



HAL
open science

LES UNITÉS GRAPHIQUES DE L'ANCIEN ET DU MOYEN FRANÇAIS : UN LIEU D'OBSERVATION PRIVILÉGIÉ POUR UNE MEILLEURE APPROCHE DES PHÉNOMÈNES DE LEXICALISATION ET DE GRAMMATICALISATION DU FRANÇAIS?

Laurent Balon

► **To cite this version:**

Laurent Balon. LES UNITÉS GRAPHIQUES DE L'ANCIEN ET DU MOYEN FRANÇAIS : UN LIEU D'OBSERVATION PRIVILÉGIÉ POUR UNE MEILLEURE APPROCHE DES PHÉNOMÈNES DE LEXICALISATION ET DE GRAMMATICALISATION DU FRANÇAIS? . Dossiers d'HEL, 2016, Écriture(s) et représentations du langage et des langues, 9, pp.304-316. hal-01305404

HAL Id: hal-01305404

<https://hal.science/hal-01305404>

Submitted on 21 Apr 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**LES UNITÉS GRAPHIQUES DE L'ANCIEN ET DU MOYEN FRANÇAIS :
UN LIEU D'OBSERVATION PRIVILÉGIÉ POUR UNE MEILLEURE APPROCHE DES
PHÉNOMÈNES
DE LEXICALISATION ET DE GRAMMATICALISATION DU FRANÇAIS ?**

Laurent Balon

Docteur en Sciences du langage, chercheur indépendant

RÉSUMÉ

Cette contribution interroge le rapport (réflexif ou spontané ?) qu'entretiennent, en ancien et moyen français, les segments graphiques avec certaines unités linguistiques envisagées en tant qu'unités grammaticales. L'objet est de montrer qu'ils apparaissent comme un lieu privilégié de l'observation des phénomènes de lexicalisation et de grammaticalisation du français et que leur étude en diachronie contribuera à une meilleure approche de ces faits de langue.

MOTS-CLEFS

Ancien et moyen français – unités/segments graphiques – unités grammaticales – lexicalisation – grammaticalisation

ABSTRACT

This paper investigates the relation between graphical sequences and grammatical units in Old and Middle French, in order to establish whether this relation is spontaneous or reflexive. The objective is to demonstrate that such relations are useful in gaining a better understanding of lexicalization and grammaticalization processes. These phenomena are thus provided a further angle of inquiry.

KEY WORDS

Old and Middle French – graphical sequences – grammatical units – lexicalization – grammaticalization

Les unités graphiques de l'Ancien et du Moyen Français

0. INTRODUCTION

Les travaux de P. Rickard (1982), N. Andrieux-Reix et S. Monsonégo (1997, 1998, 1999, 2000, 2003a) ont mis en évidence une lacune dans les études médiévales : alors que la question d'une éventuelle *scriptura continua* « originaire » du français n'est pas tranchée, et que dans l'évolution du français écrit ce sont les segmentations les plus radicales qui sont devenues la norme, pendant longtemps la pratique des copistes concernant la « syntaxe graphique » des manuscrits a été renvoyée à l'aléatoire de la « variation graphique » sans plus de recherches systématiques ou perçue comme un geste vain ne méritant pas notre attention (H. Lewicka 1963¹ ; R.-L. Wagner 1974²). Or cette caractéristique de la matérialité graphique du manuscrit, souligne N. Andrieux-Reix (2003a, p. 137),

se révèle étonnamment permanente, [et] n'a jamais été abordée dans les traités et les manuels, pas plus médiévaux que postérieurs au Moyen Âge : en-dehors de ce qui concerne l'apostrophe et le trait d'union, qui restent la part mineure de la question, le blanc et les constituants de l'entre-deux blancs n'ont été définis – normalisés ? – que par défaut : par la fixation d'images véhiculées et consacrées dans les entrées de dictionnaires et dans tous textes imprimés.

Les premiers travaux d'une étude systématique des séquences graphiques ont dressé un premier bilan des mécanismes régissant ces phénomènes, et P. Rickard (1982, pp. 486-488) conclut à leur quasi détermination par la récurrence de deux observations : la brièveté du mot, généralement monosyllabique et « outil », entrant en séquence avec un élément subséquent, et l'utilité grammaticale, par un rapport syntagmatique étroit (par ex. : *leroi*), de cette séquence. Or les constantes relevées, manifestes derrière les variations, invitent à s'interroger sur l'existence d'une norme implicite, non prescriptive, relevant de la conscience linguistique d'une époque et d'un lieu et sur d'éventuelles motivations grammaticales, alors que se manifestent de nombreuses incertitudes concernant le statut grammatical de certains éléments (par ex. : *très* (préfixe intensif ou adverbe ?) et l'unité ou non de formes en cours de lexicalisation (par ex. : *si tost/sitost*) dans une langue en construction et un état de langue lui-même en constante évolution. C'est ce que souligne avec force N. Andrieux-Reix (2003a, p. 137) :

Le choix d'écrire en continu ce qui pouvait être séparé, a été à chaque fois le fait d'une initiative individuelle, sans doute en grande partie spontanée, mais qui s'est trouvée en concordance avec d'autres initiatives individuelles. Ces produits convergents, en se multipliant, en se transmettant, par voie d'enseignement comme par communauté de fréquentation des mêmes ateliers d'écriture et par voie de copie d'un atelier à l'autre, sont devenus des usages consensuels – assez récurrents pour nous permettre de définir une grammaticalité de l'écrit médiéval – fédérant une communauté langagière/scripturaire beaucoup plus large que celle de l'atelier où ils avaient été fabriqués et peut-être à chaque fois recréés. N'y a-t-il pas là une forme de norme, autre que la norme prescriptive – de toute façon chronologiquement seconde : une norme « pragmatique » ?

Et l'évaluation aujourd'hui de ce que cette pratique dit de la conscience linguistique et grammaticale se révèle riche en enseignements : si elle confirme d'une part notre hypothèse selon laquelle la conscience d'une grammaire du français précède les premières réflexions métalinguistiques publiées au XVI^e s., et que la pratique du copiste du ms. Londres, BL, Royal

¹La chercheuse concluant qu'« il n'y a [...] rien à tirer de la soudure des mots », p. 142.

²Ce fondement aléatoire, reposant sur la variabilité des pratiques de séquenciations d'un scribe à l'autre, d'un atelier à l'autre et d'une époque l'autre, constitue le socle de la constatation formulée par Robert Léon Wagner p. 41, p. 74.

Laurent Balon

20 D XI, donne les preuves d'une réflexion grammaticale sur la procédure de « mise en texte » de son manuscrit qui annoncerait une théorisation de l'objet langue française (Balon, à paraître), l'étude des unités graphiques permet, d'autre part, de mettre plus précisément en relief les types de problèmes linguistiques qui se posent au scripteur médiéval et les interrogations du copiste sur le fait même de la langue française.

On en arrivera à la conclusion que les unités graphiques pratiquées en ancien et moyen français apparaissent comme un lieu privilégié de l'observation des phénomènes de lexicalisation et de grammaticalisation du français, et que leur étude en diachronie contribuera à une meilleure approche de ces faits de langue.

1. ÉTUDE DE QUELQUES FORMES DU TYPE « PARTICULE PRÉVERBALE³ EN + VERBE⁴ » :
L'EXEMPLE DE EN PORTER/EMPORTER, DE EN PRENDRE/EMPRENDRE ET DE EN
PENSER/EMPENSER, EN PENSSE/EMPENSSE

Dans la partie du manuscrit de Londres qui consigne *Garin de Monglane*, ces formes présentent la particularité de s'accompagner bien souvent d'ambiguïtés morpho-syntaxiques et sémantiques corrélées à la grammaticalisation de *en* et à leur caractère mouvant d'unités lexicales en cours de formation.

Les formes du type « *en* + verbe » entretiennent ainsi parfois un doute sur la valeur de *en* pouvant être soit un pronom régime clitique, et constituer un élément grammatical distinct du verbe, soit un préfixe se rattachant au verbe dans un lien de dépendance graphique plus ou moins étroit.

Cette différence de traitement serait fonction du degré d'ambiguïté morpho-syntaxique et sémantique sous-jacent à ces formes mettant en relief deux types de formes « *en* + verbe » :

- un premier groupe dont l'écriture, ordinairement soit en séquence, soit en segments, se poserait en termes essentiellement sémantique et lexical comme dans *en prendre* (« prendre une partie d'un tout ») ou « en prendre » (par ex. FM : *en prendre peur*) mais *emprendre* (« entreprendre », « commencer »), la distinction des deux verbes paraissant plutôt aisée en contexte ;
- un second groupe dont l'écriture se poserait en termes d'ambiguïtés sémantiques accrues corrélées à la valeur et au statut de *en*, comme dans *enchacier* (« expulser »), *en chacier* (« chasser de »), *en--chacier*⁵ (« expulser » ou « chasser de » ?), la distinction des deux verbes paraissant plus problématique que pour *en prendre/emprendre*, par ex. (Balon 2012).

N. Andrieux-Reix (2000, p. 58) montre comment, dans la partie du manuscrit de Londres qui consigne le *Moniage Guillaume* (2003b), la perte d'autonomie graphique de *en* passant à *em* dans *emprendre* par fait de grammaticalisation (*en* substantif > *em* préfixe) serait un indice de la soudure des deux éléments *en* et *prendre* par laquelle la séquence *emprendre* vaudrait pour « en prendre », et *enprendre* pour « entreprendre ».

Cette observation se heurte à la pratique du même copiste, dans la partie du manuscrit consignant *Garin de Monglane*, qui se caractérise par la systématisation, à une exception près

³Cette terminologie a été préférée à celle, plus ambiguë, de « pronom personnel régime », même si, ainsi que le précisent N. Andrieux-Reix et S. Monsonégo (1997, p. 300) *en* présente un « fonctionnement référentiel et morpho-syntaxique identique à celui de *me*, *le* etc. »

⁴La catégorie « verbe » incluant ici des participes passés substantivés.

⁵Le tiret redoublé indique une indécision de lecture entre *enchacier* et *en chacier*.

Les unités graphiques de l'Ancien et du Moyen Français

(*enporte*, vers 6587), de la transformation graphique de *en* à *em* employé en séquence avec *p* et de la permanence de la graphie *en* employée entre deux blancs graphiques⁶.

On observe ainsi, dans notre partie du manuscrit, la mise en place d'un système pour ainsi dire inverse à celui décrit par N. Andrieux-Reix à l'endroit du *Moniage* et qui oppose *en prendre* (*Moniage Guillaume : emprendre*) à *emprendre* (*Moniage Guillaume : enprendre*).

Autrement dit, si un phénomène graphique tel que celui décrit par N. Andrieux-Reix peut constituer un indice de différenciation lexicale, il s'observe dans la partie de ce même manuscrit consignant *Garin de Monglane* la mise en place d'un (micro-) système cumulant une dimension paradigmatique (substitution graphique *en* > *em*) et syntagmatique (fragmentation de la chaîne graphique : *en prendre/emprendre*), absent du *Moniage*. Et l'étude des données de *Garin de Monglane* révèle un fonctionnement du blanc graphique comme marque sémantico-morphologique, qu'on peut qualifier de plus grammatical que le système consistant seulement en la stabilité ou la perte d'autonomie graphique de l'élément *en*, parce qu'il assure désormais une meilleure lisibilité dans la distinction des deux verbes « en prendre/en prendre » et « entreprendre/emprendre ».

Ainsi, l'autonomie graphique de *en* devant *p* en séquence ne se rencontre qu'une seule fois dans *Garin de Monglane*, dans *enporte* au vers 6587 vs. 16 occurrences de *emporter* en toutes lettres⁷. On observe en outre qu'aucune des 23 occurrences, en toutes lettres ou abrégées *ēp-* à l'initiale, ne présente une indécision entre une séquenciation et une segmentation.

Si dans cette occurrence de *enporte* la forme a bien le sens d'« emporter » (unité en cours de formation), en suivant le principe mis en relief par N. Andrieux-Reix concernant les graphies en *enp-* dans le *Moniage*, on note que 20 des 22 occurrences restantes en *emp-* s'entendent aussi pour « emporter » à l'exception, sans doute, des occurrences des vers 4779 et 11387.

Autrement dit, la graphie *em-* ne constitue pas ici à elle seule un indice de différenciation lexicale mais une variante combinatoire de la séquence « particule préverbale *en* + *p* » s'ajoutant à la séquenciation représentant elle-même une unité de sens « emporter », conformément au code graphique moderne dans 16 des 17 occurrences en toutes lettres au plan de la graphie proprement dite (soit 94 %, exception : *enporte* vers 6587) et dans 21 occurrences sur 23 au plan de la segmentation de la chaîne graphique en unités significatives (soit 91 %, exceptions : *emporte* vers 4779, 11387). Aussi peut-on poser l'hypothèse que la distinction entre une scription *en porter*, inexistante dans *Garin de Monglane*, et *emporter* apparaît pour le copiste pertinente, peut-être parce que la distinction « *en* pronom adverbial + *porter* » et « *en* particule de sens aspectuel devant un verbe de mouvement » est elle-même bien perçue.

Si l'on considère, par hypothèse mais suivant aussi en cela la pratique du copiste, que les 6 occurrences abrégées *ēp-* valent pour *emp-*, 20 occurrences sur 23 (soit 87 %) seraient conformes au code graphique moderne sous le double rapport « sens/graphie/découpage de la chaîne graphique », la pratique du copiste opposant du même coup clairement *emporter* à *en porter*, qui certes n'apparaît pas mais dont le sens n'est en même temps convoqué que dans 2 occurrences sur 23, au plus (vers 4779, 11387).

Cette observation peut de la même manière être exemplifiée par les formes de *en prendre/emprendre* (unité en cours de formation).

⁶Pour cette raison, nous n'avons pas décompté les cas où, par ex., *en prendre* est abrégé *ē prendre* dans le ms. étant donné que la graphie *em* entre deux blancs graphiques n'est pas autorisée à apparaître selon le système graphique du copiste. En revanche, les graphies du type *ēprendre* ont été décomptées dans un premier temps, de même que celles présentant un « petit espacement » du type *em--prendre*, *ē--prendre*.

⁷L'élément *en* se transformant en *em* employé en séquence avec une forme commençant par *p* ou *b* dans 58 occ. en toutes lettres sur 59 dans *Garin de Monglane*.

Laurent Balon

Pour ces deux graphies, on observe qu'une seule occurrence (vers 11083) sur 24 présente une indécision de script, *em--prendre*, que 19 occurrences apparaissent en segments, *en prendre* (vers 680, par ex.), 4 en séquence, *emprendre* (vers 811, 2213, 5265, 5555). Si l'on considère que l'occurrence du vers 11083 vaut sans doute pour une séquence, alors *emprendre* vaut pour « entreprendre », « commencer », dans 3 occurrences sur 5, *en prendre* valant pour « en prendre », « prendre une partie d'un tout » dans 17 occurrences sur 19.

Si l'emploi d'un signe abrégatif ne permet pas de tirer de conclusion assurée en regard de la graphie attendue en français moderne (*m*) dans 2 occurrences d'emploi en séquence⁸, on observe néanmoins que la forme graphique et le contenu sémantique des verbes *en prendre* et *emprendre* est conforme aux attentes du code moderne, l'unité graphique et lexicale de *emprendre* répondant à l'unité de sens d'« entreprendre », et la discontinuité graphique de *en prendre* répondant de la même manière à la construction attendue en deux éléments séparés en français moderne, la conformité au code moderne tenant compte de la graphie, de la fragmentation de la chaîne écrite en unités significatives en regard du sens attendu concernant au bout du compte 20 des 24 occurrences des verbes (*en*) *prendre* et *emprendre*.

Ainsi, la confrontation des données du manuscrit à l'endroit des binômes *en prendre/emprendre* et *en porter/empporter* atteste que la pratique de la séquenciation, comme de la segmentation, y est corrélée à une nécessité de différenciation lexicale en conséquence de quoi, malgré quelques emplois « fautifs », la pratique tend à entériner la formation de nouvelles unités lexicales, dans 20 occurrences sur 24 pour le premier binôme, dans 22 occurrences sur 24 pour le deuxième.

Cette observation peut de la même manière être exemplifiée par les formes de *en pensser/empensser* (unité en cours de formation) et de *en penssé* (*en* + participe passé masculin substantivé de *pensser*)/*empenssé* (participe passé masculin de *empensser*).

Ces formes présentent, plus encore que pour *en prendre/emprendre* où l'unité alors en cours de formation n'est pas restée dans la langue moderne, des difficultés accrues corrélées au maintien en français moderne du seul verbe *penser* (et de son doublet *panser*, non attesté dans *Garin de Monglane*) et à la difficulté d'appréhension de la graphie (*en/em-*) *penssé* par les lexicographes (substantif ou forme verbale?), ainsi qu'en témoignent les approches contradictoires du *Godefroy*⁹ et du DMF y compris sur des exemples identiques. Soit l'exemple des vers 3012-3013 : *si li demanda / Que ele a empenssé* ; *Godefroy* considérerait plutôt la graphie comme une forme de substantif « ce qu'elle a à l'esprit/dans la tête », le DMF la considérant plutôt comme une forme verbale « ce qu'elle a décidé », entérinant ce faisant sans doute la lexicalisation complète de *empenser*, ce qui devait être le sentiment du copiste.

Dans *Garin de Monglane*, on assiste en effet à la mise en place d'un système graphique visant à discriminer les emplois de *penssé* comme substantif des emplois comme forme verbale. Si quelques occurrences peuvent prêter à discussion en raison de l'interchangeabilité souvent possible en contexte entre « (avoir) à l'esprit » et « (avoir) décidé », comme au vers 3013, on observe que le copiste oppose *en penssé*, qu'il emploie comme substantif dans 2 occurrences sur 2 dans (*venir/avoir*) *en cuer et en penssé* (vers 385, 7090) à *empenssé*, perçue et sans doute employée comme forme verbale dans ses 10 occurrences en toutes lettres, dans 12 occurrences sur 13 (soit 92 %) si l'on considère que les 3 occurrences abrégées *ēp-* valent pour *emp-*¹⁰, dans 12 occurrences sur 14 (soit 86 %) si l'on considère que *ē--penssé* vers 8344 vaut pour une séquence *empenssé*.

⁸Cf. note 6.

⁹Cf. « empenser », III, 58a.

¹⁰L'occurrence du v. 6177, *ēpenssé*, semblant être employée comme substantif (*en penssé*).

Les unités graphiques de l'Ancien et du Moyen Français

Par ailleurs, si le copiste oppose *en penssé* à *empenssé*, il oppose aussi *en penssé*, participe passé masculin substantivé de *pensser*, aux autres tiroirs verbaux du verbe (*en pensser* qui tendent à se spécialiser dans la notion de « soin » (FM *panser*), comme par exemple dans *Se Damediex n'en pensse* vers 613, dans 9 occurrences sur 9 complètement exploitables au plan du découpage graphique, dans 9 occurrences sur 10 (soit 90 %) si l'on considère que *ē--pensse* vers 3270 vaut pour une séquence *empensse*.

La pratique du copiste oppose ainsi, dans 100 % des occurrences complètement exploitables du double point de vue de la graphie et du découpage de la chaîne graphique, d'une part, *en penssé* (*en* + participe passé masculin substantivé de *pensser*) à *empenssé* (participe passé masculin de *empensser*, unité en cours de formation) dans la notion de « pensée » et, d'autre part, *en penssé* aux autres tiroirs verbaux du verbe (*en pensser* dévolus à la notion de « soin », les principes graphiques mis en œuvre y apparaissant comme des signes hautement distinctifs.

Ces observations suggèrent que les accointances entre la pratique du copiste et le code graphique moderne dépassent les caractéristiques ordinairement attachées à l'écriture du français médiéval, la pratique du copiste ici constituant dès lors davantage qu'une forme de « norme "pragmatique" » ainsi que N. Andrieux-Reix (2003a, p. 137) l'avait suggéré à l'endroit de la langue du *Moniage Guillaume* du manuscrit de Londres, la partie du manuscrit contenant *Garin de Monglane* apparaissant au bout du compte plus à même « à former système, ce qui permet de mieux étayer les raisonnements avancés sur certains types de graphies » ainsi que N. Andrieux-Reix (2000, p. 57) l'avait montré à l'endroit de la partie du manuscrit consignant le *Moniage*.

2. DE L'« UTILITÉ GRAMMATICALE » A L'« UTILITÉ LEXICALE » DES SÉQUENCES ET SEGMENTS GRAPHIQUES : UN AFFINEMENT DE LA PERCEPTION DES UNITÉS GRAMMATICALES ET LEXICALES COMME ÉLÉMENTS CONSTITUTIFS DE LA LANGUE ?

Outre cette tendance au fonctionnement du blanc graphique comme marque sémantico-morphologique, ce qui frappe dans le manuscrit de Londres, c'est la réduction considérable du nombre de séquences graphiques et la radicalisation de certaines segmentations.

La pratique de la séquenciation apparaît ainsi restreinte, le copiste segmentant et individualisant souvent les éléments de la chaîne graphique en parties du discours, autrement dit en unités significatives, cette procédure de mise en texte tendant donc vers un usage « moderne » et plus grammatical qu'on le suppose ordinairement.

Ce phénomène suggère une tentative d'organisation en système ainsi que l'avait souligné N. Andrieux-Reix (2000, pp. 56-57) à l'endroit du passage du manuscrit de Londres consignant le *Moniage Guillaume*. Mais, ainsi que nous venons de l'illustrer, on observe entre ces deux différentes parties du manuscrit pourtant rédigées de la même main quelques différences de pratique qui permettent d'aller sans doute ici plus loin dans l'interprétation que l'on peut faire de l'écriture du copiste.

De son côté, P. Rickard (1982) avait bien perçu l'« utilité grammaticale » (*Id.*, p. 486) des séquences graphiques pratiquées en français médiéval dont le premier élément est un mot grammatical dont la fonction, selon le chercheur, est de souligner les « liens syntagmatiques » et « grammaticaux » (*Id.*, p. 470) unissant les différents éléments de la phrase¹¹, « car un mot-outil avait de quoi rassurer le lecteur en restreignant, plus efficacement qu'un mot plein, le champ de son incertitude quant à la catégorie grammaticale du mot suivant » (*Id.*, p. 487).

¹¹À l'endroit de l'ancienne langue, le terme est évidemment employé dans une acception très large.

Mais, si on retrouve dans le manuscrit de Londres l'« utilité grammaticale » de ce type de séquences, on constate surtout, ce qui semble constituer une étape supérieure dans la perception et l'appréhension grammaticale de la langue, une tendance à la fragmentation des groupes syntaxiques en éléments grammaticaux et lexicaux individués ; les mécanismes linguistiques qui régissent ces séquences révélant à la fois une conscience linguistique en émergence et une tendance plus ou moins spontanée à la systématité grammaticale.

2.1 L'exemple de la séquence *sitres* comme indice de lexicalisation

On observe qu'une séquence du type « *et* + E », par exemple *etsouvent*, linguistiquement improductive car incapable tant à indiquer la formation de nouvelles unités lexicales qu'à signifier une quelconque utilité grammaticale au sens décrit par P. Rickard, n'apparaît pas dans *Garin de Monglane*.

En revanche, si la plupart des séquences du type « adverbe + adjectif ou adverbe » du type *tropgrant*, *tresgrant*, *moutgrant*, *tresbien*, n'apparaissent pas non plus¹², on note un emploi assez fréquent de l'une d'entre elles : la séquence « adverbe *si* + adverbe¹³ (ou préfixe) *tres* » qu'on peut interpréter en termes d'indice de lexicalisation.

L'intérêt de la séquence *sitres*¹⁴ (vers 893, 17 occurrences) est d'unir deux éléments parfois considérés comme appartenant à la même classe adverbiale, mais aussi sémantiquement très proches. Or le copiste ne séquence jamais deux éléments appartenant à la même classe grammaticale, excepté lorsque cette séquence fait ou pourrait faire sens au plan du lexique.

Dès lors, la tendance à la systématisation d'une écriture en séquence avec la forme *si*, dont la valeur adverbiale ne fait aucun doute, conduirait à penser que la forme *tres* appartiendrait, pour le copiste, à la classe des préfixes. Mais la quasi synonymie des éléments *si* et *tres* a aussi pu motiver un rapprochement adverbial pour créer une unité hyperbolique à deux termes redondants, la séquenciation des deux éléments *si* et *tres* ne constituant plus dès lors que l'indice d'une unité peut-être sentie par le copiste comme étant en cours de lexicalisation, bien que minoritaire face aux emplois en segments ; la séquenciation des éléments *si* et *tres* ne constituant plus dès lors nécessairement un indice d'une catégorisation de *tres* dans la classe des préfixes.

Ce sentiment lexical, entériné ou non par les dictionnaires et l'histoire ultérieure de l'unité *sitres*¹⁵, pourrait dès lors justifier les segmentations systématiques, par exemple de la catégorie « adverbe + adjectif *grant*, *grande(s)* » (à trois exceptions près¹⁶) : *si grant/grande(s)* vers 1288, 8194 (107 occurrences, dont *si grande(s)* 3 occurrences, mais *sigrant* vers 8185, 8194, *sigrande* vers 3633), *tres grant* vers 2036 (38 occurrences), *mout grant/grande* vers 2563, 5286 (76 occurrences dont *mout grande* 6 occurrences), *trop grant*

¹²Les formes composées du type *tres* + adjectif ont pu être pressenties dans l'histoire du français comme pouvant former une unité grammaticale (cf. Grevisse 1991, p. 147).

¹³C'est par commodité que nous avons classé ici *tres* dans la classe des adverbes, car cette appartenance ne fait pas l'unanimité : Gérard Moignet (1988, p. 270) souligne qu'« il constitue plutôt un préfixe intensif qu'un adverbe ». Philippe Ménard (1994, p. 264) le classe sans restriction parmi la catégorie des adverbes. Susan Baddeley (1998, p. 59) constate que dans des imprimés du début du XVI^e siècle « l'élément *tres* n'est jamais utilisé de manière autonome, mais [qu'il] fonctionne comme un préfixe, agglutiné à un adjectif ».

¹⁴Les variantes *issi*, *aussi* et *ainsi tres* apparaissent toujours en segments, et on relève 43 occ. de *si tres* en segments, 4 occ. indéterminées *si--tres*. – D'une manière générale, on note une tendance à l'emploi en segments de « *si* adverbe », par ex. : *sitost* (v. 3936, 10259, 10587, mais *si--tost* v. 12418, *si tost* v. 752, 12 occ.) ; *sicom(me)* (v. 3712, 12239, mais *si--com(me)* v. 67, 15 occ., *si com(me)*, v. 203, 142 occ.) ; *sique* (v. 2570, 4 occ., mais *si--que* v. 2629, 7 occ., *si que* v. 257, 41 occ.).

¹⁵La forme n'est pas considérée comme une unité lexicale par les dictionnaires d'ancien et de moyen français (cf. Godefroy ; TL ; DMF), contrairement à *sitost*, *sicom* (cf. Godefroy, « *sitost* », X, 679b ; « *sicom* », « *sicum* », VII, 417b), voire *sique* (cf. Hilaire van Daele 1901).

¹⁶*sigrant* v. 8185, 8194, *sigrande* v. 3633, et un cas indéterminé *si--grandes* v. 11677.

Les unités graphiques de l'Ancien et du Moyen Français

vers 2827 (4 occurrences), les cas de cooccurrences « *si + tres* en séquence + adjectif » du type, par exemple, *sitres grant* vers 3410 (17 occurrences), allant de la même manière dans le sens d'une permanence de la séquenciation de *sitres* et de l'isolement de l'adjectif, une séquence du type *sitresgrant* n'étant pas autorisée à apparaître selon le système graphique du copiste.

L'hypothèse d'une contrainte liée à l'environnement graphique pour justifier, contre l'hypothèse de la lexicalisation, l'absence de mise en contact de deux consonnes dans ces formes (soit *sg*, *pg*, *tg*) est invalidée par les exemples de formes composées ou dérivées telles que *trestout*, *itrestant*, *montgravier*, de même que par les cas où la mise en contact de deux consonnes s'accompagne d'une réduction à une seule de ces consonnes, comme dans *tretant* (*tres + tant*), *tretout* (*tres + tout*), *toudroit* (*to(u)t + droit*), *toudis* (*tou(s/z) + dis*). Dans ces conditions, les quelques occurrences de *sigrant* (mais *si grant*, 107 occurrences) ne sauraient remettre en question la non-lexicalisation de la forme, la séquence apparaissant ici davantage comme un vestige d'une pratique traditionnelle facilitée par la suite graphique « voyelle + consonne ».

Or, ce fonctionnement qui semble s'organiser en système est de la même manière perceptible dans l'emploi de *par* comme introducteur de locutions prépositionnelles et adverbiales et comme adverbe intensif dans les constructions du type *tant par aura*.

2.2. Cas de *par* introducteur de locutions prépositionnelles ou adverbiales et *par* adverbe intensif : un fonctionnement linguistique reposant sur un sentiment de la construction et de la lexicalisation

Les séquences graphiques de plus de deux éléments ne comportant pas d'éliision¹⁷, du type *alaterre*, sont exclues de la pratique du copiste si elles ne sont pas sous-tendues par une considération lexicale (formes lexicalisées, unités lexicales en cours de formation ou perçues comme telles par le copiste).

On ne relève ainsi qu'une poignée de formes composées de plus de deux éléments dont la lexicalisation est entérinée par l'histoire ultérieure de ces unités : *maudehaiz* vers 2495 (mais *mal dehé* vers 5298), *nonpourquant* vers 4174 (mais *non por quant* vers 3567), *debon(n)airement* vers 305 (10 occurrences), ainsi que sans doute *pardevers* vers 9514 et peut-être *ensifaitierement* vers 2914, l'occurrence de *en si* en segments vers 5578 et l'occurrence de *ainsi faitierement* vers 9235 suggérant la perception d'une forme *en-si-faitierement* à trois éléments, et *debonaire* vers 277, si tant est que cette forme ne masque pas en réalité la présence d'une apocope *bon'aire*, l'accord au féminin de l'adjectif *bone* avec *aire*, variable en genre selon les manuscrits, étant possible¹⁸ par ailleurs mais impossible à déterminer ici.

Cette réduction considérable des séquences de plus de deux termes, restreintes, qui plus est, à des formes de composition entérinées par l'histoire, apparaît comme un indice de modernité dans la production de l'écrit.

S. Baddeley (1998, p. 66) souligne en effet, qu'« on ne trouve plus à cette époque [au début du XVI^e siècle], ou presque plus, de cas d'agglutinations de plusieurs termes (sinon dans certaines locutions adverbiales, du type *cepourquoy* [...], ni d'agglutinations de mots pleins (mis à part les noms et verbes composés). »

2.2.1 *par* introducteur de locutions prépositionnelles ou adverbiales du type *par dessus*

¹⁷On rencontre, à ce titre, des séquenciations à trois termes dans le ms. type *dencoste* (*de* élidé + *en* + *coste*) ou *dalez* (*de* élidé + *a* + *lez*), dont la lexicalisation semble en outre déjà assurée (cf. *TL*, « dalez » II, 1172, 9).

¹⁸Cf. *TL*, « aire », I, 253, 5.

Laurent Balon

La quasi inexistence de certaines séquenciations composées de plus de deux éléments apparaît en cela caractéristique qu'elle s'observe y compris dans des cas où une séquenciation pourrait faire sens et annoncer une nouvelle unité en cours de lexicalisation, particulièrement dans le cas des locutions introduites par *par* comme, par ex., dans *par dedevant* vers 2108 (3 éléments), mais *pardevant* vers 5737 (2 éléments).

Ce principe que nous pouvons supposer de lisibilité semble ne constituer un frein qu'à la scription de séquences d'au moins trois éléments. On constate ce faisant que si *par* suivi d'un élément apparaît aussi bien en séquence qu'en segments (*pardevant* mais *par devant* vers 1120), il apparaît systématiquement entre deux blancs graphiques lorsqu'il est employé dans une construction à trois ou quatre éléments pouvant former une unité. On note ainsi que dans les constructions originellement à trois termes, la séquenciation s'effectue sans aucune exception (29 occurrences) :

- entre le deuxième et le troisième élément dans : *par dessus* vers 1789, 4037, 6673, 8328, 9835, 11183, 11454, *par desouz* vers 1235, 6393, 10124, 10566, 10689, 10881, 11644, *par deça* vers 2548, 10908, *par dela* vers 6637, 6917, 9909, 10910, *par delez/par dalez*¹⁹ vers 7581, 11180, 11958, *par decoste* vers 587, 836, *par dedenz* vers 10565, *par devers* vers 3473, *par enval* vers 4790, *par dehors* vers 2987.

- entre le troisième et le quatrième élément dans les constructions originellement à quatre termes : *par de delez* vers 3373, *par de dessus* vers 5348, *par de desouz* vers 9605, *par de dessous* vers 11475, *par de devant* vers 5085, mais, exception : *par dedevant* vers 2108.

Cette systématisme d'emploi de *par* entre deux blancs graphiques semble pouvoir souligner l'existence d'un seuil d'illisibilité qui prévaudrait sur toute autre considération, le détachement quasi systématique de *de*, apparaissant lui aussi entre deux blancs, s'inscrivant dans ce même objectif assurant dès lors une lisibilité totale à l'ensemble de la locution.

Ce souci de lisibilité pourrait être corrélé à l'emploi en cooccurrence d'éléments relevant de la même classe grammaticale qui empêcherait toute séquenciation.

On observe en effet, d'une part que dans ces locutions prépositionnelles le copiste ne séquencie jamais plus de deux éléments à la suite susceptibles d'être individués, d'autre part que l'emploi conjoint des deux prépositions *par* et *de* rend leur séquenciation impossible.

L'unique réelle exception²⁰ à cette constatation, *pardevers* vers 9514, peut dès lors être interprétée en termes de lexicalisation complète de *devers*, ainsi que semble d'ailleurs en attester l'absence dans le manuscrit d'un emploi en segments *de vers*. Or, *devers* constituant désormais une unité et non plus une simple séquenciation de deux éléments sécables, *pardevers* équivalant à la séquence de deux éléments et non plus de trois, l'occurrence devient donc possible à l'image de *pardevant*, *devers* et *devant* étant sans doute désormais perçues comme des formes lexicales à part entière. Pour le copiste, une séquenciation « préposition *par* + préposition *de* » semble n'être grammaticalement acceptable que dans ce cas.

La possible séquenciation de deux éléments du type *pardevant* et l'absence de séquenciation de trois éléments dans ce type de locutions apparaît en cela caractéristique que « dans le cas des locutions [...] l'écriture soudée devait représenter non seulement un rapprochement avec la prononciation, mais aussi un « groupe de sens », plusieurs éléments fonctionnant comme un tout, d'un point de vue syntaxique » (Baddeley 1998, p. 54).

¹⁹La construction est initialement sans doute composée de quatre éléments *par + de + a + lez*, mais l'élision de *de* devant *a* réduit de fait le nombre d'éléments à trois à l'image de *par delez*, justifiant que l'on ait classé cette forme dans les constructions à trois termes.

²⁰On relève aussi *pardevant* v. 5737, mais la forme *devant* apparaît comme une unité stable et insécable.

Les unités graphiques de l'Ancien et du Moyen Français

Il apparaît ainsi que ce fonctionnement « comme un tout » s'observe dans le manuscrit de Londres jusqu'à un certain degré de lisibilité ; le dépassement d'un seuil de lisibilité semblant entraîner l'éclatement en plusieurs segments de ce qui sous ce seuil pouvait être transcrit en un seul tenant.

Il peut de la même manière s'observer une séquenciation *dede-*, mais possible uniquement dans le même contexte, avec les mêmes éléments et pour la même locution : *par dedevant*. Ainsi, encore une fois, seule la lexicalisation de *devant* paraît seule autoriser la suite *dede-* dans une séquence *dedevant*, tout en interdisant une séquence *parde-* que l'absence d'occurrence du type *pardedevant*, *pardedevers* semble attester.

Il semble ainsi se dégager de la pratique du copiste deux principes de fonctionnement linguistiques opposés reposant pourtant sur des matériaux²¹ identiques : d'une part, un sentiment de la construction des unités attesté par des phénomènes de segmentations particulières et, d'autre part, un sentiment de la lexicalisation attesté par des phénomènes réguliers de séquenciations, la préposition *par* conservant dans ces constructions sa nature prépositionnelle, alors que *de*, selon la portée grammaticale et lexicale que le copiste veut lui donner, conserve tantôt sa valeur première de préposition, tantôt devient l'instrument d'une préfixation.

Or ce fonctionnement est de la même manière perceptible dans l'emploi de *par* comme adverbe dans les constructions du type *tant par aura*.

2.2.2 *par* adverbe intensif : l'exemple des constructions du type « *mout/tant + par + verbe* »

Dans ce type de construction, on remarque encore que les deux adverbes ne sont jamais séquenciés, *par* se retrouvant en outre entre deux blancs graphiques dans 38 des 40 occurrences présentes dans le manuscrit, à l'exception de *mout parfu* vers 5809, 12166, non signifiante, sans doute, au plan du lexique. On constate en outre que *par* demeure entre deux blancs y compris dans des unités pourtant considérées comme lexicalisées et lemmatisées comme telles dans les dictionnaires d'ancien et de moyen français²² dans :

. que Diex par ama tant vers 6 ; que il par ama tant vers 8063 ; dont il par avoit tant vers 9455²³.

Cette lisibilité de l'écrit apparaît ici d'autant plus surprenante que les trois exemples présentent une construction où l'un des deux adverbes est pourtant disjoint de l'autre dans une construction « adverbe + verbe + adverbe ». L'ordre des mots ne semble donc pas avoir d'influence, d'où l'hypothèse que la cooccurrence de deux adverbes intensifs employés conjointement ou disjointement induirait un emploi en segments de *par amer* et de *par avoir*. Cette hypothèse semble pouvoir se vérifier dans les deux occurrences présentant une séquence « *par + verbe* » attestée en tant qu'unité lexicale dans :

. *quant il parot .V. anz*²⁴ vers 12242 ; *mout se partint*²⁵ *pres de son destre costal* vers 3453.

Dans le premier cas, la séquenciation *parot* pourrait être motivée par l'absence, dans la proposition, d'un élément « adverbe » qui provoquerait une agglutination *paravoir*.

Dans le second cas, l'emploi d'une construction pronominale interposant *se* entre les deux adverbes pourrait neutraliser l'effet cumulatif engendré par l'occurrence des intensifs *mout* et

²¹Matériaux ici représentés par les prépositions *par* et *de*.

²²Cf., par ex., *TL*, « paramer », VII, 198, 51 ; « paravoir », VII, 199, 51.

²³« dont il était en pleine possession ».

²⁴« à cinq ans révolus ».

²⁵« se tenir complètement ».

par en libérant ce dernier de son emploi strictement adverbial, *par* devenant dès lors l'instrument d'une préfixation lexicale, engendrant, presque nécessairement, une séquenciation *partint*, l'emploi en segments *mout se par tint pres de* devenant sinon illisible, du moins freinant l'intelligibilité de la proposition en raison d'une segmentation qui deviendrait excessive.

Ces deux hypothèses peuvent se trouver confortées par l'occurrence de *si li ont par--tolu* du vers 1669 : dans ce cas, les éléments *par* et *tolu* ne sont concrètement employés ni en séquence ni en segments, mais dans un « entre-deux » (signalé par un tiret redoublé) qui pourrait manifester de l'embarras du copiste en regard des observations formulées ci-dessus. La construction apparaît en effet en prise à deux principes contradictoires : l'absence d'un second élément « adverbe », entraînant partout ailleurs une séquenciation, se trouve ici en partie contrecarrée par l'inexistence dans le lexique d'un verbe *partolir* en tant que tel²⁶. La segmentation incertaine, caractérisée ici par un petit espacement entre *par* et *tolu*, pourrait dès lors constituer une sorte de « non-choix » réflexif.

3. CONCLUSION

Il apparaît ainsi que les unités et segments graphiques du *Garin de Monglane* du manuscrit de Londres apparaissent conditionnés par des critères d'utilité grammaticale et lexicale, leur agencement syntaxique reposant sur la lisibilité et l'intelligibilité de l'ensemble, mais aussi sur la compatibilité grammaticale des différents éléments entre eux au plan d'une certaine idée que se fait le copiste de la grammaire et qui devait connaître quelque fortune dans l'histoire ultérieure de la langue.

Ce fruit d'une pratique individuelle d'écriture invite à un examen portant sur un corpus de textes plus vaste permettant de se faire une idée plus fiable de la compétence des locuteurs/scripteurs, de la chronologie du changement linguistique, de considérer les facteurs corrélés à l'évolution et à terme de réexaminer les hypothèses expliquant le changement.

Ainsi, en comparant deux manuscrits du *Roman de la Rose*, on observe que le manuscrit du XIII^e siècle (BNF fr. 1573) pratique la séquence *tresgrant* mais, qu'en revanche, le manuscrit du XV^e siècle (Bibl. Sainte Geneviève 1127), qui a considérablement réduit le nombre de séquences graphiques, a généralement remplacé la séquence *tresgrant* par *mout grant* en deux segments, ce qui permet d'en déduire deux choses : d'une part, que l'élément *tres* était bien compris au XV^e siècle comme un synonyme de l'adverbe *mout*, mais que le statut grammatical encore incertain de *tres*, ainsi que nous l'avons évoqué à l'endroit de la pratique du manuscrit de Londres, lui a fait préférer l'emploi de *mout*, adverbe depuis les origines ; d'autre part, que l'étude des séquences et segments graphiques en diachronie contribuera à une meilleure approche des phénomènes de lexicalisation et de grammaticalisation.

²⁶Or un vocable « partolir » pourrait faire sens : « enlever la possibilité de », « empêcher complètement », « dissuader ».

Les unités graphiques de l'Ancien et du Moyen Français

BIBLIOGRAPHIE

Sources primaires

Londres, BL, Royal 20 D XI (XIV^e s.) : *Garin de Monglane, Moniage Guillaume*

Paris, BNF, fr. 1573 (XIII^e s.) : *Roman de la Rose*

Paris, Bibl. Sainte Geneviève, 1127 (XV^e s.) : *Roman de la Rose*

Sources secondaires

ANDRIEUX-REIX, Nelly (1999). « Le manuscrit 150 de Valenciennes, f^o141 v^o : premières images des mots graphiques médiévaux », HERBIN, Jean-Christophe (éd.), *Actes de la Journée d'Étude du 3 avril 1998*, Valenciennes, Université de Valenciennes, Éditions Lez Valenciennes, 25, Presses universitaires de Valenciennes, 9-23.

ANDRIEUX-REIX, Nelly (2000). « Transcription, lisibilité, transgression : quelques problèmes posés par les éditions de textes médiévaux », BURIDANT, Claude (éd.), *Le moyen français. Le traitement du texte (édition, appareil critique, glossaire, traitement électronique)*. *Actes du IX^e colloque international sur le moyen français*, Strasbourg, Université Marc Bloch, 29-31 mai 1997, Strasbourg, Presses Universitaires de Strasbourg, 55-63.

ANDRIEUX-REIX, Nelly (2003a). « Ateliers d'écriture et fabriques individuelles de normes collectives du français au Moyen Âge », *Faits de Langues* 22, Paris, Éditions OPHRYS, 129-137.

ANDRIEUX-REIX, Nelly (2003b). *Moniage Guillaume. Édition de la rédaction longue*, Paris, H. Champion.

ANDRIEUX-REIX, Nelly, MONSONÉGO, Simone (1997). « Écrire des phrases au Moyen Âge. Matériaux et premières réflexions pour une étude des segments graphiques observés dans des manuscrits français médiévaux », *Romania* 115, 289-336.

ANDRIEUX-REIX, Nelly, MONSONÉGO, Simone (1998). « Les unités graphiques du français médiéval : mots et syntagmes, des représentations mouvantes et problématiques », ANDRIEUX-REIX, Nelly et MONSONÉGO, Simone (éd.), *Segments graphiques du français. Pratiques et normalisations dans l'histoire*, *Langue française* 119, 30-51.

BADDELEY, Susan (1998). « Théorie et pratique de la segmentation graphique dans les textes français du premier tiers du XVI^e siècle », *Langue française* 119, 52-68.

BALON, Laurent (2012). « Des indices probants de l'existence d'une « grammaire » du français dans de l'écrit littéraire dès le début du XIV^e siècle ? », COLOMBAT, Bernard, FOURNIER, Jean-Marie et RABY, Valérie (éd.), *Actes du colloque SHESL-HTL 2011*, Paris, H. Champion, coll. « linguistique historique », 673-688.

BALON, Laurent (à paraître). *Garin de Monglane*, chanson de geste du XIII^e siècle éditée d'après le manuscrit de Londres, Éditions Classiques Garnier, coll. « Textes Littéraires du Moyen Âge ».

BIEDERMANN-PASQUES, Liselotte (1998). « Des segmentations particulières d'un incunable (1488) à l'écriture du français en unités lexicales et grammaticales », *Langue française*, 119, 69-87.

CARERI, Maria *et al.* (2001). *Album de manuscrits français du XIII^e siècle : mise en page et mise en texte*, édit. CNRS et Institut de Recherche et d'Histoire des Textes, Roma, Viella.

CATACH, Nina (1998). « Les signes graphiques du mot à travers l'histoire », *Langue française* 119, 11-23.

DAELE, Hilaire van (1901). *Petit dictionnaire de l'ancien français*, Paris, Garnier frères.

Laurent Balon

- GASPARRI, Françoise, HASENOHR, Geneviève, RUBY, Christine (1993). « De l'écriture à la lecture : réflexions sur les manuscrits d'*Érec et Énide* », *Les manuscrits de Chrétien de Troyes*, édit. K. Busby, 2 vol., Amsterdam, t. 1, 97-148.
- GODEFROY, Frédéric (1982). *Dictionnaire de l'ancien français et de tous ses dialectes du IX^e siècle au XV^e siècle*, 10 t., Paris-Genève, Slatkine, (1^{re} éd., Vieweg, 1881/1902) [Godefroy].
- GREVISSE, Maurice (1991). *Le bon usage*, 12^e éd. refondue par André GOOSSE, Paris, Duculot, (1^{re} éd., 1936).
- LEWICKA, Halina (1963). « Réflexions théoriques sur la composition des mots en ancien et en moyen français », *Kwartalnik neofilologiczny* 10, 2, 131-142.
- LLAMAS POMBO, Elena (2003). « Séquences graphiques dans le *Roman de Fauvel* (BNF, fonds fr. 146) », *Thélème, Revista Complutense de Estudios Franceses*, Numéro Extraordinario, Madrid, Ediciones de la Universidad Complutense, 225-236.
- PELLAT, Jean-Christophe (1998). « Les mots graphiques dans des manuscrits et des imprimés du XVII^e siècle », *Langue française* 119, 88-104.
- MARTIN, Robert *et al.*, *Dictionnaire du moyen français (1330-1500)*, version 2012 (DMF 2012). ATILF - CNRS et Université de Lorraine ; dictionnaire électronique, <http://www.atilf.fr/dmf/> [DMF].
- MÉNARD, Philippe (1994). *Syntaxe de l'ancien français*, Bordeaux, Bière, 4^e éd.
- MOIGNET, Gérard (1988). *Grammaire de l'ancien français*, Paris, Klincksieck, 2^e éd.
- RICKARD, Peter (1982). « Système ou arbitraire ? Quelques réflexions sur la soudure des mots dans les manuscrits français du Moyen Âge », *Romania* 103, 470-512.
- SAENGER, Paul H. (1997), *Space Between Words: The Origins of Silent Reading* (Series: *Figurae: Reading Medieval Culture*), Stanford, California, Stanford University Press.
- SEGUIN, Jean-Pierre (1998). « Les incertitudes du mot graphique au XVIII^e siècle », *Langue française* 119, 105-124.
- TOBLER, Adolf, Lommatzsch Erhard, *Altfranzösisches Wörterbuch*, Wiesbaden, 1915-1995 pour la 1^{ère} éd. ; éd. électronique, Stuttgart, Franz Steiner Verlag, 2002 [TL].
- WAGNER, Robert-Léon (1974). *L'ancien français. Points de vue. Programmes*. Paris, Larousse.