



E379

### **ANÁLISE ESTATÍSTICA E IDENTIFICAÇÃO DE CLASSES RÍTMICAS EM GRANDE CORPORA DE FALA**

Daniela Bento Fonsechi (Bolsista PIBIC/CNPq) e Profa. Dra. Nancy Lopes Garcia (Orientadora), Instituto de Matemática, Estatística e Computação Científica - IMECC, UNICAMP

O objetivo do projeto é a identificação das classes rítmicas de um conjunto de línguas (Português Europeu – PE, Português Brasileiro – PB, Inglês, Holandês, Polonês, Francês, Italiano, Espanhol, Catalão e Japonês). A caracterização foi feita através do estudo da função sonoridade, continuando o trabalho feito em Galves, Garcia, Duarte e Galves (2002), realizando a identificação através da modelagem paramétrica da função sonoridade, através de um grande conjunto de dados. Primeiramente analisamos dados do PB e percebemos que havia erros nas gravações (ruídos, falhas etc), portanto a classificação do PB não pode ser realizada. A abordagem não paramétrica foi utilizada para comparar se o PE tem mesma distribuição que as línguas supostamente acentuais (Inglês, Holandês e Polonês), nessa etapa percebemos que havíamos cometido um erro na abordagem paramétrica feita em 2004. Precisamos refazer o teste paramétrico, eliminando o erro encontrado para que assim possamos comparar o dois métodos (paramétrico e não paramétrico). A comparação não paramétrica será feita baseando-se em uma “proxy” da distância de Kullback-Leibler e estimação de densidade utilizando splines como feito em Dias e Garcia (2004).

Inferência paramétrica - Inferência não paramétrica - Teste da razão de verossimilhança