

LA PRODUCTIVITÉ MORPHOLOGIQUE DES COMPOSÉS A-N ET N-de-N DANS LA LITTÉRATURE FRANÇAISE DU 17^e AU 20^e SIÈCLE

Elena Voskovskaia
Grande Prairie Regional College

La productivité morphologique des mots composés demeure un domaine très peu exploré (Bauer 2008 ; Krott, Schreider & Baayen 1999 ; Fernandez-Dominguez 2007, 2009), surtout en français. Les études antérieures sur la productivité en dérivation utilisent principalement la mesure quantitative basée sur les hapax, soulevant la question de savoir si l'application d'une seule méthode est suffisamment fiable pour évaluer la productivité en composition. Dans cet article, la productivité des composés est examinée au moyen de deux mesures différentes : la mesure P basée sur les hapax (Baayen et Lieber 1991 ; Baayen 1992) et la fréquence relative (Hay 2003).

1. Introduction

Le phénomène de la productivité est très bien étudié en dérivation, surtout en ce qui concerne l'anglais (Baayen 1992, 2001 ; Bauer 1988, 2001, 2003 ; Hay 2003 ; Plag 1999 ; Baayen et Hay 2002). Quelques recherches ont également été effectuées sur la productivité des affixes en français (Corbin 1975, 1987 ; Dal 2003 ; Namer 2003 ; Grabar et al. 2006). Très peu de travaux ont été réalisés sur la productivité des composés dans une perspective diachronique (Estopà 2009).

La mesure quantitative de la productivité morphologique P la plus couramment utilisée est basée sur la notion de l'hapax (Baayen & Lieber 1991 ; Baayen 1992 ; Baayen et Renouf 1996). Toutefois, en analysant la corrélation entre la fréquence relative et la productivité des formes dérivées en anglais, 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. Notre recherche vise à vérifier une corrélation potentielle 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^e au 20^e siècle.

Notre hypothèse est la suivante : Il existe une corrélation inverse entre la productivité morphologique et la fréquence relative des composés A-N et N-de-N : 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 français.

2. Corpus et méthodes

2.1 Corpus

La recherche est basée sur le corpus textuel *Frantext* comportant plus de 250 millions de mots. Cette base de données est un large corpus informatisé composé de textes en français, créé initialement par le laboratoire ATILF (Analyse et Traitement Informatique de la Langue Française) en 1998. *Frantext* contient plus de 5000 textes appartenant à des

domaines différents tels que les sciences, les arts, la littérature, les techniques, la politique et la philosophie. Ce corpus est très représentatif puisqu'il reflète des niveaux et des registres différents de la langue française et il a une taille suffisante pour appliquer une série de mesures quantitatives.

La productivité des formes recensées a été analysée selon quatre sous-corpus chronologiques conformément à 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). Deux types de noms composés français ont été analysés. La liste de 137 composés A-N et 27 composés N-de-N a été créée à partir du *Dictionnaire de Littré* (1877).

2.2 Méthodes

Deux mesures quantitatives ont été utilisées pour évaluer la productivité des noms composés dans *Frantext*. La première mesure, celle basée sur les hapax (Baayen 1992 ; Baayen et Lieber 1991), est calculée comme suit :

$$(1) \quad P = n_1 / N$$

où P=la productivité; n_1 = le nombre d'hapax (les formes qui apparaissent une fois seulement dans un grand corpus) ; N= le nombre total d'occurrences dans le corpus).

Cette mesure ne prend en considération que le nombre d'hapax (n_1) et le nombre total d'occurrences de la structure analysée (N). Ce calcul évalue la productivité du gabarit (du type) des noms composés en démontrant le taux de croissance du vocabulaire pour chaque période étudiée. Le nombre d'hapax correspond à la probabilité de rencontrer de nouveaux types de lexèmes qui n'ont pas été enregistrés dans les échantillons précédemment analysés (Baayen 1993).

La deuxième mesure évalue la fréquence relative des composés. Cette formule a été élaborée par Hay (2003) pour les mots dérivés comme $f_{relative} = f_{dérivé} / f_{base}$ et adaptée aux composés comme suit :

$$(2) \quad f_{relative} = f_{composé} / f_{base}$$

Selon l'analogie avec la dérivation, un niveau moins élevé de FR doit indiquer que le composé représente une structure décomposable et potentiellement productive. 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 possibles : a) en ajoutant les fréquences des constituants ; b) en calculant la fréquence moyenne des constituants ; c) en utilisant uniquement la fréquence de la tête du composé. Néanmoins, à notre avis, il est nécessaire d'inclure dans le calcul un élément qui n'est pas la tête (selon l'analogie avec la base en dérivation) :

$$(3) \quad \begin{array}{ll} \text{a.} & FR_1 = f_{relative} = f_{composé} / f_{base} \quad (B_1 = \text{tête}) \\ \text{b.} & FR_2 = f_{relative} = f_{composé} / f_{base} \quad (B_2 = \text{non tête}) \\ \text{c.} & FR_3 = f_{relative} = f_{composé} / f_{base} \quad (B_3 = B_1 + B_2) \\ \text{d.} & FR_4 = f_{relative} = f_{composé} / f_{base} \quad (B_4 = ((B_1 + B_2) : 2)) \end{array}$$

2.3 Identification de la base dans les mots composés

La base morphologique est clairement identifiée dans les mots dérivés formés par affixation. Ainsi, la base du mot dérivé est un morphème libre auquel s'attache un affixe qui est un morphème lié. Le suffixe remplit généralement la fonction de tête en portant les caractéristiques syntaxiques et sémantiques du mot.

Par contre, en composition, la base n'est pas clairement identifiée (Bauer 1983, 2001; Dressler 2006; Spenser 1991; Krott et al. 1999). Théoriquement, la base dans les composés peut être identifiée de deux façons différentes. Par exemple, dans le composé *belle-sœur*, la base morphologique est : a) soit le morphème libre qui n'est pas la tête du composé (*belle*) ; b) soit tout morphème libre (*belle* et *sœur*). Donc, en composition, la base peut être définie de plus d'une façon :

- (4) a. la base du composé est un morphème libre (donc, tous les constituants du composé peuvent être des bases) ;
b. la base du composé est le constituant qui n'est pas la tête du composé

De ce fait, en analysant la productivité des noms composés français, nous utilisons quatre variantes de mesures de la base pour évaluer la fréquence relative des formes recensées. Nous avons recours aux trois variantes proposées par Fernández-Domínguez et al. (2007, 2009) mentionnées ci-dessus en ajoutant une quatrième variante où la base du composé est le constituant qui n'est pas la tête.

De cette façon, la comparaison des résultats de chaque variante à ceux obtenus par la mesure basée sur les hapax nous permet d'identifier la base qui procure les meilleurs résultats pour la mesure de la fréquence relative de Hay (2003).

3. Résultats

3.1 Productivité morphologique des composés A-N

Parmi 137 composés A-N, 102 sont réguliers et 35 sont irréguliers (*bonne-vilaine*, *double-bec*, *gros-jean*). Parmi les 102 composés réguliers, 87 ont la tête à droite (*bel-outil*, *blanc-manger*, *demi-frère*) et 15 sont des exocentriques (*claire-voie*, *haute-taille*, *longue-langue*). Le taux de productivité des composés A-N réguliers et irréguliers est résumée dans le Tableau 1.

Tableau 1 : Productivité morphologique des composés A-N

Position de la tête	1606-1694			1695-1798			1799-1872			1873-1920		
	N	n ₁	P	N	n ₁	P	N	n ₁	P	N	n ₁	P
tête à droite	2986	8	0.0027	6241	4	0.0006	8659	8	0.0009	6116	1	0.0002
exocentriques réguliers	7	7	1.0000	52	3	0.0577	360	4	0.0111	268	1	0.0037
Total réguliers :	2993	15	0.0050	6293	6	0.0010	9008	9	0.0010	6402	2	0.0003
exocentriques irréguliers	860	5	0.0058	1110	2	0.0018	1932	0	-	900	2	0.0022

Total exocentriques	867	12	0.0138	1162	5	0.0043	2292	4	0.0017	1168	3	0.0026
----------------------------	-----	----	--------	------	---	--------	------	---	--------	------	---	--------

Du tableau ci-dessus se dégagent les conclusions suivantes :

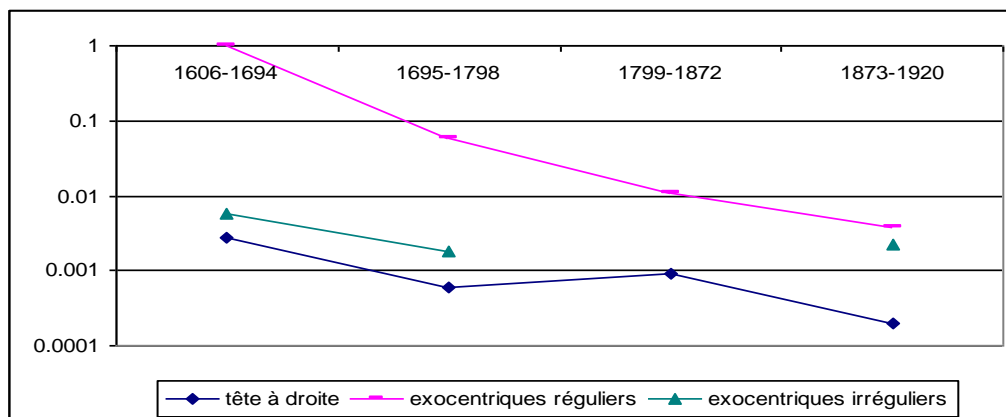
La productivité des composés A-N réguliers est très similaire à celle des composés A-N irréguliers pendant les années 1606-1694 ($P_1 = 0.0050$ vs $P_1 = 0.0058$). À partir de la période 1695-1798, le taux de productivité des composés irréguliers devient plus élevé que celui des composés réguliers.

La productivité des composés A-N dépend de la position de la tête morphologique. Les composés exocentriques réguliers sont considérablement plus productifs ($P_1 = 1.0000$; $P_2 = 0.0577$; $P_3 = 0.0111$; $P_4 = 0.0037$) durant toutes les périodes analysées. Cependant, le résultat $P_1 = 1.0000$ doit être pris en considération prudemment. En premier lieu, le nombre d'occurrences est moins de 50 ($N = 7$). En deuxième lieu, comme le montre Baayen (1992 :143), le taux très élevé de productivité présenté par un nombre assez bas d'occurrences doit être envisagé avec beaucoup de précautions. Le problème consiste dans le fait que pour une petite valeur de N les mathématiques sous-jacentes ne s'appliquent pas au calcul de la mesure P. Ce phénomène a été aussi noté par Bauer (2001 : 150).¹

Globalement, les A-N exocentriques réguliers ($P = 0.0218$) sont 11.5 fois plus productifs que les exocentriques irréguliers ($P = 0.0019$). Toutefois, la productivité totale des composés A-N irréguliers est légèrement plus élevée que celle des A-N réguliers ($P = 0.0019$ vs $P = 0.0013$).

La productivité des composés A-N est présentée dans le Graphique 1, basé sur un axe logarithmique pour mieux différencier les composés réguliers et irréguliers.

Graphique 1 : Productivité des composés A-N selon la position de la tête



¹ Bauer (2001 :151) attire l'attention sur le cas du suffixe *-iana* qui ressort une fois seulement ($N = 1$) dans le corpus Wellington (*Victoriana*) et dont $n_1 = 1$. Théoriquement, ce suffixe doit être considéré comme 100 % productif. Néanmoins, Bauer indique que tous les cas où $N = 1$ doivent être examinés avec prudence : plus le nombre d'occurrences des formes recensées est petit, moins les résultats peuvent être considérés comme concluants.

Comme le montre le graphique ci-dessus, les composés A-N exocentriques réguliers sont considérablement plus productifs que les A-N irréguliers et surtout les A-N avec la tête à droite au cours des années 1606-1798. Par contre, lors de la période 1873-1920, la productivité des A-N endocentriques diminue énormément lorsque le taux de productivité des A-N exocentriques réguliers et irréguliers devient assez similaire.

3.2 Productivité morphologique des composés N-de-N

Parmi 27 composés N-de-N cinq formes sont régulières (*belle-de-jour*, *mestre de camp*, *pou-de-soie*) et 22 formes sont irrégulières (*bec-de-cane*, *cou-de-pied*, *cul-de-sac*). Les données statistiques sur la productivité des composés N-de-N sont récapitulées dans le Tableau 2.

Tableau 2 : Productivité morphologique des composés N-de-N

Position de la tête	1606-1694			1695-1798			1799-1872			1873-1920		
	N	n ₁	P	N	n ₁	P	N	n ₁	P	N	n ₁	P
tête à gauche	52	1	0.0192	101	0	-	9	0	-	2	0	-
tête à droite	-	-	-	-	-	-	1	1	1.0000	-	-	-
exocentriques réguliers	4	0	-	-	-	-	-	-	-	1	0	-
Total réguliers	56	1	0.0180	101	0	-	10	1	0.1000	3	0	-
exocentriques irréguliers	113	4	0.0354	234	1	0.0043	626	2	0.0032	654	3	0.0046
Total exocentriques :	117	4	0.0342	234	1	0.0043	626	2	0.0032	654	3	0.0046

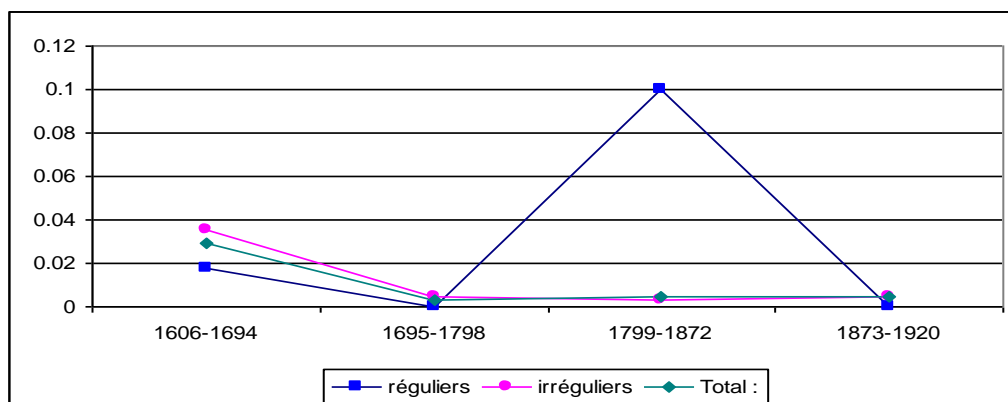
Des chiffres présentés ci-dessus permettent d'avancer des conclusions suivantes. Il n'y a pas suffisamment de données pour tirer des conclusions précises concernant la productivité des composés N-de-N réguliers. Malgré le niveau de productivité très élevé de ce type de composés pendant la période 1799-1872 ($P_3 = 1.0000$), ces données ne peuvent pas être considérées comme fiables à cause du nombre d'occurrences trop petit ($N=1$).

En ce qui concerne l'aspect régulier/irrégulier des exocentriques, les composés N-de-N réguliers sont globalement deux fois plus productifs que les N-de-N irréguliers ($P = 0.0061$). Néanmoins, il faut prendre en considération le fait que lors des années 1799-1872, il n'y a qu'une seule occurrence dans les composés N-de-N avec la tête à droite.

Quant à la corrélation entre le niveau de productivité et la position de la tête morphologique des composés N-de-N réguliers, aucune conclusion n'est pas possible en raison de données lacunaires pour les composés réguliers.

Le Graphique 2 montre le rapport entre le niveau de productivité et le facteur régulier/irrégulier des composés N-de-N.

Graphique 2 : Productivité des composés N-de-N selon l'aspect régulier/ irrégulier



Conformément à la représentation ci-dessus, le taux de productivité le plus élevé est observé dans les N-de-N réguliers entre 1799 et 1872. Cependant, pendant la première période, la productivité des N-de-N réguliers est presque deux fois plus bas que celui des N-de-N irréguliers. En même temps, rappelons que dans le cas des N-de-N réguliers, on observe les lacunes de données statistiques entre les années 1695-1798 et 1873-1920. Compte tenu de ces faits, il est impossible de déduire des conclusions fiables à partir des résultats obtenus pour ce type de composés.

3.3 Fréquence relative des composés A-N

Parmi les 137 composés 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. La corrélation entre les mesures P et FR dans les composés A-N réguliers et irréguliers est présentée dans le Tableau 3.

Tableau 3 : Corrélation entre la productivité et la fréquence relative dans les composés A-N

mesure	composés réguliers				composés irréguliers			
	1606-1694	1695-1798	1799-1872	1873-1920	1606-1694	1695-1798	1799-1872	1873-1920
FR ₁	0.0338	0.0831	0.0752	0.0802	0.0112	0.0114	0.0063	0.0042
FR ₂	0.0311	0.0241	0.0191	0.0231	0.0850	0.1554	0.0638	0.0613
FR ₃	0.0063	0.0065	0.0069	0.0078	0.0080	0.0051	0.0044	0.0036
FR ₄	0.0123	0.0127	0.0137	0.0157	0.0162	0.0101	0.0087	0.0072
P	0.0050	0.0009	0.0010	0.0003	0.0058	0.0018	-	0.0022

Dans les composés A-N réguliers, le niveau de productivité le plus élevé (P1 = 0.0050) correspond à la fréquence relative la plus basse FR3 = 0.0063. (FR₃ = $f_{\text{composé}} / f_{\text{base}}$ (B₃ = B₁ + B₂)).

Telle qu'indiqué au Tableau 3, seulement deux sur quatre périodes correspondent au schéma de la corrélation inverse de Hay (2003 :151), tant pour FR₃ que FR₄: 1606-1694 et 1873-1920.

Dans les composés A-N irréguliers, cette tendance n'est pas présente. Le taux de productivité le plus grand (P1 = 0.0058) est celui de la première période et la fréquence relative la plus petite (FR₃ = 0.0036) est celle de la quatrième : il n'y a donc pas de corrélation inverse générale.

3.4 Fréquence relative des composés N-de-N

En ce qui concerne 27 composés N-de-N, la majorité absolue (99.6 %) de composés sont moins fréquents que leurs bases, ce qui permet de les considérer comme potentiellement décomposables et productifs.

La corrélation entre les mesures P et FR dans les composés N-de-N réguliers et irréguliers est résumée dans le Tableau 4.

Tableau 4 : Corrélation entre la productivité et la fréquence relative dans les composés N-de-N

mesure	composés réguliers				composés irréguliers			
	1606-1694	1695-1798	1799-1872	1873-1920	1606-1694	1695-1798	1799-1872	1873-1920
FR ₁	0.4031	0.4331	0.0975	0.3493	0.0082	0.0175	0.0092	0.0096
FR ₂	0.0168	0.0252	0.0061	0.0018	0.0457	0.1206	0.0435	0.0820
FR ₃	0.0159	0.0230	0.0015	0.0017	0.0047	0.0077	0.0044	0.0052
FR ₄	0.0319	0.0461	0.0030	0.0033	0.0094	0.0153	0.0088	0.0103
P	0.018	-	0.1000	-	0.0354	0.0043	0.0032	0.0046

Pour les composés N-de-N réguliers, le taux de productivité le plus élevé (P3 = 0.1000) correspond à la fréquence relative la plus basse FR₃ = 0.0015.

Quant à la mesure FR₁, le patron inverse entre la productivité P et la fréquence relative FR coïncide aussi avec la troisième période 1799-1872 (FR₁ = 0.0975 vs P3 = 0.1000) alors que dans le cas de la mesure FR₂, la corrélation inverse n'est pas présente.

À cause de données lacunaires liées à la mesure P pour la deuxième (1695-1798) et la quatrième (1873-1920) période, il est difficile de dégager une corrélation fiable entre les mesures FR et P pour ce type de composés.

Pour ce qui est des composés N-de-N irréguliers, le taux de productivité le plus grand (P1 = 0.0354) de la première période correspond à la fréquence relative la plus petite (FR₃ = 0.0044) de la troisième : il n'y a donc pas de corrélation inverse générale.

3.5 Productivité des composés A-N et N-de-N en diachronie

Pour évaluer l'évolution des composés A-N et N-de-N en diachronie, nous avons recours à la mesure de productivité P au sens strict.

3.5.1 Productivité des composés A-N en diachronie

Le changement de la productivité des composés A-N (réguliers et irréguliers) en diachronique est récapitulé dans le Tableau 5.

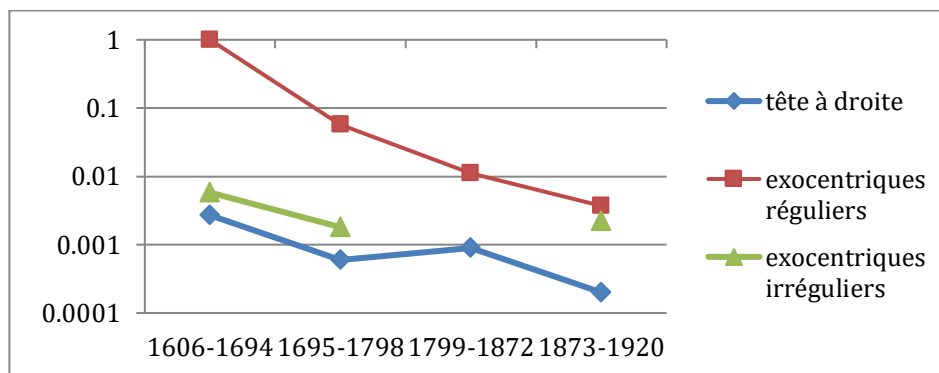
Tableau 5 : Productivité des composés A-N par périodes

composés	1606-1694	1695-1798	1799-1872	1873-1920
réguliers	0.0050	0.0010	0.0010	0.0003
irréguliers	0.0058	0.0018	-	0.0022

La productivité la plus élevée des composés A-N réguliers a été enregistrée pendant la première période (voir les zones ombrées). Elle s'abaisse considérablement durant les années 1695-1872 ($P_2 < 5$ fois et $P_3 = < 5$ fois), et surtout au cours des années 1873-1920 ($P_4 = < 17$ fois). En ce qui concerne les A-N irréguliers, le taux de productivité le plus élevé est aussi observé pendant la première période 1606-1694 (voir les zones ombrées). Au cours des années 1695-1798, leur productivité s'abaisse par rapport à la première période ($P_2 = < 3$ fois) en augmentant un peu pendant la quatrième période ($P_4 = > 1.2$ fois). Aucun hapax n'apparaît dans ce type de composés entre 1799-1872.

L'évolution diachronique des composés A-N selon la position de la tête morphologique est résumée dans le Graphique 3.

Graphique 3 : Évolution globale des composés A-N (réguliers et irréguliers) selon la position de la tête morphologique



Comme le montre le Graphique 3, le niveau de productivité le plus élevé a été observé dans les formes exocentriques régulières dans les années 1606-1920. En fait, il apparaît qu'entre 1606 et 1694, le taux de productivité P des composés clair-N (*claire-voie*), courte-N (*courte-boule*, *courte-queue*), haut-N (*haute-taille*), saint-N (*sainte-nitouche*) était égal à 1.0000. Le taux de productivité des A-N dont la tête est à droite est notablement plus bas par rapport aux A-N exocentriques réguliers ($P_1 = < 37$ fois ; $P_2 = < 96$ fois ; $P_3 = < 12$ fois ; $P_4 = < 19$ fois) pendant chaque période étudiée. Pour les A-N exocentriques irréguliers, leur niveau de productivité est nettement plus élevé que celui des A-N endocentriques (P_1

= >2.1 fois ; P2 = >3 fois ; P4 = >11) durant les années 1606-1694, 1695-1798 et 1873-1920. Néanmoins, il est plus bas que celui des exocentriques réguliers tout au long de quatre périodes examinées.

3.5.2 Productivité des composés N-de-N en diachronie

Le groupe de composés N-de-N est assez petit (27), de plus, plusieurs données statistiques manquent pour ce type de composés. Le Tableau 6 résume les données de cette étude (où N = le nombre total d'occurrences ; n₁ = nombre d'hapax ; P = productivité morphologique au sens strict).

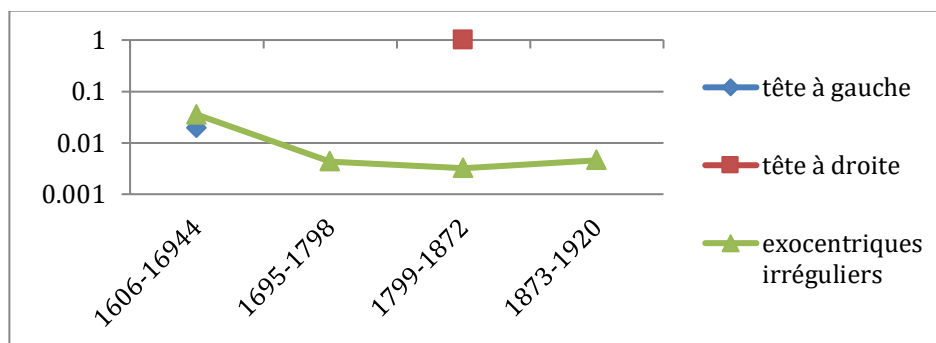
Tableau 6 : Productivité des composés N-de-N réguliers par périodes

période	N	n ₁	P
1606-1694	56	1	0.0180
1695-1798	101	-	-
1799-1872	10	1	0.1000
1873-1920	3	-	-

Contrairement à ce qu'on a observé précédemment dans le type A-N, le taux de productivité des N-de-N réguliers est le plus élevé pendant la troisième période (voir les zones ombrées). Toutefois, ce résultat est à interpréter avec précaution à cause du petit nombre d'occurrences (N = 10). Entre 1606 et 1694, le niveau de productivité des composés N-de-N est considérablement plus bas en comparaison avec la période 1799-1872 la plus fructueuse (P1 = <6 fois). Aucun hapax n'apparaît dans ce type de composés pendant la deuxième (1695-1798) et la quatrième (1873-1920) période.

Pour examiner l'évolution des composés N-de-N réguliers et irréguliers, nous recourons à l'analyse comparative de la valeur P selon la position de la tête morphologique. Les résultats de cette étude sont récapitulés dans le Graphique 4.

Graphique 4 : Évolution globale des composés N-de-N (réguliers et irréguliers) selon la position de la tête morphologique



L'analyse du Graphique 4 démontre que le type N-de-N avec la tête à droite est la seule structure qui possède un taux de productivité assez élevé. Néanmoins, ce résultat ne peut pas être considéré comme fiable à cause du petit nombre d'occurrences (N = 1) et

d'hapax ($n_1 = 1$). En outre, cette structure n'apparaît que durant une seule période dans Frantext (1799-1872). Le seul type qui apparaît dans Frantext au cours des quatre périodes est les N-de-N exocentriques irréguliers dont le taux de productivité le plus élevé est noté entre 1606 et 1694. Leur productivité diminue de façon considérable au cours des années 1799-1872 ($P_3 = <11$ fois), en restant presque identique pendant la deuxième (1695-1798) et la quatrième (1873-1920) période.

3.5.3 Productivité globale des composés A-N et N-de-N

Pour déterminer les types de composés les plus productifs d'une manière globale, l'évolution générale de la productivité des composés A-N et N-de-N (réguliers et irréguliers) est examinée tout au long des années 1606-1920. Les résultats sont résumés dans le Tableau 7 (où P = productivité au sens strict ; $P_{\text{élevée}}$ = changement du niveau de productivité par rapport à la valeur de P la plus élevée ; $P_{\text{pér.}}$ = changement du niveau de productivité par rapport à la valeur de P de la période précédente).

Tableau 7 : Évolution globale de la productivité des composés A-N et N-de-N

période	A-N			N-de-N		
	P	$P_{\text{élevée}}$	$P_{\text{pér.}}$	P	$P_{\text{élevée}}$	$P_{\text{pér.}}$
1606-1694	0.0052			0.0296		
1695-1798	0.0011			0.0030		
		-79 %	-79 %		-90 %	-90 %
1799-1872	0.0008			0.0047		
		-85 %	-27 %		-84 %	+57 %
1873-1920	0.0005			0.0046		
		-90 %	-38 %		-85 %	-2 %
Moyenne	0.0019			0.0105		

La période 1606-1994 est la période la plus fructueuse pour le processus de composition en français (voir les zones ombrées). À partir des années 1695-1798, le taux de productivité subit une baisse allant de 79 % à 90 % pour les types recensés. Vers 1873-1920, le déclin le plus considérable a été constaté dans le type A-N (-90 %) par rapport à la première période qui est la plus fructueuse.

À partir des années 1799-1872, le type N-de-N devient le type le plus productif (+57%). Toutefois, ces résultats sont à interpréter avec précaution puisque d'un côté, ce groupe de composés est assez petit (cinq réguliers et 22 irréguliers) ayant un nombre d'occurrences notablement faible, d'autre côté, les données statistiques sont très lacunaires pour ce type de composés.

4 Conclusion

Nous avons avancé l'hypothèse qu'il existe d'une corrélation inverse entre la productivité et la fréquence relative dans les composés français A-N et N-de-N. L'étude a révélé que la corrélation inverse entre les mesures FR et P (avec les mesures FR_3 et FR_4) a été observée pour 80 % des composés A-N (60% sont A-N réguliers et 20% A-N irréguliers).

Toutefois, pour les composés N-de-N, la corrélation inverse entre les mesures FR et P ne peut pas être évaluée de façon fiable pour cause de données lacunaires de P. Une faible présence d'une corrélation inverse entre les mesures FR et P a été confirmée pour les composés N-de-N réguliers (17%).

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.

L'absence de corrélation significative inverse entre les mesures FR et P peut être expliquée par le fait que dans la majorité des formes recensées A-N et N-de-N (97% et 99 %) les bases sont plus fréquentes que le mot composé. Par contre, la fréquence de la base en dérivation est souvent plus basse que celle du dérivé.

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.

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 ; ce phénomène est observable indépendamment du type de composés ou la position de la tête morphologique.

En ce qui concerne les années 1695-1798, les formes recensées ont subi une chute considérable de la productivité, soit entre 79 % et 90 %.

Lors des années 1799-1872, la productivité du type A-N connaît un autre affaiblissement de 27 % par rapport à la valeur de P de la période précédente. Par contre, le taux de productivité du type N-de-N est devenu plus élevé (+57 %). Toutefois, ces résultats sont à interpréter avec précaution à cause des données statistiques lacunaires pour ce type de composés.

Finalement, au cours de la période 1873-1920, le développement du processus de composition se stabilise en démontrant un petit déclin.

Références

- Amiot, Dany. 2005. Between Compounding and Derivation: Elements of Word Formation Corresponding to Prepositions. Dans *Morphology and its Demarcations: Selected Papers from the 11th Morphology Meeting, Vienna, February 2004*, sous la direction de Wolfgang U. Dressler, Dieter Kastofsky et Franz Rainer, 183-196. Amsterdam : John Benjamins.
- Baayen, Harald, et Rochelle Lieber. 1991. Productivity and English derivation: a corpus-based study. *Linguistics* 29. 801-843.
- Baayen, Harald. 1992. Quantitative aspects of morphological productivity. Dans Geert E. Booij & Jaap van Marle (eds.), *Yearbook of morphology 1992*, 109-149. Dordrecht: Kluwer.
- Baayen, Harald. 1993. On frequency, transparency and productivity. Dans Geert E. Booij & Jaap van Marle (eds.), *Yearbook of morphology 1993*, 181-208. Dordrecht: Kluwer.

- Baayen, Harald R. 2009. Corpus Linguistics in Morphology: Morphological Productivity. Dans *Corpus Linguistics : An International Handbook*, sous la direction de Anke Lüdeling et Merja Kytö, 900-919. Berlin: Mouton De Gruyter.
- Bauer, Laurie. 2001. *Morphological productivity*. Cambridge: Cambridge University Press.
- Bauer, Laurie. 2009. Typology of Compounds. Dans *The Oxford Handbook of Compounding*, sous la direction de Rochelle Lieber et Pavol Štekauer, 343-356. Oxford : Oxford University Press.
- Dal, Georgette. 2003. Productivité morphologique : définitions et notions connexes. Dans *Langue française*, v.140, 3-23. Larousse : Paris.
- Fernandez-Dominguez, Jesus, Ana Diaz-Negrillo, et Pavol Štekauer, P. 2007. How is Low Morphological Productivity Measured? *Atlantis* 29.1. 29-54. Revista de la Asociación Española de Estudios Anglo-Norteamericanos.
- Fernandez-Dominguez, Jesus. 2009. *Productivity in English Word-formation: An approach to N+N compounding*. European University Studies: Peter Lang Publishing.
- Hay, Jennifer, et Harald Baayen 2002. Affix Productivity and Base Productivity. Paper presented at the Morphological Productivity Seminar, ESSE 6, Aug 30-Sept 3, Strasbourg.
- Hay, Jennifer. 2003. *Causes and consequences of word structure*. New York: Routledge.
- Krott, Andrea, Robert Schreuder, et Harald Baayen. 1999. Complex words in complex words. *Linguistics* 37-5. 905-926.