

Morphological productivity and family size: evidence from French compound nouns *garde-x* and *N-de-N*

Recent works in derivational morphology have shown a new interest for the concept of morphological productivity and for its various measuring tools, both quantitative (Baayen 1992, 1993; Baayen & Lieber 1991; Baayen & Renouf 1996) and qualitative (Bauer 2001; Dal 2003). However, the productivity of compound words has been very little explored, especially in French, except for the work of Krott, Schreuder & Baayen (1999). Recent studies, have proposed a new approach to morphological productivity in term of family size (De Jong, Schreuder, & Baayen, 2000, del Prado M., Bertram, et al). This paper considers a possible correlation between morphological productivity and family size and based on the results of a quantitative study of two types of compounds in French (synthetic *garde-x* and *N-de-N* compounds) during four different chronological periods.

Two primary sources (the Gallica corpus of French National library (50 million words) and Littré's *Dictionnaire de la langue française* (1863-1877) with 80 000 entries) were searched manually to create a primary database of French compound nouns; the larger corpus Frantext, with more than 120 millions of words, was later used to validate the frequency of the forms *garde-x* and *N-de-N* across four chronological periods (1606-1694; 1695-1798; 1799-1872; 1873-1920). This case study was applied to test both quantitative measure of morphological productivity (Baayen 1992, 1993) and family size approach (De Jong, Schreuder, & Baayen, 2000, del Prado M. et al 2004).

In the study, the quantitative measure based on the notion of the hapax legomenon (Baayen 1992) was used to measure the productivity of compound nouns. Productivity is determined according to the formula [$P = n_1 / N$] where **P** represents productivity in a narrow sense (the rate of the growth of the vocabulary), **n₁** represents the number of hapax legomenon (the forms that appear once only in a big corpus), and **N** represents the total number of observed occurrences in the corpus containing the derivational analyzed type. According to this measure of productivity, a category with a large number of words of high frequency will have a big value of **N** and, accordingly, a low degree of productivity.

Our study compared the correlations between the productivity of French compound nouns and the size of their morphological families, as well as investigates whether the family size can predict the productivity of French compound nouns. The results demonstrated that for the *garde-x* compounds, the biggest family size (F=39) is not correlated to the highest productivity (P=0.041). The experimental evidence from *N-de-N* compounds showed very similar pattern: the highest productivity (P=0.044) was presented by smallest size family (F=33). The results of this study demonstrated the necessity of complementary research of the different type of compound and their morphological families.

Selected References:

- Baayen, R.H.:1992, Quantitative aspects of morphological productivity, *Yearbook of Morphology*, Kluwer Academic Publishers, pp.109-149.
- De Jong, N. H., Schreuder, R. and Baayen, R. H.: 2000, The morphological family size effect and morphology, *Language and Cognitive Processes* 15, 329–365.
- Krott, A., Schreuder, R., Baayen, R. H.:1999, Complex words in complex words, *Linguistics*, 37-5, pp. 905-926.
- del Prado, M., Bertram, R., Haikio, T., Schreuder, R., Baayen, H.:2004, Morphological Family Size in a Morphologically Rich Language: The Case of Finnish Compared with Dutch and Hebrew, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 30, No. 6, p.1271–1278.