

Universidade de São Paulo - USP
Universidade Federal de São Carlos - UFSCar
Universidade Estadual Paulista - UNESP



Descrição morfológica preliminar dos termos da Educação a Distância

Ana Catarina Gianoti
Ariani Di Felippo

NILC-TR-11-02

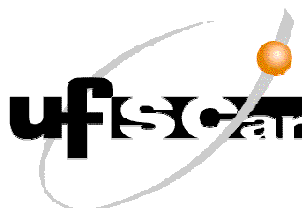
Agosto, 2011

Série de Relatórios do Núcleo Interinstitucional de Lingüística Computacional
NILC - ICMC-USP, Caixa Postal 668, 13560-970 São Carlos, SP, Brasil

Resumo

Neste relatório, estão descritas parte das atividades realizadas no projeto de iniciação científica (IC) intitulado “*Extração automática de termos segundo o paradigma linguístico*”, o qual integrou o projeto “*Instanciação e aplicação de uma metodologia para o desenvolvimento de wordnets terminológicas em português do Brasil*” (doravante, TermiNet) (FAPESP Proc. 2009/06262-1 e CNPq Proc. 471871/2009-5). O objetivo principal do referido projeto de IC foi a extração automática de terminologia (EAT) baseada no paradigma linguístico. Essa tarefa foi realizada no âmbito do projeto TermiNet cujo objetivo final é a construção de uma base de dados lexicais terminológicas do domínio da Educação a Distância (EaD) em português do Brasil (PB) no formato da WordNet de Princeton (FELLBAUM, 1998). A EAT consiste em reconhecer/ extrair possíveis unidades terminológicas, a partir de um *corpus* (“extensos conjuntos de textos em formato digital”), de forma automática e é uma das tarefas previstas na metodologia proposta no TermiNet (DI-FELIPPO, ALMEIDA, 2010) para a construção de bases lexicais terminológicas no formato *wordnet*. Da EAT, são obtidos os termos que, uma vez validados, são organizados em função da sinonímia, pois uma *wordnet* armazena essencialmente (i) conceitos lexicalizados, codificados em “conjuntos de sinônimos” (em inglês, *synonym sets* ou *synsets*), e (ii) diferentes relações semântico-conceituais entre *synsets*. Como resultado concreto, o projeto foi responsável pela identificação de 59 unigramas, 102 bigramas, 63 trigramas e 5 quadrigramas, os quais serão diretamente inseridos na base de dados terminológica do domínio da EaD denominada WordNet.EaD. Neste relatório, em especial, descreve as atividades realizadas para o cumprimento do objetivo secundário do referido projeto, no caso, a descrição da morfologia dos 59 unigramas extraídos pela ferramenta ExATOlp e validados pelo especialista de domínio.

Este trabalho contou com o apoio financeiro da FAPESP e do CNPq/UFSCar.



Sumário

Resumo	ii
1. Introdução	2
2. Fundamentação teórica	4
2.1. Os termos e sua estrutura morfossintática	4
2.2. A Morfologia	6
2.2.1. Os processos de formação lexical.....	6
3. Metodologia.....	10
4. Resultados.....	11
4.1. Derivação prefixal.....	11
4.2. Derivação sufixal	11
4.3. Composição acronímica/sigla	11
4.4. Composição coordenativa.....	11
4.5. Composição subordinativa.....	12
4.6. Empréstimo	12
4.7. Derivação regressiva.....	12
4.8. Palavra-valise.....	12
5. Considerações finais	12
Agradecimentos	12
6. Referências bibliográficas	12
APÊNDICE 1: Processos morfológicos de 59 unigramas da EaD.....	16

1. Introdução

A definição e a implementação de metodologias para a extração de candidatos a termos (EAT), a partir de *corpora*, têm assumido um papel central no desenvolvimento de produtos terminológicos, como dicionários e glossários. Recentemente, o desenvolvimento de metodologias semiautomáticas, pautadas em técnicas do Processamento Automático das Línguas Naturais (PLN), tem acelerado o processo de extração dos candidatos a partir de *corpora*, contribuindo para o desenvolvimento de produtos terminológicos clássicos (ou seja, dicionários e glossários) e/ou linguístico-computacionais (ou seja, léxicos computacionais, bases de dados, etc.). O crescente interesse pela sistematização do conhecimento especializado em recursos linguístico-computacionais pauta-se principalmente no fato de que aplicações de PLN, como a tradução e sumarização, precisam lidar com textos reais, sejam eles de língua geral ou especializados.

Desde o surgimento do TERMINO¹, considerado o primeiro sistema de extração automática de candidatos a termos, diversos projetos têm sido elaborados com a finalidade de projetar extratores semiautomáticos. Tais extratores têm sido elaborados com base em um dos três paradigmas de EAT (CABRÉ *et al.*, 2001; JACQUEMIN, BOURIGAULT, 2003; PAZIENZA *et al.*, 2005; BERNHARD, 2006): o linguístico, o estatístico e o híbrido.

No paradigma linguístico, busca-se identificar os candidatos a termo com base em certos conhecimentos linguísticos. Para que os extratores desenvolvidos segundo o paradigma linguístico funcionem, o *corpus* sob análise deve ser pré-processado. Dependendo do método linguístico implementado, esse pré-processamento pode incluir os processos de: (i) tokenização (no inglês, *tokenization*), (ii) etiquetagem morfosintática (no inglês, *tagging*), (iii) lematização (no inglês, *lemmatization*), (iv) análise sintática (no inglês, *parsing*) e/ou (v) exclusão de *stopwords*. Para a extração dos candidatos simples (unigramas), os extratores comumente utilizam dois tipos de conhecimento linguístico, os quais dão origem a dois métodos de EAT, a saber: categoria sintática ou nuclearidade sintagmática. Quando baseados no método que identifica os candidatos a termos pela categoria sintática (verbo, adjetivo, advérbio e verbo), os extratores possibilitam aos usuários a especificação da categoria que desejam extrair do *corpus*; uma vez especificada, somente as unidades da categoria em questão são extraídas. Quando baseados no reconhecimento de estruturas sintagmáticas, as ferramentas apenas extraem candidatos considerados núcleo de sintagmas (em especial, de sintagmas nominais (SNs)). Para a extração de candidatos complexos (n-gramas > 1), os extratores podem utilizar três tipos de informação linguística, ou seja, três métodos distintos, a saber: padrão morfosintático, expressão indicativa ou nuclearidade sintagmática. Os padrões morfosintáticos são sequências de etiquetas morfosintáticas, exemplificadas aqui pela expressão regular [Nome+Adjetivo] (p.ex.: “educação infantil”). As expressões indicativas (ou padrões léxico-sintáticos) introduzem definições e os termos definidos, por exemplo: “é definido como”, “é um tipo de” (p.ex.: “A Educação a Distância (EaD) é um tipo de modalidade de ensino [...]”), etc. Dentre os trabalhos que se encaixam nessa linha, citam-se, por exemplo, os de Aussenac-Gilles e Séguéla (2000), Suárez e Cabré (2002), Cederberg e Widdows (2003); Morin e Jacquemin (2004), Morin e Jacquemin (2004), Mititelu (2006), etc.

¹ A versão 1.0 do TERMINO foi criada em 1989 pelo o grupo de Recherche et développement en linguistique computationnelle (RDLC) do Centro ATO (Analyse de textes par ordinateur) da Universidade de Quebec Montreal.

No paradigma estatístico, os candidatos são extraídos com base na aplicação de medidas como frequência, informação mútua, *log-likelihood ratio* e coeficiente Dice². Para a extração de unidades simples, utiliza-se comumente a frequência simples, que pode ser entendida como a quantidade de vezes que um *token* (isto é, sequência de caracteres separada por espaços em branco) ocorre no *corpus*. A frequência simples também pode ser usada para a extração de termos complexos. As demais estatísticas (informação mútua, *log-likelihood ratio* e coeficiente Dice) são utilizadas apenas para a extração dos candidatos complexos, pois buscam identificar a estabilidade de expressões sintagmáticas, ou seja, a correlação entre n-gramas.

Para paradigma híbrido, os sistemas de EAT combinam métodos baseados em ambos os paradigmas: linguístico e estatístico. Os métodos híbridos são desenvolvidos com base no pressuposto de que a confiabilidade das medidas estatísticas é maior quando estas são aplicadas a candidatos a termo linguisticamente “justificados”. Um exemplo de sistema híbrido é o YATE (do inglês, “*Yet Another Term Extractor*”) (VIVALDI, 2001), que foi desenvolvido para extrair candidatos a termos dos tipos (Nome) e (Nome + Adjetivo) a partir de textos em espanhol do domínio da Medicina. Especificamente, o YATE combina vários métodos de extração, cada um deles baseado em um tipo diferente de informação, a saber: (i) semântica; (ii) morfológica (especificamente, morfemas greco-latinos); (iii) contextual (léxico-semântico) e (iv) estatística. Segundo Vivaldi e Rodríguez (2001), qualquer combinação desses métodos produz melhores resultados de extração do que qualquer um dos métodos aplicados individualmente.

No caso específico da extração de candidatos a partir de textos em PB, há trabalhos sendo desenvolvidos segundo os três paradigmas, sendo que os principais métodos/ferramentas utilizados em tais trabalhos estão descritos no Quadro 3.

Quadro 3: Métodos/ferramentas de extração utilizados no TerminiNet.

Paradigma	N-Grama	Conhecimento	Ferramenta
Linguístico	1, 2, 3, 4	Núcleo de sintagma nominal	ExATOp (Lopes <i>et al.</i> , 2009)
Híbrido	1	Classe gramatical/tf-idf	OntoLP (Ribeiro Jr., 2008)
	2	Padrões morfossintáticos/tf-idf	
	3	Padrões morfossintáticos/tf-idf	
	4	Padrões morfossintáticos/tf-idf	

Quanto ao método estatístico, a EAT a partir de *corpora* em PB comumente utiliza as medidas disponíveis no pacote NSP (do inglês, *N-gram Statistics Package*) (BANERJEE, PEDERSEN, 2003), que é constituído por um conjunto de programas que auxilia na análise de *n*-gramas (isto é, sequência de *n* elementos em texto) em arquivos no formato *.txt*.

Diante desse cenário, vê-se que os métodos/ferramentas disponíveis para o processamento de textos em PB não incluem estratégias de extração baseadas na estrutura morfológica dos termos, assim como faz o YATE.

A relevância do conhecimento de nível morfológico para a extração de candidatos a termos é enfatizada, por exemplo, por Almeida e Vale (2008). Os autores destacam as principais características morfológicas de três domínios especializados: Revestimentos Cerâmicos, Fisioterapia e Nanociência e Nanotecnologia.

² Mais informações sobre essas medidas estatísticas são encontradas em Ribeiro Jr. (2008).

Segundo os autores, o domínio dos Revestimentos Cerâmicos caracteriza-se pela alta frequência de lexias complexas (p.ex.: *argila refratária aluminosa, análise granulométrica por peneiramento*, etc.) e lexias simples com determinados morfemas, como os sufixos derivacionais *agem, -ção, -mento*, que podem indicar que o conceito veiculado é um processo (p.ex.: *secagem, moagem, trituração, atomização, peneiramento, destorroamento*, etc.). Na Fisioterapia, a terminologia é caracterizada pelas formações eruditas comumente emprestadas da Medicina, com radicais e/ou afixos de origem grega ou latina, p.ex.: (i) *atr(i/o)-* (do grego, *árthron*, ou seja, elemento de composição que significa “juntura” ou “articulação”), como em *artralgia, artrectomia, artrite*, etc.; e (ii) *-óide* (do grego, *-(o)eîdos*, pelo lat. cient. *-īdes*, ou seja, sufixo que pode significar “aspecto ou forma de”), como em *deltóide, escafóide, articulação*, etc. Por fim, Almeida e Vale (2008) salientam que a terminologia da Nanociência e Nanotecnologia (N&N), por sua vez, apresenta lexias simples e complexas, sendo que uma das características dessa terminologia é a alta frequência do prefixo *nano-*, como em *nanocompósitos, nanocristais, nanoeletromecânicos, nanossistema biológico*, etc.

As características morfológicas da N&N, aliás, são utilizadas pelos autores para ilustrar a relevância das mesmas na identificação semiautomática de candidatos a termos a partir de *corpus*. Em especial, Almeida e Vale (2008) exemplificam buscas em *corpus* por meio do software Unitex³ com base especificamente no conhecimento morfológico dos termos do referido domínio.

No entanto, para elaborar métodos de EAT com base em conhecimento morfológico, é preciso descrever pormenorizadamente as estruturas morfológicas das terminologias. Nesse sentido, apresentam-se neste relatório, as primeiras aproximações à descrição da morfologia da terminologia da Educação a Distância (EaD). Diz-se “aproximações” porque se apresenta aqui apenas a descrição de parte da lista de lexias simples (ou unigramas) identificadas no projeto TerminiNet, que visa à construção de uma *wordnet* terminológica (ou *terminet*) para a domínio da EaD (DI-FELIPPO, 2010).

Para apresentar essa descrição parcial, este relatório está dividido em 5 Seções. Na Seção 2, apresenta-se uma breve fundamentação teórica sobre os termos e sua estrutura morfossintática, assim como sobre Morfologia e os principais processos de formação de palavras. Na Seção 3, apresenta-se a metodologia empregada na referida descrição. Os resultados da descrição morfológica parcial são apresentados na Seção 4.

2. Fundamentação teórica

2.1. Os termos e sua estrutura morfossintática

Para a construção de uma base *wordnet*, seja ela de língua geral ou especializada, é preciso especificar o conhecimento lexical e o semântico-conceitual a serem sistematizados. Quanto ao conhecimento lexical, é preciso especificar os elementos constitutivos dos *synsets*, ou seja, os tipos de unidades lexicais que formarão os conjuntos de sinônimos. Quanto ao conhecimento semântico-conceitual, é necessário especificar a noção de sinonímia para a construção dos *synsets* e as relações semântico-conceituais que interligarão os *synsets* (cf. DI-FELIPPO, 2010).

Especificamente quanto ao conhecimento lexical, as unidades constitutivas dos *synsets* em uma *terminet* são as “unidades terminológicas” ou “termos” do domínio especializado. Um “termo” pode ser definido como signo linguístico, composto de forma

³ O software Unitex foi desenvolvido na Universidade Marne-La-Vallée (França) por Sébastien Paumier (PAUMIER, 2002).

e conteúdo, que representa certo conceito no interior de um domínio específico (Cabré, 1999). Os tipos de unidades a serem armazenadas em uma *terminet* foram delimitados em função de dois critérios: formal e funcional. Do ponto de vista formal ou de sua estrutura morfossintática, uma *terminet* poderá armazenar termos simples⁴, constituídos de apenas um radical, com ou sem afixos (p.ex.: “ambiente”), e complexos, constituídos de dois ou mais radicais, aos quais podem-se acrescentar outros elementos⁵ (p.ex.: “ambiente digital”, “ambiente de aprendizagem”, “ambiente virtual de aprendizagem”).

Os termos compostos também são unidades lexicais formadas por dois ou mais radicais. No entanto, eles diferem dos complexos pelo alto grau de lexicalização e pelo conjunto de morfemas lexicais e/ou gramaticais que os constitui. No caso, os termos compostos por aglutinação (p.ex.: “teleeducação” [tel(e)+educação]) e por justaposição (com ou sem hífen) (p.ex.: “livro-texto” e “weblog” [web+log]) são considerados termos simples⁶.

Os termos complexos podem ter tamanhos distintos; os sintagmas terminológicos chegam a compor-se de até 5 unidades lexicativas (p.ex.: *amilóide cutânea genuína localizada maculosa*) (Alves, 1999:101). Tendo em vista que os termos complexos de tamanho 5 ou pentagramas são menos frequentes, optou-se por restringir a constituição dos *synsets* de uma *terminet* às unidades dos tipos unigrama, bigrama, trigrama e quadrigrama, como as exemplificadas em (1).

- (1) a. ambiente (unigrama ou 1-grama)
- b. ambiente virtual (bigrama ou 2-grama)
- c. ambiente de aprendizagem (trigrama ou 3-grama)
- d. ambiente virtual de aprendizagem (quadrigrama ou 4-grama)

Quanto ao critério funcional, optou-se por restringir a construção de uma *wordnet* terminológica às unidades lexicais da categoria dos nomes (substantivos), pois estas ocupam um lugar de destaque nas terminologias, ou seja, no conjunto de termos de uma área especializada (Barros, 2004). Devido à forte interdependência entre a formação dos conceitos e a formação de suas expressões linguísticas, as classes de palavras são associadas às classes de conceitos. Dessa associação, fala-se em “conceitos nominais”, “conceitos verbais”, etc. (Sager, 1993). Uma *terminet* codificará, assim, conceitos nominais (isto é, expressos por termos da classe dos nomes).

Dessa forma, fez-se um recorte no conhecimento lexical comumente armazenado em uma base do tipo *wordnet*. Como consequência desse recorte, a WordNet.EaD deve armazenar apenas unigramas, bigramas, trigramas e quadrigamas da categoria dos nomes.

Após a delimitação do conhecimento lexical a ser armazenado em uma *terminet*, o próximo passo, segundo a metodologia proposta no TerMiNet, consiste na seleção e/ou construção dos recursos-fonte a partir dos quais esse conhecimento deve ser extraído e na seleção da estratégia de extração.

⁴ Do ponto de vista computacional, os termos simples são definidos como n-gramas de tamanho 1 (unigramas), ou seja, sequências únicas de caracteres separados por espaços em branco.

⁵ Do ponto de vista computacional, os termos complexos são n-gramas de tamanho maior que 1, isto é, sequências de 2 unigramas (bigramas), 3 unigramas (trigramas), etc.

⁶ Do ponto de vista computacional, os termos compostos são considerados unigramas.

2.2. A Morfologia

O termo “morfologia”, que do ponto de vista etimológico, significa “estudo da forma”. Criado em 1830 por Goethe, esse termo foi aplicado inicialmente no domínio das ciências biológicas “para designar o estudo das formas dos organismos vivos”. No âmbito da Linguística, esse termo somente foi aplicado aos estudos sobre a linguagem no século XIX, por volta de 1860, englobando os processos de flexão e a derivação (LAROCA, 2005).

Nas décadas de 1940 e 1950, a morfologia atinge o seu auge com os trabalhos realizados no cenário da morfologia estruturalista americana, também conhecida como “distribucionalismo”. Nessa forma de estruturalismo, a descrição linguística baseia-se na distribuição das unidades nos contextos linguísticos em que aparecem. Dessa característica é que nasce o nome “distribucionalismo”. Os distribucionalistas foram os responsáveis pela definição de “morfema”, sobre a qual está pautada a morfologia de cunho estruturalista. No Brasil, destaca-se a obra *Princípios de linguística geral* de Joaquim Mattoso Câmara Jr, que aborda a descrição sincrônica das estruturas morfológicas do português (LAROCA, 2005). No estruturalismo/ distribucionalismo, considera-se a língua como um sistema de valores que se opõem entre si, formando as mais variadas estruturas da língua. Sob esse enfoque, os objetivos de uma descrição morfológica consistem em segmentar e classificar os elementos que compõem a estrutura morfológica (sincrônica) das línguas.

Na linguística gerativa, que compreende as investigações embasadas na obra de Chomsky, a língua é algo muito mais profundo, inerente à condição humana, relacionado com o aspecto criador da linguagem. Consequentemente, quanto à morfologia, os linguistas gerativos buscam explicitar as regras internalizadas que os falantes possuem e que são responsáveis pela capacidade criadora das palavras. No âmbito da morfologia de cunho gerativo, destacam-se as obras de Jackendoff e Aronoff na década de 70. No Brasil, destacam-se os estudos de Margarida Basílio na década de 80 (LAROCA, 2005).

Em linhas gerais, a morfologia pode ser dividida em dois domínios: (i) a morfologia flexional, que se ocupa do estudo das relações entre diferentes formas de uma mesma palavra (ou seja, o paradigma flexional) e (ii) a morfologia lexical, que trata das palavras e de seus processos de formação (LAROCA, 2005).

Tendo em vista que a descrição morfológica dos termos da EaD focaliza a morfologia lexical, apresenta-se, a seguir, a tipologia dos processos de formação de palavras que norteou a descrição em questão.

2.2.1. Os processos de formação lexical

Com base em Rocha (2008) e Alves (2004), é possível classificar os processos de formação de palavras em 6 tipos, a saber: (i) processo fonológico, (ii) processos sintáticos, (iii) conversão, (iv) processo semântico, (v) empréstimo e, por fim, organizados em um último grupo, estão os processos de (vi) truncação, palavra-valise, reduplicação e derivação regressiva. Com exceção dos processos fonológicos, os principais tipos estabelecidos por Alves estão sistematizados na Figura 1.

O processo fonológico ocorre quando um item lexical é formado a partir de um significante totalmente inédito. Os processos sintáticos caracterizam-se pela combinação de elementos já existentes no sistema linguístico. Nesse grupo, listam-se os processos de derivação (prefixal, sufixal, parassintética) e de composição (coordenativa, subordinada, sintagmática e acronímica). O processo de conversão ou derivação imprópria é definido como o enriquecimento vocabular ocasionado pela mudança de classe e é tratado como sendo de natureza sintático-semântica e não puramente morfológica. A formação de um

item lexical por meio de processo semântico ocorre quando o item é formado sem que se opere nenhuma mudança formal em unidades léxicas já existentes; assim, qualquer transformação semântica manifestada em um item lexical ocasiona a criação de um novo elemento (ALVES, 2004). Por empréstimo, entende-se o processo por meio do qual um item lexical de uma língua passa a ser empregado em outra língua. Por fim, no último grupo, listam-se processos menos produtivos, como a truncação, palavra-valise, reduplicação e derivação regressiva. A seguir, são descritos com mais detalhes os principais processos da Figura 1.

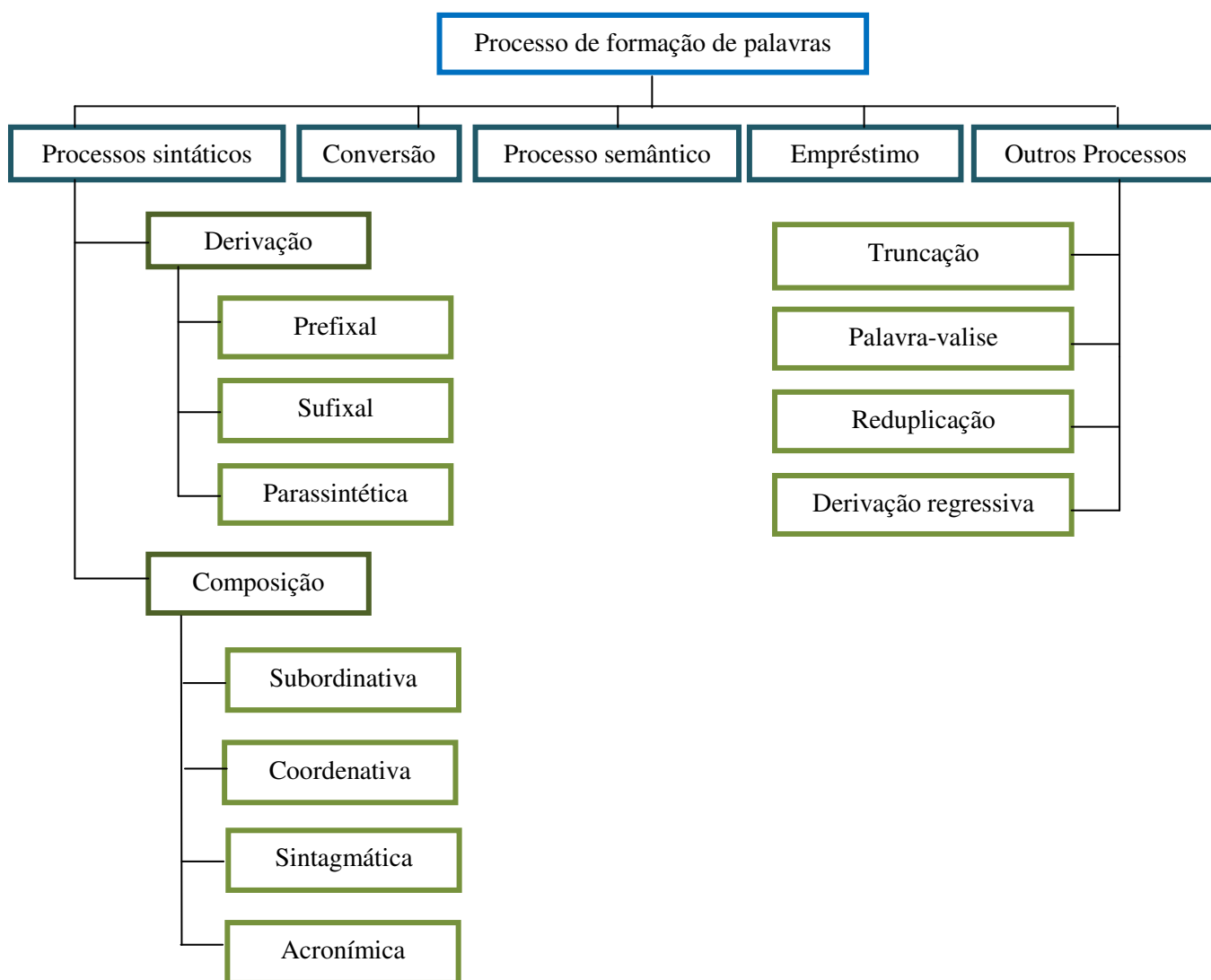


Figura 1: Tipologia de processos de formação segundo Alves (2004).

a) Processos sintáticos: derivação e composição

A partir de processos sintáticos, formam-se palavras por meio da combinatória de elementos já existentes na língua. Segundo Alves (2004), tais processos são denominados sintáticos porque a combinação de seus membros constituintes não está circunscrita exclusivamente ao âmbito lexical. Os processos sintáticos aqui identificados foram classificados em *derivação* (prefixal, sufixal e parassintética) e *composição* (coordenativa, subordinativa, sintagmática e acronímica). A seguir, serão apresentados os tipos de derivação e de composição.

- *Derivação prefixal*

Esse processo tem como característica a união de um prefixo a uma base, acrescentando a esta, diversos significados.

“(...) trataremos como prefixos as partículas independentes ou não-independentes que, antepostas a uma palavra-base, atribuem-lhe uma idéia acessória e manifestam-se de maneira recorrente, em formações em série” (ALVES, 2004, p. 15).

Ressalta-se com base em Sandmann (1989) que elementos de origem estrangeira, como auto- e pseudo-, ao se integrarem ao português gerando formações em série, fazem parte dos prefixos e não são considerados radicais estrangeiros presos. Alves (2004) e Rocha (2008) citam os seguintes prefixos: anti-, auto-, contra-, de-, des-, dis-, em-, ex-, extra-, hiper-, in-, justa-, mega-, micro-, mini-, multi-, não-, neo-, pós-, pré-, pró-, quase-, re-, recém-, sobre-, sem-, semi-, super-, trans-, ultra-, entre outros.

- *Derivação sufixal*

A derivação sufixal consiste na anexação de um sufixo a uma base (ROCHA, 2008). Em outras palavras, o sufixo que é um elemento de caráter não-autônomo e recorrente e atribui à palavra-base a que se associa uma idéia acessória e, com frequência, altera-lhe a classe gramatical (ALVES, 2004). Alves (2004) e Rocha (2008) listam os seguintes sufixos: -ão, -aco, -ama, -ano, -âneo, -ança, -ção, -ar, -aria, -ário, -ada, -agem, -al, -dor, -eo, -ento, -esco, -eiro, -estre, -iano, -ice, -ície, -ico, -idade, -idão, -ino, -inho, -ismo, -ivo, -ista, -logia, -mente, -mento, -nça, -ncia, -nte, -óide, -oso, -izar, -tivo, -vel, entre outros.

- *Derivação parassintética*

A derivação parassintética constitui-se da junção simultânea de um prefixo e um afixo a uma base nominal. Esse é o caso, por exemplo, de o item lexical *esclarecer* que, a partir de *claro*, não se acrescenta primeiro um prefixo (**esclaro*) e depois um sufixo (*esclarecer*), e tão pouco não se acrescenta primeiro um sufixo (**clarecer*) e depois um prefixo (*esclarecer*). No caso, o item *esclarecer* é formado diretamente de *claro*, ou seja, há o acréscimo simultâneo dos afixos (ROCHA, 2008). Outros exemplos que ilustram esse tipo de derivação são: *achocolatado* (derivado de *chocolate*), *desalmado* (derivado de *alma*), *abotoar* (derivado de *botão*), etc.

b) Composição

A composição consiste na “justaposição de bases autônomas ou não-autônomas. A unidade léxica composta, que funciona morfológica e semanticamente como um único elemento, não costuma manifestar formas recorrentes, o que a distingue da unidade constituída por derivação” (ALVES, 2004). Segundo Cunha e Cintra (2001), a palavra composta representa sempre uma idéia única e autônoma, muitas vezes dissociada das noções expressas pelos seus componentes. Dentre os diferentes tipos de composição, listam-se a composição, coordenativa, subordinativa, sintagmática e acronímica (ou siglagem).

- *Coordenativa*

Segundo Alves (2004), na composição coordenativa, não há uma relação de subordinação do tipo determinado/determinante. Assim, as bases que compõem a unidade lexical

desempenham a mesma função que a do elemento recém formado e associam-se cumulativamente a fim de formarem esse elemento. Consequentemente, a função sintática da coordenação é exercida pela justaposição de substantivos, adjetivos ou membros de outra classe gramatical.

- *Subordinativa*

A relação subordinativa ocorre entre duas bases, em que a primeira exerce o papel de determinado e a segunda, de determinante. Em outras palavras, a base determinada constitui um elemento genérico ao qual o determinante acresce uma especificação, ou seja, uma característica da classe adjetival (ALVES, 2004).

- *Sintagmática*

As formações sintagmáticas são, segundo Alves (2004), segmentos frasais cujos membros se encontram em uma íntima relação sintática, tanto morfológica quanto semanticamente, de forma a constituírem uma única unidade léxica. Ainda de acordo com a autora, os membros integrantes do composto sintagmático conservam as peculiaridades flexionais de suas categorias de origem. Ao diferenciar uma composição sintagmática de itens lexicais compostos, Alves (2004) estabelece que a unidade lexical sintagmática encontra-se ainda em vias de lexicalização e, por isso, não costuma ser unida por hífen. O item léxico composto, ao contrário, é geralmente transcrito com essa marca gráfica.

- *Acronímica (ou siglagem)*

As composições por acronímia (ou siglagem ou, simplesmente, siglas) são, segundo Alves (2004), um tipo especial de composição sintagmática, resultantes do processo de economia discursiva. Em outras palavras, diz-se que o sintagma é reduzido de modo a se tornar mais simples e mais eficaz no processo da comunicação. Rocha (2008) afirma que nesse tipo de formação a base será um substantivo ao mesmo tempo composto e próprio, e o produto, um lexema simples e próprio, formado, em princípio, pelos grafemas e sílabas iniciais dos lexemas composto que constitui a base. Com isso, o autor distingue alguns tipos de siglagem:

- *Siglagem grafêmica*: composta pelos grafemas iniciais das bases compostas (Ex.: Ordem dos Advogados do Brasil – OAB);
- *Siglagem silábica*: composta pelas sílabas iniciais das bases (Ex.: Departamento de Transito – Detran);
- *Siglagem grafossilábica*: composta pelos grafemas e pelas sílabas iniciais das bases (Ex.: Banco do Estado de São Paulo – Banespa);
- *Siglagem fortuita*: compostas pelos mais variados critérios de formação com inclusões e cortes aleatórios de sílabas ou palavras, preocupando-se com o aspecto sonoro, procurando constituir uma sigla chamativa e original (Ex.: Superintendência de Desenvolvimento da Capital – Sudecap); neste caso, temos em cap- (de capital) um corte aleatório;
- *Siglagem significativa*: em que o produto resulta em um nome relacionado com o sentido da sigla (Ex.: FaLe – Faculdade de Letras); a sigla nos remete a *falar*, o verbo-síntese dos estudiosos da área.

c) Conversão

A conversão também é chamada de derivação imprópria e consiste no emprego de uma palavra de determinada classe lexical em outra classe (ROCHA, 2008). Em outras

palavras, isso quer dizer que a conversão designa um tipo de formação lexical pelo qual uma unidade lexical sofre alterações em sua distribuição sem que haja manifestação de mudanças formais (ALVES, 2004), ou seja, esse processo ocorre quando há alteração da classe gramatical.

d) Empréstimo

O contato com comunidades linguísticas diversas faz com que novas unidades lexicais sejam inseridas ou criadas em uma determinada língua. Alves (2004) destaca três diferentes modos de integração dos empréstimos. No entanto, para descrição ora apresentada, considerou-se como empréstimo o que Alves (2004) denominou “estrangeirismo”, ou seja, o item lexical estrangeiro empregado em uma língua.

e) Outros processos

Nesse grupo, estão listados os processos de truncação, palavra-valise, reduplicação e derivação regressiva.

A truncação é um tipo de abreviação em que uma parte da sequência lexical, geralmente a final, é eliminada (p.ex.: *depressão* > *deprê*).

O processo denominado palavra-valise é um tipo de redução em que duas bases são privadas de parte de seus elementos para constituírem um novo item léxico (p.ex.: *portunhol*).

A reduplicação caracteriza-se por um recurso morfológico em que uma mesma base é repetida duas ou mais vezes a fim de constituir um novo item lexical. Um exemplo desse processo é a reduplicação *trança-trança* que, segundo Alves (2004), tem como origem a repetição de uma base verbal denotativa de “andar para diversos lados”, já a repetição dessa base, em função substantival, implica uma “sucessão de viagens”.

A derivação regressiva, também chamada de derivação sufixal zero, consiste na formação de uma palavra por meio da supressão de um elemento de caráter sufixal. Segundo Alves (2004), a maior parte dos casos de derivação regressiva, em português, constitui-se por substantivos deverbais, resultantes da substantivação de suas respectivas formas verbais e formados pelo acréscimo das vogais *-a*, *-e* e *-o* ao radical do verbo (p.ex.: *amassar* → *amasso*). Rocha (2008) defende esse tipo de derivação como sendo derivação sufixal zero, pois, ao observar a nominalização, nota-se que além desta se efetuar com sufixos explícitos, como *-mento*, *-ção*, *-agem*, etc. (p.ex.: *patrulhar* → *patrulhamento*), pode também ser efetuada com o sufixo implícito zero, p.ex.: *patrulhar* + sufixo zero → *patrulhaØ*, *desmatar* + sufixo zero → *desmateØ*, etc.

3. Metodologia

No projeto TermiNet, os candidatos a termos foram extraídos no *Corpus.EaD*, o qual engloba 347 textos em PB manualmente compilados de fontes de qualidade, ou seja, de conteúdo confiável (DI-FELIPPO, SOUZA, 2010). Tais fontes foram restritas a sites de instituições governamentais reconhecidamente vinculadas a projetos de EaD.

A EAT foi feita por meio dos principais método/ferramentas disponíveis para o PB. Dessa extração, 172 unigramas, 167 bigramas, 147 trigramas e 15 quadrigramas foram validados.

Dos 172 unigramas validados, apenas 59 foram submetidos à descrição morfológica. Esses 59 termos foram especificamente extraídos do *Corpus.EaD* pelo extrator linguístico ExATOlp (LOPES *et al.*, 2009). A descrição morfológica consistiu na identificação dos processos morfológicos por meio das quais tais unigramas foram

formados. Para tanto, utilizou-se a tipologia da Figura 1. Salienta-se que na descrição dos unigramas, o processo sintático de composição sintagmática não é observado, posto que diz respeito às formações complexas.

Para a descrição dos processos de formação lexical, em especial, organizou-se uma tabela contendo: (i) todos os unigramas validados provenientes do ExATOlp e (ii) a indicação do processo que caracteriza cada um dos unigramas. Tal tabela consta no Apêndice 1 deste documento. A seguir, apresentam-se os processos de formação de palavras identificados nos termos da EaD, sendo que, para cada processo, listam-se os termos em questão.

4. Resultados

De um modo geral, os processos identificados foram:

4.1. Derivação prefixal

No total, foram identificados 8 casos de derivação prefixal que englobam os prefixos *hiper-*, com duas ocorrências (*hipermídia* e *hipertexto*), e *inter-*, com 3 ocorrências (*interação*, *interatividade* e *interdisciplinaridade*). Nos outros 3 casos, a formação ocorre com a base presa *tele-* (*teleaula*, *teleconferência* e *teleduc*).

4.2. Derivação sufixal

Esse processo é o mais produtivo no cenário da EaD em questão, com 22 ocorrências. Os sufixos que participam da formação das unidades lexicais são os seguintes: *-nte*, com 2 ocorrências na lista (*docente* e *aprendente*), *-mento* (*compartilhamento*), *-ção*, com 3 ocorrências dentre as investigadas (*assimilação*, *conversação* e *mediação*), *-dor*, com 5 ocorrências identificadas (*avaliador*, *educador*, *formador*, *mediador* e *orientador*), *-or*, com 2 ocorrências (*gestor* e *mentor*), *-tor* (*instrutor*), *-ismo* (*construtivismo*), *-ista* (*cursista*), *-ando* (*educando*), *-ança* (*mudança*), *-al* (*portal*), *-dade* (*subjetividade*), *-ria* (*tutoria*) e *-ial* (*tutorial*).

4.3. Composição acronímica/sigla

Após a derivação sufixal e a prefixal, a composição acronímica/sigla ocupa o terceiro lugar entre os processos de formação mais produtivos no conjunto de unigramas investigado. No total, foram 11 ocorrências, sendo (i) 7 ocorrências de sigla gráfica, no caso, *ead* (*educação a distância*), *nead* (*núcleo de educação a distância*), *uab* (*universidade aberta do Brasil*), *ava* (*ambiente virtual de aprendizagem*), *eja* (*educação de jovens e adultos*), *tic* (*tecnologia da informação e da comunicação*) e *moodle* (do inglês, *Modular Object-Oriented Dynamic Learning Environment*) e (ii) 4 ocorrências de sigla fortuita, no caso, *webct* (do inglês, *web course tools*), *seed* (*secretaria de estado da educação*) e *mobral* (*Movimento Brasileiro de Alfabetização*) e *cederj* (*Centro de Educação Superior a Distância do Rio de Janeiro*).

4.4. Composição coordenativa

Os casos de composição coordenativa encontrados foram *videoaula* e *videoconferência*. Esses casos foram assim classificados porque parece não haver uma relação de subordinação do tipo determinado/determinante. Dessa forma, *vídeo* e *aula*, por exemplo, desempenham a mesma função que a do elemento recém formado (*videoaula*).

4.5. Composição subordinativa

Os 5 casos de composição subordinativa foram *andragogia* ([andr(o)]agogia), *pedagogia* ([ped]agogia), *pedagogo* ([ped]agogo), *ciberspaço* ([ciber]espaço) e *cibercultura* ([ciber]cultura). Nesses casos, o elemento à esquerda parece determinar/ especificar o da direita.

4.6. Empréstimo

Quanto aos empréstimos, no caso, estrangeirismos, foram identificadas 6 ocorrências, são elas: *blog*, *chat*, *link*, *designer*, *moodle* e *webct*.

4.7. Derivação regressiva

Dentre os casos investigados, foi encontrado apenas *apostila* (<*apostilar*) como caso de derivação regressiva.

4.8. Palavra-valise

Na lista de unigramas investigada, foram identificados 5 casos de palavra-valise, a saber: *aulanet* (*aula(inter)[net]*), *ciberspaço* ([*ciber(nético)*]espaço), *cibercultura* ([*ciber(nético)*]cultura), *teleduc* (*tele[duc(ação)]*) e *telemática* (*tele[(infor)mática]*).

Ao final, os processos identificados ficaram assim distribuídos:

- Derivação prefixal: 8
- Derivação sufixal: 22
- Composição acronímica/sigla: 11
- Composição coordenativa: 2
- Composição subordinativa: 5
- Empréstimos: 6
- Derivação regressiva: 1
- Palavra-valise: 5

5. Considerações finais

Até o momento, foram descritos apenas parte dos unigramas validados no TermiNet. No entanto, a descrição dos demais unigramas, bigramas, trigramas e quadrigramas está em plena realização. Ao final do processo de descrição, buscar-se-á caracterizar a terminologia da EaD em PB quanto aos processos morfológicos e elementos formativos envolvidos nesses processos.

Agradecimentos

Agradecemos à Coordenadoria de Iniciação Científica e Tecnológica da Pró-Reitoria de Pesquisa da UFSCar pela bolsa concedida no âmbito do Programa Institucional de Bolsas de Iniciação Científica – PIBITI/CNPq/UFSCar.

6. Referências bibliográficas

ALMEIDA, G. M. B.; VALE, O. A. Do texto ao termo: interação entre Terminologia, Morfologia e Linguística de Corpus na extração semi-automática de termos. In: ISQUERDO, A. N.; FINATTO, M. J. B. (Orgs.). **As ciências do Léxico: Lexicologia, Lexicografia e Terminologia**. Campo Grande: Editora da UFMS, 2008, v. IV, p. 483-499.

- ___ ; CORREIA, M. Terminologia e corpus: relações, métodos e recursos. In: Stella E. O. Tagnin; Oto Araújo Vale. (Org.). **Avanços da Linguística de Corpus no Brasil**. 1 ed. São Paulo: Humanitas/FFLCH/USP, 2008, v. 1, p. 63-93.
- ___.; ALUÍSIO, S. M.; OLIVEIRA, L. H. M. O método em Terminologia: revendo alguns procedimentos. In: ISQUERDO, A. N.; ALVES, I. M. (Orgs.). **Ciências do léxico: lexicologia, lexicografia, terminologia**. 1 ed. Campo Grande/São Paulo: Editora da UFMS/Humanitas, 2007, v. III, p. 409-420.
- ALVES, I. M. **Neologismo – criação lexical**. São Paulo: Ática, 2004.
- AUSSENAC-GILLES, N.; SÉGUÉLA, P. Les relations sémantiques: du linguistique au formel. **Cahiers de Grammaire**, n.25, p. 175–198, 2000.
- BARROS, L.A. **Conhecimento de Terminologia Geral para a prática tradutória**. São José do Rio Preto, SP: Editora NovaGraf, 2007.
- BANERJEE, S.; PEDERSEN, T. The design, implementation, and use of the N-gram Statistics Package. In: INTERNATIONAL CONFERENCE ON INTELLIGENT TEXT PROCESSING AND COMPUTATIONAL LINGUISTICS, 4, 2003, **Proceedings...** Mexico City, Mexico, 2003, p. 370-381.
- BARROS, L. A. **Curso básico de Terminologia**. São Paulo: EDUSP, 2004.
- BENTIVOGLI, L.; BOCCO, A.; PIANTA, E. ArchiWordnet: integrating Wordnet with domain-specific knowledge. In: INTERNATIONAL GLOBAL WORDNET CONFERENCE, 2, 2004. **Proceedings...** Masaryk University, Brno, 2004. p. 39-47. Disponível em: <<http://www.fi.muni.cz/gwc2004/proc/101.pdf>>. Acesso em: 10 maio 2008.
- BERNHARD, D. Multilingual term extraction from domain-specific corpora using morphological Structure. In: CONFERENCE OF THE EUROPEAN CHAPTER OF THE ACL, 11, 2006. **Proceedings...** Trento, Italy, 2006, p. 171-174.
- BICK, E. **The parsing system PALAVRAS: automatic grammatical analysis of Portuguese in a constraint grammar framework**. 2000. PhD Thesis. Arhus University, 2000.
- BUITELAAR, P.; SACALEANU, B. Extending synsets with medical terms. In: INTERNATIONAL CONFERENCE ON GLOBAL WORDNET, 1, 2002. **Proceedings...** Mysore, India, 2002.
- CABRÉ, M. T. **La terminología: representación y comunicación: elementos para una teoría de base comunicativa y otros artículos**. Barcelona: Institut Universitari de Linguística Aplicada, 1999.
- ___; ESTOPÀ, R.; PALATRESI, J. V. Automatic term detection: a review of current systems, In: BOURIGAULT, D. *et al.* (Eds.). **Recent Advances in Computational Terminology**. Amsterdam & Philadelphia: John Benjamins Publishing Co., 2001, p. 53-87.
- CARROL, J. Parsing. In: MITKOV, R. (Ed.). **The Oxford handbook of computational linguistics**. Oxford, New York: Oxford University Express, 2004, cap. 12, p. 233-248.
- CEDERBERG, S.; WIDDOWS, D. Using LSA and noun coordination information to improve the precision and recall of automatic hyponymy extraction. In: CONFERENCE ON COMPUTATIONAL NATURAL LANGUAGE LEARNING, 11, 2003. **Proceedings...** Edmonton, Canada, 2003, p. 111-118.
- COLETI, J. S. **A terminologia da Nanociência e Nanotecnologia: descrição e análise morfológica**. São Carlos, 2009, 59p. Trabalho de conclusão de curso (Licenciatura em Letras) – Departamento de Letras, Universidade Federal de São Carlos, São Carlos, 2009.
- CUNHA, C.; CINTRA, L. **Nova gramática do português contemporâneo**. 3aed. Rio de Janeiro: Nova Fronteira, 2001.

- DI FELIPPO, A.; SOUZA, J.W.C. O projeto do corpus para a construção de uma wordnet terminológica. In: SHEPHERD, T., BERBER SARDINHA, T. e VEIRANO PINTO, M. (Orgs.) **Caminhos da Linguística de Corpus**. Campinas, SP: Mercado de Letras. ISBN 978-85-7591-158-7. *In press*
- ESTOPÁ BAGOT, R. **Extracció de terminologia: elements per a construcció d'un SEACUSE** (Sistema d'Extracció Automàtica de Candidats a Unitats de Significació Especialitzada). Tese de (Doutorado) – Universidade Pompeu Fabra, Barcelona, 1999.
- FELLBAUM, C (Ed.). **Wordnet: an electronic lexical database**. Ca, MA: MIT Press, 1998.
- HANKS, P. Lexicography. In: MITKOV, R. (Ed.). **The Oxford handbook of computational linguistics**. Oxford: Oxford University Press, 2004, p. 48-69.
- HAYES-ROTH, F. Expert systems. In: Shapiro, E. (Ed.). **Encyclopedia of artificial intelligence**. New York, Wiley, 1990, p. 287-298.
- JACQUEMIN, C.; BOURIGAULT, D. Term extraction and automatic indexing. In: MITKOV, R. (Ed.). **Handbook of Computational Linguistics**. Oxford University Press, 2003, p.599-615.
- LAROCA, M. **Manual de morfologia do português**. Campinas: Pontes; Juiz de Fora: Editora da UFJF, 2005.
- LOPES, L. *et al.* ExATOLp: An automatic tool for term extraction from Portuguese language corpora. In: LANGUAGE & TECHNOLOGY CONFERENCE LTC'09, 4, 2009, Polish. **Proceedings...** Polish, 2009, pp 427–431.
- LOPES, L. *et al.*, EXATOLP - an automatic tool for term extraction from Portuguese language corpora. In: INTERNATIONAL LANGUAGE AND TECHNOLOGY CONFERENCE (LTC), 4, 2009. **Proceedings ...** Poznam, Poland, 2009. p. 1-5.
- LOPES, L.; OLIVEIRA, L. H. M.; VIEIRA, R. Portuguese Term Extraction Methods: Comparing Linguistic and Statistical Approaches. In: PROPOR, 9, 2010, Porto Alegre. **Proceedings...** Porto Alegre, 2010, p. 1-6
- MAGNINI, B.; SPERANZA, M. Integrating generic and specialized wordnets. In: CONFERENCE ON RECENT ADVANCES IN NATURAL LANGUAGE, 2, 2001, Tzigov Chark. **Processing...** Tzigov Chark, Bulgaria. 2001.
- MILL, D. et al (Orgs.). **Polidocência na Educação a Distância: múltiplos enfoques**. São Carlos: Edufscar, 2010. *No prelo*.
- MITITELU, V.B. Automatic extraction of patterns displaying hyponym-hypernym co-occurrence from corpora. In: CENTRAL EUROPEAN STUDENT CONFERENCE IN LINGUISTICS, 1, 2006, Budapest. **Proceedings...** Budapest, Hungary, 2006.
- MORATO, J. *et al.* Wordnet applications. In: INTERNATIONAL GLOBAL WORDNET CONFERENCE, 2, 2004, Masaryk University. **Proceedings...** Masaryk University, Brno, 2004. p. 270-278.
- MORIN, E.; JACQUEMIN, C. Automatic acquisition and expansion of hypernym links. **Computer and the Humanities**, v. 38 (4), p. 343-362, 2004.
- PAUMIER, S. *Unitex user manual*. disponível em: <http://www-igm.univ-mlv.fr/~unitex>. 2002.
- PAZIENZA, M. T. *et al.* Terminology extraction: an analysis of linguistic and statistical approaches. **Studies in Fuzziness and Soft Computing**, v.185, p. 255-280, 2005.
- POPRAT, M. *et al.* Building a BioWordnet using Wordnet data structures and Wordnet's software infrastructure - a failure story. In: ACL WORKSHOP ON SOFTWARE ENGINEERING, TESTING, AND QUALITY ASSURANCE FOR NATURAL LANGUAGE PROCESSING, 2008, Ohio. **Proceedings...** Ohio, 2008. p.31-39.

- RIBEIRO JR., L.C. **OntoLP**: construção semi-automática de ontologias a partir de textos da língua portuguesa. São Leopoldo, 2008, 131p. Dissertação (Mestrado em Computação Aplicada) – Univ. do Vale do Rio dos Sinos, 2008.
- RIGAU, G. **Automatic acquisition of lexical knowledge from MRDs**. Tesis doctoral, Departament de Llenguatges i Sistemes Informàtics, UPC, Barcelona, 1998.
- ROCHA, L. C. A. **Estruturas morfológicas do português**. São Paulo: Editora WMF Martins Fontes, 2008.
- ROVENTINI, A.; Marinelli, R. Extending the Italian Wordnet with the specialized language of the maritime domain. In: INTERNATIONAL GLOBAL WORDNET CONFERENCE, 2, 2004. **Proceedings...** Masaryk University, Brno, 2004. p. 193-198.
- SAGER, J.C. **Curso práctico sobre el procesamiento de la Terminología**. Madrid: Fundación Germán Sánchez Ruipérez, 1993.
- SAGRI *et al.* Jur-Wordnet. In: INTERNATIONAL GLOBAL WORDNET CONFERENCE, 2, 2004. **Proceedings...** Masaryk University, Brno, 2004. p. 305-310.
- SANDMANN, A. J. **Formação de palavras no português brasileiro contemporâneo**. Curitiba: Scientia et Labor, 1989.
- SMITH, B.; FELLBAUM, C. Medical Wordnet: a new methodology for the construction and validation of information resources for consumer health. In: INTERNATIONAL CONFERENCE ON COMPUTATIONAL LINGUISTICS, 20, 2004. **Proceedings...** Geneva, 2004.
- SINCLAIR, J. Corpus and text: basic principles. In: Wynne, M. (Ed.). **Developing linguistic corpora: a guide to good practice**. Oxford: Oxbow Books, 2005. P.1-16. Disponível em: <www.ahds.ac.uk/creating/guides/linguistic-corpora/chapter1.htm>. Acesso em: 02 ago. 2010.
- SUÁREZ, M.; CABRÉ, M.T. La variación denominativa en los textos de especialidad: indicios lingüísticos para su recuperación automática. In: Simposio Iberoamericano de Terminología, 8, 2002. **Actas...** Cartagena de Indias, 2002. p.1-12.

APÊNDICE 1: Processos morfológicos de 59 unigramas da EaD.

	Processos sintáticos					Empréstimo	Outros processos	
	Derivação		Composição				Derivação regressiva	Palavra-valise
	Prefixal	Sufixal	Acronímica/Sigla	Coordenativa	Subordinativa			
andragogia	0	0	0		1	0	0	0
apostila	0	0	0			0	1	0
aprendente	0	1	0			0	0	0
aprendiz	0	0	0			0	0	0
assimilação	0	1	0			0	0	0
aulanet	0	0	0			0	0	1
ava	0	0	1			0	0	0
avaliador	0	1	0			0	0	0
blog	0	0	0			1	0	0
cederj	0	0	1			0	0	0
chat	0	0	0			1	0	0
cibercultura	0	0	0		1	0	0	1
ciberespaço	0	0	0		1	0	0	1
cognição	0	0	0			0	0	0
compartilhamento	0	1	0			0	0	0
construtivismo	0	1	0			0	0	0
conversação	0	1	0			0	0	0
cursista	0	1	0			0	0	0
designer	0	0	0			1	0	0
docência	0	0	0			0	0	0
docente	0	1	0			0	0	0
ead	0	0	1			0	0	0
educador	0	1	0			0	0	0
educando	0	1	0			0	0	0
eja	0	0	1			0	0	0
formador	0	1	0			0	0	0
gestor	0	1	0			0	0	0
hipermídia	1	0	0			0	0	0
hipertexto	1	0	0			0	0	0
instrutor	0	1	0			0	0	0
interação	1	0	0			0	0	0
interatividade	1	0	0			0	0	0
interdisciplinaridade	1	0	0			0	0	0
link	0	0	0			1	0	0
mediação	0	1	0			0	0	0
mediador	0	1	0			0	0	0
mentor	0	1	0			0	0	0
mobral	0	0	1			0	0	0
moodle	0	0	1			1	0	0
mudança	0	1	0			0	0	0
nead	0	0	1			0	0	0
orientador	0	1	0			0	0	0
pedagogia	0	0	0		1	0	0	0
pedagogo	0	0	0		1	0	0	0
portal	0	1	0			0	0	0
seed	0	0	1			0	0	0
subjetividade	0	1	0			0	0	0
teleconferência	1	0	0			0	0	0
telecurso	1	0	0			0	0	0
teleduc	1	0	0			0	0	1
telemática	0	0	0			0	0	1
tic	0	0	1			0	0	0
tutor	0	0	0			0	0	0
tutoria	0	1	0			0	0	0
tutorial	0	1	0			0	0	0
uab	0	0	1			0	0	0
veredas	0	0	0			0	0	0
videoaula	0	0	0		1	0	0	0
videoconferência	0	0	0		1	0	0	0
webct	0	0	1			1	0	0
TOTAL	8	22	11	2	5	6	1	5