

Atlas prosódico multimédia: curvas de uma trajectória

*Lurdes de Castro Moutinho, Rosa Lídia Coimbra,
Suzana Secca Ruivo & Urbana Pereira Bendiha*

Centro de Investigação de Línguas e Culturas – Universidade de Aveiro

O *Atlas Prosódico Multimédia das Línguas Românicas* encontra-se em fase inicial de realização, envolvendo diversas universidades europeias cujo denominador comum é o interesse pela investigação sobre a estrutura prosódica dos falares românicos.

Em Portugal, o projecto é assegurado por um grupo de investigadoras da Universidade de Aveiro, em parceria com o Centre de Dialectologie – Université Stendhal de Grenoble (M. Contini; A. Romano), a quem cabe a coordenação do projecto global.

Objectivos

O objectivo fundamental da pesquisa é o da construção de um instrumento multimédia onde se caracterizam as variações linguísticas a nível da prosódia nos diversos falares românicos. No que respeita à variação prosódica do Português Europeu (PE), o Atlas contemplará registos representativos de diversas regiões de Portugal Continental, nomeadamente, numa fase preliminar, Minho, Trás-os-Montes, Beira Interior, Alentejo, Algarve e Beira Litoral.

A longo prazo, poder-se-á encarar a hipótese de estender esta pesquisa, quer a outras variantes do PE, quer, eventualmente, ao Português do Brasil.

Pretende-se também, desta forma, contribuir para um conhecimento mais aprofundado da variação prosódica da Língua Portuguesa. Para além disso, a disponibilização *online* do corpus poderá vir a possibilitar futuras investigações a diversos níveis da análise linguística.

O corpus experimental

Antes da realização das gravações cujos resultados de análise integrarão o Atlas Multimédia, procedeu-se à recolha de um corpus experimental, já parcialmente analisado, que obedece aos critérios definidos para a presente pesquisa. Este trabalho prévio tem como objectivo testar não só a constituição linguística do corpus, mas também a adequação da metodologia proposta ao estudo da prosódia do

Português. É um exemplo retirado deste pré-inquérito que será utilizado para explicar a metodologia de análise.

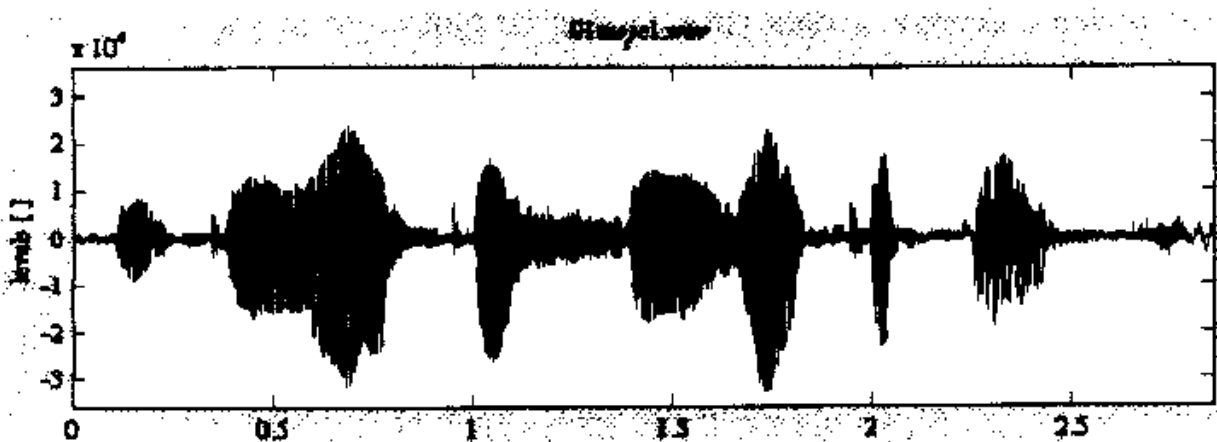
De um modo geral, o corpus deve obedecer aos seguintes critérios:

Com base em frases de estrutura SVO, criam-se sucessivas expansões de modo a permitir que os diferentes vocábulos – que devem contemplar as diversas estruturas acentuais (oxítone, paroxítone e proparoxítone) do PE – ocorram em diferentes posições frásicas. A constituição do corpus obedece a restrições, nomeadamente, de tipo fonético e sintáctico, devendo incluir, inicialmente, frases declarativas e suas correspondentes interrogativas, podendo vir a ser completado com outros tipos de frase.

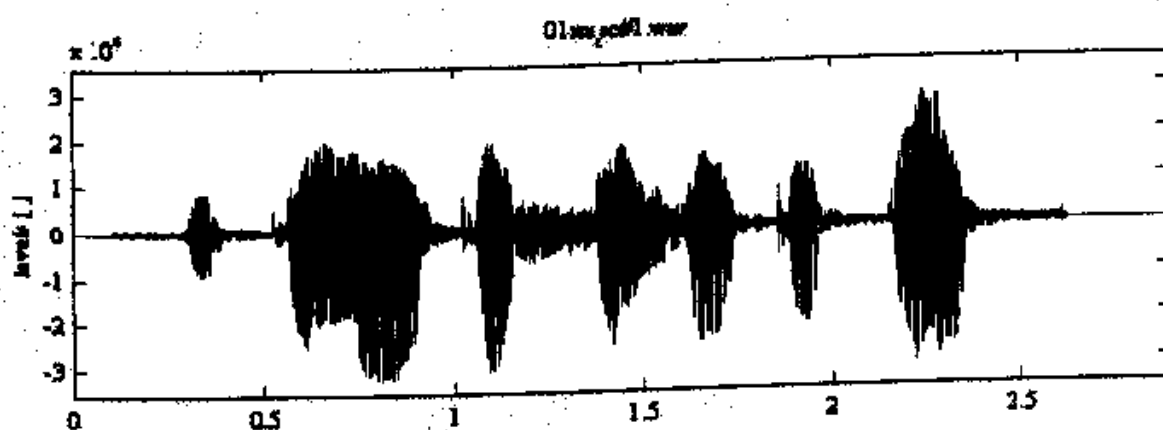
Métodos

Após gravação e pré-tratamento do sinal acústico nos programas GoldWave e CoolEdit, os dados recolhidos são analisados no programa MatLab com aplicações desenvolvidas especificamente para esse fim. No momento da sua gravação, evita-se a situação de leitura, já que a ela se associa uma prosódia com características específicas.

O estudo incidirá sobre os segmentos vocálicos realizados, por se considerar serem estes os portadores de informação mais pertinente a nível prosódico e, também, por ser esta a metodologia utilizada por todas as equipas europeias que integram o projecto. Os parâmetros retidos para análise são os de duração, intensidade e variação de F_0 das vogais previamente seleccionadas. É com base em oscilogramas, como os representados em A.1 (frase declarativa) e A.2 (frase interrogativa), que, no programa MatLab, se realiza a segmentação do sinal acústico, primeiro momento de análise de cada uma das frases seleccionadas.

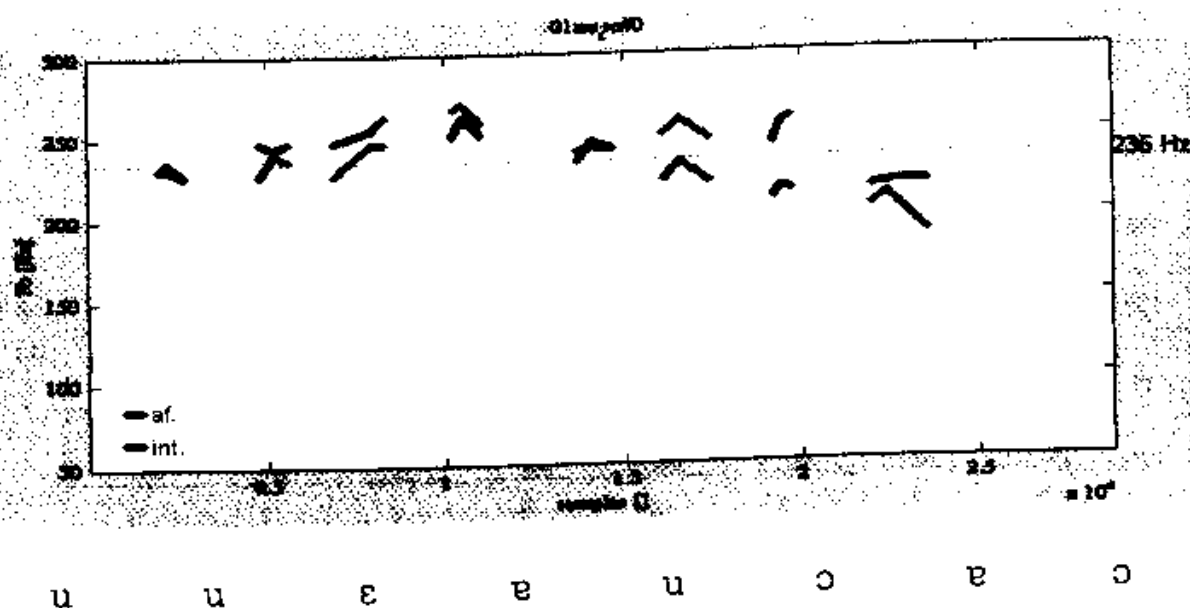


Doc. A.1 [utu'nekusu'zokɐ'pɔt]



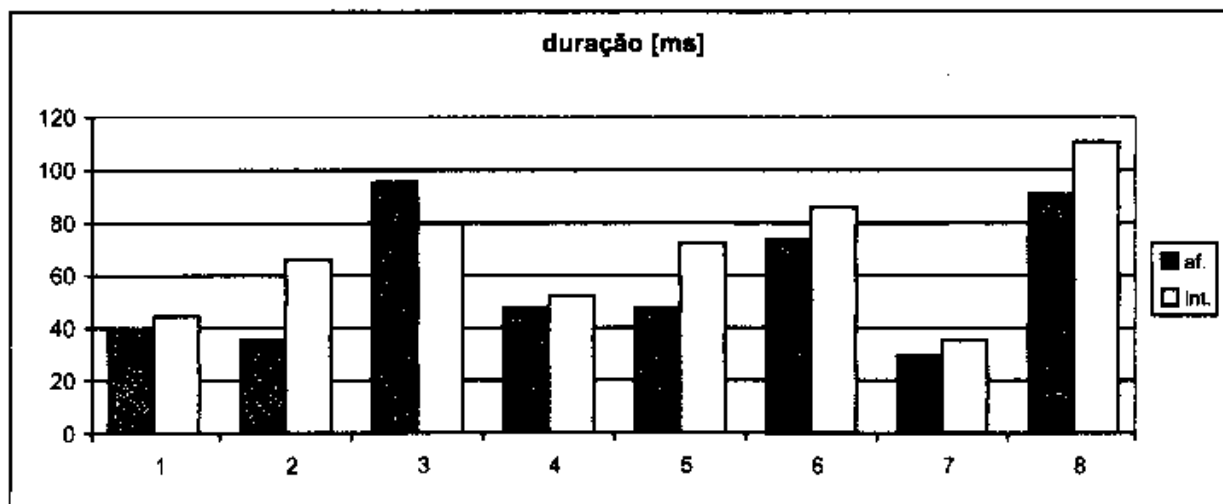
Doc. A.2 [utu'nekesu'zoke'pot]

Os dados resultantes dessa análise permitem a realização de documentos como os representados em B.1, B.2 E B.3 e servirão de base à interpretação linguística dos resultados obtidos.



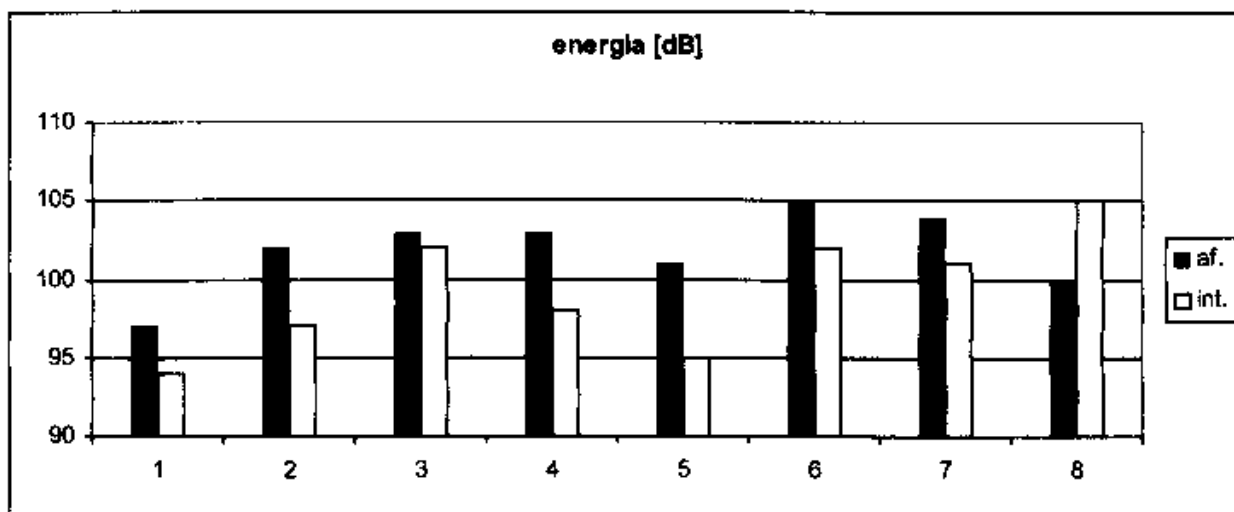
Doc.B.1

O documento B.1 permite a visualização da variação da frequência fundamental (F_0), por referência à frequência média do informante.



u u ε ρ u o ρ o

Doc. B.2



u u ε ρ u o ρ o

Doc. B.3

Os documentos B.2 e B.3 mostram, respectivamente, a variação da duração e da energia das frases analisadas, tendo sempre por referência somente os segmentos vocálicos.

As futuras gravações contemplarão também situações de fala espontânea em moldes a definir, posteriormente, e de acordo com as restantes equipas envolvidas no projecto.

Todas as gravações do inquérito final virão a ser realizadas no terreno utilizando um gravador DAT.

Com base nos dados da análise prosódica efectuada anteriormente, o próprio programa possibilitará a modelização da variação de F_0 e a restituição auditiva de cada frase, sobre a base de informações prosódicas que a caracterizam. Este procedimento facilitará, por um lado, uma comparação entre as estruturas prosódicas das variedades do Português e, por outro, a comparação entre estas últimas e o mesmo tipo de estruturas nas variedades românicas.

A informação global veiculada pelo Atlas e a sua estrutura interactiva permitirão a observação de características prosódicas multiparamétricas de cada uma das frases analisadas que, para uma maior facilidade de leitura, se farão sempre acompanhar de comentários linguísticos sucintos.

Através da implementação do projecto na Internet, a comunidade científica internacional terá acesso a uma base de dados prosódicos multilingue, a primeira consagrada ao domínio dos falares românicos.