

RESUMO PARA POSTER - SOCIOLINGUÍSTICA VARIACIONISTA

**PESQUISA EM SOCIOLINGUÍSTICA VARIACIONISTA A PARTIR DA
FERRAMENTA DE BUSCA CAPT 2.5**

João Franco Almeida (francojoao785@gmail.com)

Hebe Carvalho De Macedo (hebe@letras.ufc.br)

O presente resumo tem como objetivo apresentar a ferramenta computacional de Coleta Assistida de Padrões de Texto - CAPT 2.5 (Valério, 2025), produzida na Universidade Federal do Ceará, para auxílio da busca textual na pesquisa em sociolinguística. Compreendemos que a análise quantitativa de dados faz-se recurso necessário para pesquisas à Luz da Teoria da Variação e Mudança Linguística (Weinreich, Labov, Herzog, 2006 [1968]; Labov, 2008 [1972]), pois torna-se imprescindível medir o quão diferentes e o quão padronizados são os usos linguísticos dentro da comunidade (Mendes, 2010 apud Berlinck, Biazolli, 2018). Sendo assim, urge a necessidade de ferramentas metodológicas que auxiliem o pesquisador no desenvolvimento da pesquisa, especificamente, na edição, seleção, coleta e armazenamento dos dados linguísticos. Adotamos essa ferramenta na coleta de editoriais e artigos de opinião; e na busca dos modos subjuntivo e indicativo em orações encaixadas completivas, do projeto Retrato da variação do subjuntivo/indicativo na escrita jornalística mantido pelo PIBIC/UFC, que pretende estudar a alternância dos modos indicativo e

subjuntivo na escrita jornalística. A apresentação deste trabalho busca: (a) analisar o uso da ferramenta Capt 2.5 em pesquisa de cunho Variacionista; (b) expor o potencial da ferramenta para pesquisas sociolinguísticas demonstrando seus recursos específicos e vantajosos. Baseia-se em auxiliar no processamento de corpus linguístico, com capacidades de edição, coleta, seleção e categorização de amostras filtradas por expressões regulares. O software pode ser adotado em mineração de textos para a elaboração de grandes corpora, especificamente, em bases de dados da Sociolinguística Variacionista em interface com a Linguística de Corpus. Entendemos que desde os anos 60 os corpora eletrônicos têm sido considerados um valioso recurso para o estudo linguístico (Huston, 2002; Mcenery, e Gabrielatos, 2006 apud Sarmiento, 2010) e o uso do Capt 2.5 na pesquisa Sociolinguística tem a vantagem de (1) trabalhar diretamente com textos e transcrições possibilitando consultas rápidas de morfemas/vocábulos/frases; (2) realização da fraseologia, vocabulário, similaridade e legibilidade do corpus; (3) consultas específicas com os usos de Expressões Regulares. Tivemos como resultado a constituição de um corpus com 2782 artigos de opinião dos jornais escolhidos publicados em 2024, e 800 editoriais publicados também em 2024, no qual foi possível acessar padrões sintáticos-semânticos, as associações que as palavras estabelecem entre si, suas colocações, as variações das unidades lexicais, entre outras operações propiciadas pela ferramenta e encontradas nos textos do corpus.

Palavras-chave: sociolinguística variacionista; ferramenta de busca; linguística de corpus.