

Polysemous nominalizations in *-ção* and *-mento* in Portuguese: Correlating frequency with semantic opacity

Andrew Swearingen
Instituto de Linguística Teórica e Computacional

In Portuguese, suffixes *-ção* and *-mento* are functionally identical. Both are productive, typically involve the creation of nouns from verbs, and, (as across Romance generally) are attributed a prototypical meaning of ‘act or result of V-ing’ (Corbin 1987, Kerleroux 2008). The prototypical meanings of deverbal nouns such as *animação* ‘animation’ ↔ *animar* ‘animate’ and *andamento* ‘going’ ↔ *andar* ‘go’ exemplify such derivational transparency with their respective verbal bases. On the other hand, nouns such as *constituição* ‘constitution’ or *orçamento* ‘budget’ though expressible in a transparent sense of ‘action or result of situating/budgeting’, are overwhelmingly used in a more opaque sense as concrete entities (‘resultative objects’ Grimshaw 1990). The fact that deverbal nouns such as *constituição* and *orçamento* occur more frequently in corpora than their bases while *animação* and *andamento* exhibit the opposite tendency corroborates previous studies correlating the (relative) frequency of derivations in relation to their bases with semantic opacity (Bybee 1985, Hay 2001, Hay and Baayen 2002). However, these studies appear to reflect upon derivations whose meanings are *either* transparent *or* opaque rather than polysemous, i.e. derivations with *both* transparent and opaque meanings. In Portuguese, this is the norm for nominalizations in *-ção* and *-mento*. One question, therefore, is whether transparency and opacity correlate with relative frequency in relation to the base in the same way that exclusively opaque or transparent derivations do. Hay 2001 tested the transparency/opacity hypothesis for English by parsing for base lexemes within the dictionary definitions of their derivational counterparts (e.g. *sane* in the definition of *insane*). In Portuguese, inconsistent methodologies for defining headwords in one and the same dictionary (Correia 1999) mean that applying Hay’s method would at best lead to distortions in the results. In the *Grande Dicionário da Língua Portuguesa*, for example, definitions for *constituição* do not contain its base *constituir*, while definitions for *orçamento* do contain *orçar*, despite much higher relative frequencies for both derivations’ opaque senses. This study employs an alternative method that measures the relative frequency of polysemous nouns in *-ção* and *-mento* by analysing their contextualized use in a Portuguese corpus. The lexicometric programme *CooCS* (Martinez 2012) identifies and measures lexical co-occurrences for a given word and enables cumulative determinations of contextual use for each, reducing the unrealistic manual verification of individual tokens in a large corpus. When applied to derivations in *-ção* and *-mento*, a more nuanced and reliable picture of the correlation between relative frequency and semantic transparency/opacity in derivational morphology emerges.

Keywords: Corpus linguistics, derivational morphology, polysemy, Portuguese, relative frequency

References

- Bybee, J. (1985). *Morphology. The Relation between Form and Meaning*. Amsterdam: John Benjamins.
- Corbin, D. (1987). *Morphologie dérivationnelle et structuration du lexique*. 2 vols. Tübingen: Max Niemeyer Verlag.
- Correia, M. (1999). *A denominação das qualidades – contributos para a compreensão da estrutura do léxico português*. PhD thesis. University of Lisbon. Portugal.
- Figueiredo, C. (1996). *Grande Dicionário da Língua Portuguesa*. 25th ed., 2 vols. C/CD-Rom. Lisboa: Bertrand.
- Grimshaw, J. (1990). *Argument structure*. Cambridge, MA. MIT Press.
- Hay, J. (2001). ‘Lexical frequency in morphology: is everything relative?’ *Linguistics* 39 (6), 1041-1070.
- Hay, J. and Baayen, R. H. (2002). ‘Parsing and productivity’. *Yearbook of Morphology 2001*. Dordrecht: Kluwer Academic Publishers, 203-235.
- Kerleroux, F. ‘Des noms indistincts’. In B. Fradin (ed.), *La raison morphologique. Hommage à la mémoire de Danielle Corbin*. Amsterdam/Philadelphia: John Benjamins, 113-132.
- Martinez, W. (2012). ‘Au-delà de la cooccurrence lexicale - poly-cooccurrence et trames de cooccurrence.’ *Corpus n°11: La cooccurrence: du fait statistique au fait textuel*. Nice: Faculté des Lettres, 191-218.