
Composés N-N et N-A dans la littérature française du 17^e au 20^e siècle : la productivité morphologique

Elena Voskovskaia

Grande Prairie Regional College, Canada

1 Introduction

La productivité des mots composés est un domaine de recherche très peu exploré, surtout en français (Krott, Schreider & Baayen 1999), (Fernandez Dominguez 2007, 2009). 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 vise à étudier la productivité morphologique des noms composés N-N et N-A dans la littérature française du 17^e au 20^e siècle. Nous examinons une corrélation possible entre la productivité morphologique et la fréquence relative de ces deux types de composés.

Ce travail vise à vérifier une hypothèse suivante : Il existe une corrélation inverse entre la productivité et la fréquence relative des composés N-N et N-A : les composés dont la fréquence relative est plus basse sont plus productifs que ceux dont la fréquence relative est plus élevée. Nous cherchons à vérifier si la notion de la fréquence relative proposée par Hay (2003) en dérivation peut être appliquée aux mots composés.

Mots-clés : Morphologie, productivité morphologique, corpus textuel, diachronie.

2 Corpus et méthodes

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) et 1873-1920 (28 millions de mots). La liste de 72 composés N-N (*croix-pile*, *volte-face*) et 39 composés N-A (*arc-boutant*, *acquit-patent*, *bec-cornu*) a été créée à partir du Dictionnaire de Littré (1877).

Deux mesures sont utilisées pour évaluer la productivité des composés. 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 dont la fréquence relative est basse sont plus productives.

3 Résultats

Notre étude a démontré que parmi 72 composés N-N, 65 composés (90%) sont moins fréquents que leurs bases, ce qui permet de les envisager comme décomposables et potentiellement productifs. La corrélation entre les mesures P et FR dans les composés N-N est présentée dans le Tableau 1.

Tableau 1 Corrélation entre la productivité et la fréquence relative dans les composés N-N

mesure	composés N-N			
	1606-1694	1695-1798	1799-1872	1873-1920
FR ₁	0.0349	0.0455	0.0427	0.0296
FR ₂	0.2501	0.2268	0.1506	0.1855
FR ₃	0.0099	0.0188	0.0129	0.0118
FR ₄	0.0199	0.0376	0.0257	0.0218
P	0.0770	0.0100	0.0014	0.0012

La corrélation inverse entre les mesures FR₃ et FR₄ et la productivité P a été observée pour 73 % des composés N-N. Le niveau de productivité le plus élevé (P1 = 0.0770) correspond à la fréquence relative la plus basse si on considère la mesure FR₃ = 0.0099 pendant la période 1606-1694 et 1799-1872 (P3 = 0.0014 vs FR₃ = 0.0129). Quant à la mesure FR₂, elle ne présente aucune corrélation inverse. Toutefois, les chiffres du Tableau 1 illustrent que la corrélation n'est pas parfaite : le patron inverse apparaît seulement dans les années 1799-1872 (FR₃ = 0.0427 vs P3 = 0.0014).

L'analyse du type N-A a été basée sur 39 composés dont 35 composés (90 %) sont moins fréquents que leurs bases indépendamment de la variante FR utilisée. Ainsi, ils peuvent être envisagés comme décomposables et potentiellement productifs. La corrélation entre les mesures P et FR dans les composés N-A est résumée dans le Tableau 2.

Tableau 2 Corrélation entre la productivité et la fréquence relative dans les composés N-A

mesure	composés N-A			
	1606-1694	1695-1798	1799-1872	1873-1920
FR ₁	0.0435	0.1469	0.1894	0.2238
FR ₂	0.2138	0.2939	0.3111	0.3366
FR ₃	0.0119	0.0141	0.0345	0.0241
FR ₄	0.0238	0.0283	0.0689	0.0481
P	0.0348	0.0051	0.0012	-

Les données du Tableau 2 révèlent que le patron de la corrélation inverse fonctionne parfaitement pour les composés N-A si on considère la mesure FR₃. Le taux de productivité le plus élevé (P1 = 0.0348) correspond à la fréquence relative la plus basse si on considère la mesure FR₃ = 0.0119.

3 Conclusion

Nous avons avancé l'hypothèse qu'il existe d'une corrélation inverse entre la productivité et la fréquence relative dans les composés français N-N et N-A. L'étude a révélé une faible présence d'une corrélation inverse entre les mesures FR et P : 34% pour les composés N-N et 50% pour les composés N-A.

La mesure de la fréquence relative peut varier considérablement en fonction du choix de la base du composé. Globalement, les mesures FR₃ et FR₄ correspondent mieux à la corrélation attendue, où la fréquence relative la plus basse correspond au niveau de productivité la plus élevée. Étant donné que la mesure FR₄ est un dérivé de la mesure FR₃, la mesure FR₃ est envisagée comme la meilleure variante pour évaluer la fréquence relative en composition.

Ainsi, à l'opposé de ce qu'on observe dans la dérivation, les données obtenues indiquent qu'en composition, la corrélation inverse entre la fréquence relative et la productivité n'est pas explicitement présentée. Le critère FR₃ sert plutôt à établir la présence de la productivité (i.e. la décomposition du composé), mais il ne démontre pas son degré.

Les résultats de cette étude démontrent la nécessité de recherches supplémentaires sur les différents types de composés concernant une corrélation possible entre la productivité morphologique et la fréquence relative.

Références

- Baayen, H., Lieber, R. 1991. Productivity and English derivation: a corpus-based study. *Linguistics* 29. 801-843.
- Baayen, Harald 1992. Quantitative aspects of morphological productivity. In Geert E. Booij & Jaap van Marle (eds.), *Yearbook of morphology 1992*, 109-149. Dordrecht: Kluwer.
- Baayen, Harald 1993. On frequency, transparency and productivity. In Geert E. Booij & Jaap van Marle (eds.), *Yearbook of morphology 1993*, 181-208. Dordrecht: Kluwer.
- Bauer, L. 2001. *Morphological productivity*. Cambridge: Cambridge University Press.
- Dal, Georgette 2003. Productivité morphologique : définitions et notions connexes. Dans *Langue française*, v.140, 3-23. Larousse : Paris.
- Fernandez-Dominguez, J., Diaz-Negrillo, A. & Štekauer, P. 2007. How is Low Morphological Productivity Measured? *Atlantis* 29.1. 29-54. Revista de la Asociación Española de Estudios Anglo-Norteamericanos.
- Fernández-Domínguez, J. 2009. *Productivity in English Word-formation: An approach to N+N compounding*. European University Studies: Peter Lang Publishing.
- Hay, Jennifer & Baayen Harald. 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. 905-926.