



# Pontifícia Universidade Católica de São Paulo

Programa de Estudos Pós-Graduados em Linguística Aplicada  
e Estudos da Linguagem - LAEL

## DISCIPLINA OPTATIVA: ANOTAÇÃO DE CORPUS

**Linha de Pesquisa:** Linguagem, Educação e Tecnologia

**Responsável:** Prof. Dr. Tony Berber Sardinha

**Créditos:** 03

**Dia/Horário:** Quinta - feira das 16h00 às 19h00

**Semestre/Ano:** 2º/2012

**Nível:** M/D

### Ementa

Esta disciplina destina-se a ensinar como fazer anotação de corpus eletrônico. Anotação é o processo de identificar e registrar, em um corpus, informação relativa a diversos tipos de análise linguística. A anotação deve seguir critérios consistentes de codificação para ser recuperável por máquina. Os tipos de análise a serem enfocados incluem: gramatical, semântica, metáfora conceptual, metáfora linguística e pacotes lexicais. Serão ainda tratados aspectos de análise estatística de dados linguísticos a partir da anotação do corpus. Ferramentas de análise, automática e manual, serão exploradas para auxiliar na tarefa de anotação.

### Bibliografia

BERBER SARDINHA, T. *Linguística de Corpus*. São Paulo: Manole, 2004. 410 p.

BIBER, D.; CONRAD, S. *Register, genre, and style*. Cambridge ; New York: Cambridge University Press, 2009. ix, 344 p.

LÜDELING, A.; KYTÖ, M. *Corpus Linguistics -- An International Handbook*. Berlin / New York: Walter de Gruyter, 2008.

O'KEEFFE, A.; MCCARTHY, M. *The Routledge handbook of corpus linguistics*. 1st. ed. London ; New York, NY: Routledge, 2010. xxviii, 682 p.

SHEPHERD, T.; BERBER SARDINHA, T.; VEIRANO PINTO, M. (Orgs.) *Caminhos da Linguística de Corpus*. Campinas, SP: Mercado de Letras. 2010