015.
5. Crisostomo, V. L. (2009). Dificuldades das empresas brasileiras para financiar seus investimentos em capital físico e em inovação. *Revista Economia Contemporânea*, Rio de Janeiro, v. 13, n. 2. Disponível em: http://www.scielo.br/scielo.php?pid=S1415-98482009000200004&script=sci_abstract&tlng=pt Acesso em: 27/05/2015.
6. Vasconcelos, C. R. F.; Vasconcelos, S. P. (2005). Medidas "antidumping" e resultados colusivos: o caso do PEBDL na economia brasileira. *Nova Economia*, Belo Horizonte, v. 15, n. 3, p. 117-141. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-63512005000300005 Acesso em: 27/05/2015.

Quadro I. Amostra

Anexo 2

Sufixo	Frequência Total por Sufixo				Frequência Relativa do Sufixo / Frequência Total de Sufixos				Bases Distintas				Bases Distintas/ Total de Bases Distintas				Frequência Média das Bases Distintas			
	T1	T2	T3	T4 T5	T1	T2	T3	T4 T5	T1	T2	T3	T4 T5	T1	T2	T3	T4 T5	T1	T2	T3	T4 T5
-ção	98	112	148	157	94,23	80,57	81,76	88,70	36	47	46	33	87,80	72,30	67,64	82,50	2,72	2,38	3,21	4,75
-mento	3	11	15	16	2,88	7,91	8,28	9,03	2	9	10	3	4,87	13,84	14,70	7,50	1,50	1,22	1,50	5,33
-agem	0	6	5	1	0	4,31	4,41	0,56	0	3	4	1	0	4,61	5,88	2,50	0	2	1,25	1
-ncia	2	6	8	3	1,92	4,31	2,76	1,69	2	4	6	3	4,87	6,15	8,82	7,50	1	1,50	1,33	1
-da	1	4	5	0	0,96	2,87	2,76	0	1	2	2	0	2,43	3,07	2,94	0	1	2	2,50	0
Totais	104	139	181	177	100	100	100	100	41	65	68	40	100	100	100	100	6,22	9,10	9,79	12,08

Tabela 1. Síntese da produtividade dos sufixos nominalizadores em análise nos artigos 1, 2, 3 e 4, bem como suas frequências relativas de uso.

Sufixo	Frequência Total por Sufixo		Frequência Relativa do Sufixo/ Frequência Total de Sufixos		Bases Distintas		Bases Distintas/ Total de Bases Distintas		Frequência Média das Bases Distintas	
	T5 T2	T6 T4 T5	T5 T2	T6 T4 T5	T5	T6 T3 T4 T5	T5 T2 T3 T4	T6	T5 T2 T3 T4	T6
-ção	162	274	47,78	66,99	47	61	68,11	64,21	3,44	4,49
-mento	151	78	44,54	19,07	12	15	17,39	15,78	12,58	5,20
-agem	0	10	0	2,44	0	3	0	3,15	0	3,33
-ncia	26	20	7,66	4,88	10	12	14,49	12,63	2,60	1,66
-da	0	27	0	6,60	0	4	0	4,21	0	6,75
Totais	339	409	100	100	69	95	100	100	18,62	21,43

Tabela 2. Síntese da produtividade dos sufixos nominalizadores em análise nos artigos 5 e 6, bem como suas frequências relativas de uso.