

Lexical Database of Estimated Age-of-Acquisition and Familiarity Norms for 408 Portuguese Nouns

L. Gonzaga, L. Meireles & S. Vicente
Laboratório de Fala, Faculdade de Psicologia e de Ciências da Educação
Universidade do Porto

December 2007

Based on previous work (Vicente, 2003, 2004) we created a corpus of estimated age-of-acquisition (AOA) and familiarity ratings for 408 Portuguese dissyllabic nouns selected from the Porlex, a computerized lexical database in European Portuguese representative of the adult lexicon (Gomes & Castro, 2003; total of 29.238 lexical entries).

All the nouns are dissyllabic words CV.CV with first syllable-stress, having 4 phonemes long (between 4- and 5- letters). None of them has the vowel schwa in the final position (e.g., *banho*, *fala*, *vaca*; that is, *bath*, *speech*, *cow*). They were selected from a subgroup of Portuguese dissyllabic nouns listed in Porlex because they are a good choice as verbal materials to use in experimental tasks regarding word recognition in European Portuguese. Indeed, unlike monosyllabic words, they are frequent in the adult and child lexicons (e.g., the Viana Lexicon of the productions of the 3-, 4-, and 5-year olds; Vicente, 2003; Vicente & Castro, in preparation). Also, as Vicente, Castro, and Walley (2003) showed, in Portuguese the densest words are the shortest with 4 phonemes long, which represent about 21% of the Viana database and 5% of the Porlex. This AOA & Familiarity database is in Excel format (Microsoft Corporation, 2004), contains 408 different lexical entries and provides 11 types of information (see Table 1). AOA and Familiarity data were collected in a study conducted with 200 university students (Gonzaga, Meireles, & Vicente, 2007). Other variables such as frequency and neighborhood density were also considered. Frequency counts were imported from a frequency database obtained from written material, Corlex (Nascimento et al, s.d.). Neighborhood density values were imported from the Porlex.

This standardized pool of 408 words has been used to select verbal materials to design experimental tasks in different studies concerning spoken and written word processing (Gonzaga, & Vicente, ongoing; Meireles, & Vicente, ongoing; Vicente, Lima, & Gonzaga, 2006; Vicente, Gonzaga, & Lima, 2006; Vicente & Castro, 2007; Vicente, & Castro, submitted). A paper presenting this database is in preparation.

Table 1 – Information available in the AOA & Familiarity database for each of the 408 words

Variable	Code Name	Type ^a	Content
1	#	I	Number
2	Word	S	Orthographic Wordform
3	Fot	S	Phonetic Wordform
6	NSilO	I	Orthographic Syllable Length
7	NSilF	I	Phonological Syllable Length
8	Nfom	I	Phonemic Length
9	DO	I	Orthographic Density
10	DFot	I	Phonetic Density
11	Freq	I	Frequency
12	AOA	I	Age-of-Acquisition
13	Fam	I	Familiarity

Note. Data from columns 3 to 10 was imported from Porlex. Data in column 11 was imported from Corlex.

^a Entry type: I = integer; S = string

References:

- Bacelar do Nascimento, M. F., Casteleiro, J. M., Marques, M. L. G., Barreto, F., Amaro, R. (s.d.). *Corlex: Léxico de frequências do português* [Base lexical]. Disponível em [http:// www.clul.ul.pt](http://www.clul.ul.pt)
- Gomes, I., & Castro, S. L. (2003). Porlex, a lexical database in European Portuguese. *Psychologica*, 32, 91-108.
- Gonzaga, L. O., Meireles, L. P., & Vicente, S. G. (2007, Marsh). *Corpus de Idade de Aquisição e de Familiaridade Subjectiva para o Português Europeu* [Age-of-Acquisition and Familiarity data for European Portuguese]. Poster presented at the 2º Encontro Nacional da Associação Portuguesa de Psicologia Experimental (APPE), Universidade do Porto, Porto, Portugal.
- Vicente, S. (2003). *Reconhecimento de palavras faladas: Abordagem desenvolvimental em Português Europeu* (Vols 1-2). Tese de Doutoramento não publicada, Universidade do Porto, Porto, Portugal.
- Vicente, S., Castro, S. L., & Walley, A. (2003). A developmental analysis of similarity neighborhoods in European Portuguese. *Journal of Portuguese Linguistics*, 2, 115-133.
- Vicente, S. (2004). Estimativas de Idade-de-Aquisição por adolescentes e jovens adultos [Age-of-Acquisition estimates by adolescents and adults] [Abstract]. *II Congresso Hispano-Português de Psicologia*, 207.
- Vicente, S. G., Gonzaga, L. O., & Lima, C. F. (2006). Efeitos da frequência, densidade de vizinhança e frequência de vizinhança no reconhecimento de palavras faladas em Português Europeu [Frequency, neighborhood density and neighborhood frequency effects in the recognition of words spoken in European Portuguese]. In N. R. Santos, M. L. Lima, M. M. Melo, A. A. Candeias, M. L. Grácio, & A. A. Calado (Orgs.), *Proceedings of the VI National Symposium of Investigation in Psychology* [CD, ISBN 978-972-98136-9-6]. Évora: Universidade de Évora.
- Vicente, S. G., & Castro, S. L. (2007). Neighborhood density effects on spoken word recognition by European Portuguese seven grade students [Abstract]. *XIII ESCOP 2007 Conference*, 77.
- Vicente, S. G., Castro, S. L., & Walley, A. (submitted). Neighborhood density and the special role of age-of-acquisition in spoken word recognition in Portuguese. *XXIX International Congress of Psychology, ICP*.