

## SEÇÃO ENTREVISTA

### PORTAL DA LÍNGUA PORTUGUESA: UM RECURSO LINGUÍSTICO

Sílvia Barbosa é a professora investigadora do ILTEC (Instituto de Linguística Teórica e Computacional) e integrante da equipe que elaborou o Portal da Língua Portuguesa. Nesta seção, ela fala sobre uma importante ferramenta de apoio ao universo acadêmico: a elaboração e as funcionalidades do Portal da Língua Portuguesa desenvolvido em Lisboa, Portugal.

**Por Alexandra F. Muller**

#### **Entrelinhas: O que é e qual o objetivo do Portal da Língua Portuguesa?**

O Portal da Língua Portuguesa é um repositório organizado de recursos linguísticos. Pretende ser orientado tanto para o público em geral como para a comunidade científica, servindo de apoio a quem trabalha com a língua portuguesa e a todos os que têm interesse ou dúvidas sobre o seu funcionamento. Todo o conteúdo do Portal é de livre acesso e está em constante desenvolvimento.

#### **Entrelinhas: Qual o material disponível no Portal?**

O recurso central do Portal da Língua Portuguesa é o *Vocabulário Ortográfico do Português* (VOP), baseado numa base de dados chamada *MorDebe*, na qual é gerida toda a informação lexical. Por via do VOP, o Portal disponibiliza informações relativas à flexão e à ortografia das palavras. Além de servir de base para a criação de recursos lexicais relacionados, o VOP funciona como um guia ortográfico, já adaptado às novas regras ortográficas, permitindo ao utilizador saber se uma determinada forma existe e como se escreve. Para uma descrição detalhada do VOP, consulte [www.portaldalinguaportuguesa.org/vop.html](http://www.portaldalinguaportuguesa.org/vop.html)

Além do VOP, o Portal contém também bases de dados que o complementam, fornecendo informações que são apresentadas em cada entrada lexical. Estes recursos foram desenvolvidos pela nossa equipe, e apenas estão disponíveis no Portal. Alguns exemplos são os seguintes:

\* *Dicionário de Gentílicos e Topónimos*, que contém pares de topónimos (nomes de terras, como Lisboa ou Pernambuco) e gentílicos (nomes das pessoas que nasceram ou vivem nessas terras, como lisboeta ou pernambucano);

\* *Dicionário de Estrangeirismos*, constituído por palavras de outras línguas que os lusófonos usam (como aftershave e diesel);

\* *Dicionário de Nomes Deverbais*, que lista os nomes associados a cada verbo (como digestão para digerir);

\* *Vocabulário de Mudança*, que lista todas as palavras modificadas pelo Acordo Ortográfico da Língua Portuguesa (1990) e que apresenta as principais mudanças nas regras de escrita;

\* *Divisão Silábica*, em que se mostra como dividir em sílabas (ortográficas) as palavras do português.

O Portal disponibiliza ainda o texto dos acordos ortográficos da língua portuguesa e das nomenclaturas gramaticais em vigor. Uma nova secção apresenta, neste momento, alguns jogos que tornam divertido conhecer o léxico do português.

Os dados do Portal são constantemente atualizados e aumentados.

### **Entrelinhas: Como nasceu o projeto do Portal da Língua Portuguesa?**

O Portal foi lançado como projeto sem fins lucrativos pelo ILTEC (Instituto de Linguística Teórica e Computacional) em fevereiro de 2007, com o apoio do Ministério da Ciência, Tecnologia e Ensino Superior de Portugal. Começou por ser um substituto da página da *MorDebe*, o nome que então tinha a base de dados lexical do ILTEC. Posteriormente, foram sendo disponibilizadas várias outras fontes e ferramentas, tais como o texto completo dos vários acordos ortográficos, os textos das nomenclaturas gramaticais do Brasil e de Portugal, e, entre outros, dicionários de gentílicos e topónimos, de estrangeirismos e de relações derivacionais e um vocabulário com todas as palavras cuja grafia foi afetada com o Acordo Ortográfico de 1990. Em 2010, o Portal passou a albergar algumas ferramentas que visam facilitar a aplicação do Acordo Ortográfico de 1990, desenvolvidas com o apoio do Fundo da Língua Portuguesa.

Desde janeiro de 2010, em resultado de um projeto financiado pelo Fundo da Língua Portuguesa, o Portal passou a disponibilizar, progressivamente, os resultados do Vocabulário Ortográfico do Português (VOP). Este projeto visa a correção, expansão e adaptação ao Acordo Ortográfico de 1990 da *MorDebe*, a base de dados lexical que serviu de base à criação do Portal e que o ILTEC vem desenvolvendo desde 2004. O projeto do VOP conta com cerca de 200 mil entradas, tendo como principal fonte o Vocabulário da Língua Portuguesa de Rebelo Gonçalves, o Dicionário da Academia de Ciências de Lisboa, o Dicionário da Porto Editora e o Dicionário de Antônio Houaiss na versão do português europeu. Em simultâneo, são desenvolvidas ferramentas informáticas de adaptação ao Acordo Ortográfico de 1990, como um conversor de ficheiros de texto e verificadores ortográficos para ferramentas informáticas de edição de texto, todas a serem disponibilizadas de forma gratuita no Portal.

O léxico disponibilizado no Portal é uma coleção corrigida e homogeneizada das nomenclaturas de alguns dos principais dicionários portugueses. Com a feitura do Vocabulário Ortográfico do Português, o Vocabulário da Língua Portuguesa de Rebelo Gonçalves converteu-se na sua principal fonte. Um número reduzido de palavras incluídas na base de dados, menos de 1% do total, são palavras de elevada frequência com origem em órgãos de comunicação social, sendo a sua inclusão alvo de tratamento lexicográfico prévio. A base de dados tem cerca de 1,35 milhões de formas flexionadas.

**Entrelinhas: De um modo geral, qual o público-alvo do Portal da Língua Portuguesa?**

O Portal da Língua Portuguesa disponibiliza, de forma imediata, acessível e simples, um conjunto de recursos que permitem, ao utilizador comum da Internet, o acesso a informações relativas à língua portuguesa, nas suas diferentes variedades. Pretende ser uma ferramenta de livre acesso, direccionada para alunos e professores, para profissionais de diversas áreas e, ainda, para todos os que têm interesse e curiosidade pelo funcionamento da língua portuguesa. Refletindo alguns resultados de investigação, o Portal estabelece uma ponte entre a comunidade científica e o público em geral.