

## **MESA-REDONDA SOBRE CORPORA LINGUÍSTICOS:**

### **APRESENTAÇÃO**

**Maria Fernanda Bacelar do Nascimento**  
**Centro de Linguística da Universidade de Lisboa**

O valor dos *corpora* como lugar de observação sistemática dos dados e como base de experimentação de hipóteses linguísticas tem sido fortemente reconhecido nas últimas décadas, no que respeita quer à língua falada, quer à língua escrita.

Com o extraordinário desenvolvimento da informática e do poder dos computadores, têm sido possíveis o armazenamento, processamento e codificação de grandes *corpora* cuja exploração automática permite a investigação de fenómenos linguísticos empiricamente observáveis, a uma escala anteriormente não imaginada.

Vão participar nesta mesa-redonda sobre *corpora* que ocupará a manhã e a tarde, investigadores que, de há muito, trabalham nesta área, tanto em Portugal como no estrangeiro.

O Centro de Linguística da Universidade de Lisboa (CLUL), desde o final da década de 60 (quando ainda tinha o nome de Centro de Estudos Filológicos), por iniciativa do Professor Lindley Cintra, começou a desenvolver projectos que visavam a recolha e o estudo de *corpora* linguísticos. Assim, esta instituição detém, hoje, um património importantíssimo de *corpora* portugueses nas suas variantes regionais e nacionais, que, no conjunto, são constituídos por um Arquivo

Sonoro de mais de 3.600 horas de gravação e por um *corpus* informatizado, de português oral e escrito, de cerca de 30 milhões de palavras.

Os trabalhos do CLUL que na parte da tarde serão apresentados continuam, pois, uma tradição deste Centro e, simultaneamente, integram uma das correntes dominantes da Linguística actual.

Dadas as funções que desempenho como coordenadora de um desses projectos do CLUL, o "*Corpus* de Referência do Português Contemporâneo", a Associação Portuguesa de Linguística (APL), entendeu convidar-me a organizar esta mesa-redonda, trabalho feito conjuntamente com o meu colega José Bettencourt Gonçalves.

Durante a tarde apresentam-se alguns dos *corpora* de língua portuguesa que, pelo conhecimento que temos da sua dimensão e importância institucional, nos pareceu importante dar a conhecer. Se, por falta de conhecimento em tempo útil, alguns autores de *corpora* não tiverem sido contactados, pensamos que terão hoje uma boa oportunidade para aqui os darem a conhecer.

Durante a manhã, participam nos trabalhos Claire Blanche-Benveniste da Universidade da Provença (Aix-Marselha) que falará de *corpora* orais, John Sinclair da Universidade de Birmingham que falará da Tipologia de *Corpora*, Nicoletta Calzolari do Instituto de Linguística Computacional de Pisa que tratará da importância dos *corpora* em Lexicografia e em Engenharia Linguística, e, finalmente, José Soler da Direcção-Geral XIII da Comissão das Comunidades Europeias que traçará um panorama das infraestruturas linguísticas na Europa.

Antes destas intervenções, vai ser lida uma comunicação de Paul Rivenc da Universidade de Toulouse-le-Mirail, um dos pioneiros na constituição de *corpora* orais e também um dos autores do Francês Fundamental que, não podendo estar presente, não quis deixar de enviar o seu contributo para esta mesa-redonda.