

Procura-PALavras (P-PAL): Um *software* para avaliação de palavras do Português Europeu



Universidade do Minho



ESEIG | POLITÉCNICO DO PORTO

¹Ana Paula Soares*, ¹Montserrat Comesaña, ²José João Dias de Almeida, ³Alberto Simões & ⁴Álvaro Iriarte Sanromán

¹Escola de Psicologia, Universidade do Minho, ²Departamento de Informática, Universidade do Minho, ³Escola Superior de Estudos Industriais e de Gestão, Instituto Politécnico do Porto, ⁴Departamento de Estudos Portugueses, Universidade do Minho

O que é?

- Adaptação do *software* N-Watch (Davis, 2005) para o Português Europeu.

Para que servirá?

- Analisar listas de palavras previamente definidas pelo investigador nos índices psicolinguísticos considerados relevantes para a investigação e disponíveis no menu de análises.
- Obter listas de palavras que obedecem aos parâmetros psicolinguísticos previamente definidos pelo investigador e disponíveis no menu de análises.

Índices de frequência de uso:

- Frequência de uso total (absoluta, por milhão e logarítmica - base 10)
- Frequência escrita (absoluta, por milhão e logarítmica - base 10)
- Frequência falada (absoluta, por milhão e logarítmica - base 10)

Índices morfosintácticos:

- Categoria sintáctica
- Frequência de categoria sintáctica (substantivos, verbos, adjetivos), classe (aberta/fechada), *lemma* e *wordforms* (para palavras cuja inflexão de género e número derivam de radical distinto) e para palavras funcionais (artigos, pronomes e contrações)

Índices ortográficos:

- Extensão (número de letras)
- Estrutura C_V
- Frequência de bigrama (tipo e *token* por posição e extensão da palavra)
- Homógrafos (número e lista)
- Ponto de unicidade

Índices silábicos:

- Silibificação ortográfica
- Extensão silábica ortográfica (número de sílabas)
- Silibificação fonológica
- Extensão silábica fonológica (número de sílabas fonológicas)
- Frequência silábica (tipo e *token* por posição: 1ª, 2ª e última sílabas)
- Frequência de bigrama (tipo e *token* por posição e extensão da palavra)

Índices semânticos:

- Distância semântica entre palavras
- Lista das palavras que mais co-ocorrem com uma dada palavra

Índices subjectivos:

- Familiaridade
- Idade de Aquisição (AOA)
- Imaginabilidade
- Valência
- Concreteza
- Controlo

Índices de vizinhança:

- Vizinhos ortográficos por substituição (número e lista por posição, frequência média, máxima e mínima dos vizinhos)
- Vizinhos ortográficos por adição e subtração (número e lista por posição, frequência média, máxima e mínima dos vizinhos)
- Vizinhos fonológicos por substituição (número e lista por posição, frequência média, máxima e mínima dos vizinhos)
- Vizinhos fonológicos por adição e subtração (número e lista por posição, frequência média, máxima e mínima dos vizinhos)
- Vizinhos por transposição de pares de letras adjacentes e não adjacentes

Índices fonológicos:

- Transcrição fonética (máximo 3 alternativas)
- Extensão (número de fonemas)
- Estrutura C_V (máximo 3 alternativas)
- Padrão de acento
- *Schwas* (número e posição)
- Regularidade
- Primeiro fonema
- Frequência de bifone (tipo e *token* por posição e extensão da palavra)
- Homófonos (número e lista)
- Ponto de unicidade

ITEM	CELEX	LEN_S	LEN_P	LEN_L	CV_O	IMG
house	558,72	1	3	5	CVVCV	606
dog	71,73	1	3	3	CVC	636
pen	19,44	1	3	3	CVC	576
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						

Qual a vantagem?

- Facilitar a tarefa de seleccionar estímulos verbais para a investigação a partir da utilização de uma aplicação informática multi-plataforma e de fácil utilização.

Referências

- Davis, C. J. (2005). N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods*, 37(1), 65-70.

Para mais informações contactar: asoares@psi.uminho.pt

Escola de Psicologia, Universidade do Minho, *campus* de Gualtar, 4710-057 BRAGA

5º Encontro Nacional da Associação Portuguesa de Psicologia Experimental (APPE) | 26-27 Março, 2010 | Braga: Universidade do Minho, Portugal



Projecto de Investigação PTDC/PSI/PCO/104679/2008 financiado pela Fundação para a Ciência e a Tecnologia (FCT) e co-financiado pelo FEDER (Fundo Europeu de Desenvolvimento Regional) no âmbito do programas COMPETE (Programa Operacional Factores de Competitividade) e do QREN (Quadro de Referência Estratégico Nacional).