

## La productivité morphologique des composés A-N et N-de-N dans la littérature française du 17<sup>e</sup> au 20<sup>e</sup> siècle

Elena Voskovskaia  
Grande Prairie Regional College

La productivité des mots composés est un domaine de recherche très peu exploré, surtout en français (Krott, Schreider & Baayen 1999 ; Fernández Domínguez 2007 ; 2009). Très peu de travaux ont été réalisés sur la productivité des composés dans une perspective diachronique. La mesure quantitative la plus couramment utilisée est basée sur la notion de l'hapax (Baayen & Lieber 1991 ; Baayen 1992). Toutefois, Hay (2003) montre que c'est la fréquence relative plutôt que la fréquence absolue qui a un effet sur la décomposition et la productivité des mots complexes. Ce papier étudie une corrélation possible entre la productivité (P) et la fréquence relative (FR) des composés A-N (*franc-tireur, faux-saunage*) et N-de-N (*belle-de-jour, pou-de-soie*) dans la littérature française du 17<sup>e</sup> au 20<sup>e</sup> siècle.

La première mesure utilisée pour évaluer le niveau de productivité est la mesure P au sens strict (Baayen & Lieber 1991 ; Baayen 1992) calculée comme  $P = n_1/N$  (où P=le taux de productivité ;  $n_1$  = le nombre d'hapax legomenon ; N= le nombre total d'occurrences). Selon cette mesure, une catégorie avec un grand nombre de mots de haute fréquence a une grande valeur de N et un degré de productivité P moins élevé.

La deuxième mesure appliquée est celle de la fréquence relative FR (Hay 2003) élaborée pour la composition comme  $f_{relative} = f_{composé} / f_{base}$ . Selon cette mesure, les formes de fréquence relative basse sont plus productives. La question qui se pose est la suivante: dans le mot composé, quel élément peut être considéré comme une base? Fernández Domínguez (2007, 2009) suggère que la fréquence de la base des composés peut être mesurée selon trois variantes: a) en ajoutant les fréquences des constituants ; b) en calculant la fréquence moyenne des constituants ; c) en utilisant la fréquence de la tête du composé. Néanmoins, à notre avis, il est nécessaire d'inclure un élément qui n'est pas la tête (selon l'analogie avec la base en dérivation).

La recherche est basée sur le corpus textuel *Frantext* (plus de 170 millions de mots) divisé en quatre périodes selon leur importance dans l'histoire de la langue française : 1606-1694 (17,3 millions de mots) ; 1695-1798 (34,4 millions de mots) ; 1799-1872 (41 millions de mots) ; 1873-1920 (28 millions de mots). La liste de 137 composés A-N et 27 composés N-de-N a été créée à partir du *Dictionnaire de Littré* (1877).

Notre étude a démontré que parmi les 137 formes A-N, 133 composés (97 %) sont moins fréquents que leurs bases indépendamment de la mesure FR utilisée, ce qui permet de les envisager comme décomposables et potentiellement productifs. Néanmoins, l'hypothèse de la corrélation inverse entre les mesures P et FR n'a pas été confirmée pour les composés A-N (36 %). Quant aux composés N-de-N, elle ne peut pas être prouvée ou rejetée à cause de données lacunaires de la mesure P pour ce type de composés. Donc, à l'opposé de ce qu'on observe en dérivation, la corrélation inverse entre FR et P n'est pas explicitement présente en composition. Par suite, la mesure de la fréquence relative ne peut pas être considérée comme un facteur complètement fiable dans l'évaluation du niveau de productivité des noms composés en français ; le critère FR<sub>3</sub> sert plutôt à établir la présence de la productivité (i.e. la décomposition du composé), mais il ne démontre pas son degré.

L'étude de l'évolution diachronique de la productivité des composés A-N et N-de-N révèle la dominance de la période 1606-1694 vis-à-vis du taux élevé de productivité ( $P_1=0.0052$  vs.

P1=0.0296). À partir des années 1695-1798, le niveau de productivité des composés A-N (P2=0.0011 ; P3=0.0008 ; P4=0.0005) et N-de-N (P2=0.0030 ; P3=0.0047 ; P4=0.0046) diminue progressivement.

### Références

- Baayen, H., Lieber, R. 1991. Productivity and English derivation: a corpus-based study, *Linguistics*, 29, pp.801-43.
- Baayen, H. 1992. Quantitative aspects of morphological productivity, *Yearbook of Morphology*, Kluwer Academic Publishers, , pp.109-149.
- Baayen, H. 1993. On frequency, transparency and productivity, *Yearbook of Morphology*, Kluwer Academic Publishers, pp.181-208.
- Bauer, L. 2001. *Morphological productivity*, Cambridge: Cambridge University Press.
- Dal, G. 2003. Productivité morphologique : définitions et notions connexes, dans *Langue française*, v.140, Cedex, Larousse, Paris, pp.3-23.
- Fernández-Domínguez, J., Díaz-Negrillo, A., Štekauer, P., 2007, *How is Low Morphological Productivity Measured?*, *Atlantis*, revista de la Asociación Española de Estudios Anglo-Norteamericanos, 29.1, pp.29-54.
- Fernández-Domínguez, J., 2009, *Productivity in English Word-formation: An approach to N+N compounding*, *European University Studies*, Peter Lang Publishing.
- Hay, J., Baayen, H. 2002. *Affix Productivity and Base Productivity*. Paper presented at the *Morphological Productivity Seminar*, ESSE 6, Aug 30-Sept 3, Strasbourg.
- Hay, J., (2003). *Causes and consequences of word structure*. New York : Routledge.
- Krott, A., Schreuder, R., Baayen, R. H. 1999. Complex words in complex words, *Linguistics*, 37-5, pp. 905-926.