

**Bolsista:** Fernanda Baissvenger Pazinato.  
**Orientadora:** Maria José Bocorny Finatto.

## Explorando verbos em textos de divulgação sobre Pneumopatias Ocupacionais para leigos

### 1. Introdução

O Processamento de Linguagem Natural (PLN) consiste no desenvolvimento de ferramentas computacionais para realização de tarefas como extração e recuperação de informações de línguas. A integração entre PLN e Estudos Linguísticos é de grande utilidade para análise de textos, pois essa associação permite, entre outras coisas, fazer extração e análise de classes (verbos, substantivos, adjetivos) e extração de outros tipos de listas de palavras em grandes volumes de textos.

### 3. Metodologia

Seleção e preparação do *corpus*, composto por 57 textos de divulgação sobre PO, coletados a partir de diferentes *blogs*, para processamento.

Extração dos verbos específicos, via ExATOlp.

O ExATOlp, através de contraste entre o *corpus* de estudo e um *corpus* de referência, elege os verbos mais específicos do *corpus* de estudo. Essa extração foi o que forneceu os **21 verbos e locuções verbais**, considerados os principais, que analisamos neste trabalho.

Análise linguística da funcionalidade e do peso de opinião dos verbos.

1.	Ser: 36.659	8.	Apresentar: 4.11114	15.	Constituir: 3.5
2.	Conter: 6.36395	9.	Poder causar: 4.0759	16.	Poder ocorrer: 3.48168
3.	Poder provocar: 6	10.	Assemelhar: 4	17.	Haver: 3.39407
4.	Ter: 5.51178	11.	Inalar: 4	18.	Proibir: 3.33808
5.	Causar: 5.46353	12.	Poder produzir: 4	19.	Ajuizar: 3
6.	Subscrever: 5	13.	Poder ser classificar: 4	20.	Denominar: 3
7.	Estar: 4.85339	14.	Poder ser dividir: 4	21.	Poder originar: 3

- - modalidade epistêmica: crença do que seja.
- - modalidade epistêmica + alética: possibilidade de efetivação
- - modalidade deôntica: o deve ser (obrigação).

### 4. Resultados

Os verbos analisados não apresentam, por si só, posicionamento positivo nem negativo por parte do autor. A maior parte deles reflete o nível de conhecimento de quem escreve, via expressão de certeza e incerteza relacionadas às modalidades de Travaglia (1991). No *corpus*, os níveis de certeza e de possibilidade são os mais frequentemente manifestados. Sugere-se, então, que, com outros indicativos textuais, os verbos podem desempenhar papel importante em *Análise de Sentimento*. Tendo em vista que o público-alvo dos textos de divulgação sobre PO é, em geral, paciente ou trabalhador com pouca escolaridade, é fundamental trazer as informações de maneira clara e objetiva. Através da pesquisa, foi possível verificar o modo como os autores dos textos de divulgação veiculam o conhecimento especializado ao seu leitor. Desses resultados, percebem-se maneiras de aperfeiçoar as ferramentas computacionais para que sejam capazes de processar verbos associados a opiniões e níveis de certeza e incerteza.

### 5. Referências

- FERREIRA, E. *Análise de Sentimento em redes sociais utilizando influência das palavras*. Monografia de Bacharelado. Universidade Federal de Pernambuco. Recife, 2010, 69 f.
- FINATTO, M.; LOPES, L.; CIULLA, A. *Processamento de Linguagem Natural, Linguística de Corpus e Estudos Linguísticos: uma parceria bem-sucedida*. Domínios de Linguagem, Porto Alegre, v. 9, n. 5, p. 41-59, dez. 2015.
- LOPES, L.; FERNANDES, P.; VIEIRA, R.; FEDRIZZI, G. ExATO lp -- An Automatic Tool for Term Extraction from Portuguese Language Corpora. In: LTC'09 - 4th Language and Technology Conference, 2009, Poznan, 2009. Proceedings of the Fourth Language and Technology Conference. Poznan: Adam Mickiewicz University, 2009. p. 427-431.
- TRAVAGLIA, L. *Um estudo textual-discursivo do verbo no português do Brasil*. Tese de Doutorado. UNICAMP/ILEL Campinas, 1991, p.46-306.

### 2. Objetivos

Com base nas modalidades epistêmica, relacionada com a crença do locutor na verdade do que diz; alética, referente à verdade da proposição e na modalidade deôntica, associada às normas de conduta, apresentadas por (TRAVAGLIA, 1991): a) estudamos sobre a importância e a funcionalidade dos verbos em textos de divulgação científica para leigos, na área de Pneumopatias Ocupacionais (PO); b) buscamos contribuir com o estudo da *Análise de Sentimentos* em PLN, contando com **ExATOlp\*** (Extrator Automático de Termos para Ontologias em Língua Portuguesa) \*Ferramenta que “filtra” candidatos a palavras/expressões importantes em um texto segundo sua frequência e especificidade em um dado domínio/tema via cálculo de relevância/especificidade.

- “(...) reconhecem que todos os tipos de amianto, inclusive o crisotila, **causam** asbestose, mesotelioma e câncer do pulmão” - **modalidade epistêmica**.
- “A sílica, parte sólida da poeira vulcânica, em combinação com a umidade do ar, cria pequenos grãos que, se inalados, **podem causar** silicose” - **modalidade epistêmica + alética**.
- “Os trabalhadores que atuam em minas, construção civil e da indústria química **devem** ficar atentos ao contato diário com a sílica”. - **modalidade deôntica**.