**NomLex-PT Guidelines**

**V.2**

**03/10/2014**

**Cláudia Freitas, Valeria de Paiva**

**Agentive**: palavras agentivas são aquelas em que a paráfrase “o que X” (e X é Verbo) é possível

**Valores da coluna Agentivo: Y / N**

pintor: o que pinta → **Y**

construtor: o que constrói → **Y**

aspirador: o que aspira → **Y**

presidente: o que preside → **Y**

estudante: o que estuda → **Y**

utilização: o que utiliza → **N**

O critério para a identificação é formal, e o agentivo corresponde apenas à construção “o que X”. **Democratização** ou **globalização, por ex,** em “A democratização/globalização da mídia permitiu que...”, embora possa, de um ponto de vista, ser compatível com a ideia de agente, **não correspondem** a “*o que democratiza*”, ou “*o que globaliza*”, por isso, não são considerados agentive. Para esses casos, os agentive seriam “democratizador” e “globalizador”, esse último apenas como adjetivo.

**Animate/Inanimate**

**Valores da coluna Animate/Inanimate:**

**A / I / B (both – underspecified)**

Pintor → A

Aspirador → I

Utilização → I

Amor → B

Fabricante 🡪 B

**Lexicalized**

Consideramos palavras lexicalizadas - ou: palavras que passaram por um processo de lexicalização - aquelas em que o significado final não corresponde mais ao significado decorrente do processo morfológico. (“procuração” não corresponde ao ato de procurar; “coordenação” corresponde tanto ao ato de coordenar (a) quanto à entidade/pessoa(s) responsáveis por coordenar(b)

1. A coordenação da pesquisa foi feita pelo Departamento.
2. A coordenação decidiu não aceitar a proposta.

Nem sempre é fácil decidir se estamos diante de uma palavra lexicalizada, e frequentemente as palavras podem ter um equivalente lexicalizado em um domínio específico (ocorrência -> ato de ocorrer (ocorrência de chuvas) vs (*ocorrência* policial) ; transferência (psicanalítica).

Por ser difícil – talvez impossível – a decisão e reconhecimento de processos de lexicalização, e considerando a polissemia sistemática como um dos aspectos decorrentes do processo de nominalização, indicamos com 1 apenas os casos em que não é relevante (para a aplicação proposta) indicar o processo morfológico. Tais casos, são muito pouco frequentes.

Esta decisão difere ligeiramente de nossa primeira decisão, em que a intenção era diferenciar os casos de nominalização

- em que o processo morfológico é irrelevante (indicados com 1);

- em que só conseguimos perceber um sentido diretamente vinculado ao verbo /processo morfológico (indicados com 0);

- em que percebemos as duas situações anteriores (indicados com B)

Com a mudança de perspectiva, a diferença entre 0 e B deixa de ser relevante. Optamos por mantê-la na lista simplesmente para não perder a informação já indicada de que, para certas palavras, nos foi imediato o reconhecimento de um sentido da nominalização que nomeia entidades, e que pertencerá a um synset diferente no contexto da OpenWordNet-PT. Mas reforçamos que essa indicação entre 0 e B não é sistemática, e portanto a diferença relevante é entre (1) e os demais casos.

**Valores da coluna** **Lexicalized**:

**0 (non-lexicalized) / 1 (lexicalized only) / B (both)**

Publicação → B

Declaração → B

Coordenação → B

Organização→ B

Habitação → B

Civilização → B

Pintura → B

Vigilância → B

Militância → B

Anuência → B

Procuração 🡪 1

Restaurante 🡪 1

Como citar o trabalho:

Cláudia Freitas, Valeria de Paiva, Livy Real, Alexandre Rademaker, Anne de Araujo Correia da Silva. ***Extending a Lexicon of Portuguese Nominalizations with Data from Corpora***. In: 10th International Conference on Computational Processing of the Portuguese Language - PROPOR, 2014. Computational Processing of the Portuguese Language - Lecture Notes in Artificial Intelligence. Springer-Verlag, pp 114-124. <http://link.springer.com/chapter/10.1007%2F978-3-319-09761-9_12>