

Étude statistique de la syntaxe des lexèmes *chetif, dolent, las et maleüré* en français médiéval

Nicolas Mazziotta¹

¹Université de Liège – Belgique/Universität de Stuttgart – Allemagne

Abstract

In medieval French, the four lexemes *chetif*, *dolent*, *las* and *maleüré* can be used as adjectives as well as as so-called « interjections » (which we prefer to call *locutives*) to express distress and psychological pain. Multiple correspondance analyses performed on our corpus allow us to highlight that these lexemes each have specific morphosyntactic behaviour. Further analyses on the non-locutive occurrences on the one hand and locutive occurrences on the other hand outputs two additional observations ; namely, the *X* element in the construction *s'apeler X* behaves like a locutive (because it is clearly a *de*-locutive construction), and the rise of *helas* as a replacement for all the lexemes at study at the end of Middle-Ages.

Résumé

En français médiéval, les lexèmes *chetif*, *dolent*, *las* et *maleüré* ont la capacité d'être employés comme des adjectifs ou comme des « interjections » (que nous nommons plutôt *locutifs*) exprimant le désarroi et la douleur psychologique. L'analyse des paramètres morphosyntaxiques associés aux différents emplois avec l'ACM montre que les différents lexèmes ont un potentiel syntaxique commun, mais que chacun a tendance à se spécialiser dans une ou plusieurs fonctions spécifiques. L'examen séparé des emplois locutifs et des emplois non locutifs met en outre en évidence, respectivement, le développement du lexème *helas*, à la fin de la période médiévale pour remplacer les autres lexèmes étudiés ici et la construction *s'apeler X*, dont l'élément *X* partage des propriétés communes avec les emplois locutifs, en vertu de son origine délocutive.

Key words : syntax, medieval French, locutive, delocutive, MCA.

Mots-clés : syntaxe, français médiéval, locutif, délocutif, ACM.

1 Introduction

Cette étude porte sur les lexèmes permettant d'exprimer le désespoir du locuteur en français médiéval : *chetif* lit. 'captif', *dolent* lit. 'qui éprouve de la douleur', *las* lit. 'fatigué' et *maleüré* lit. 'infortuné'. Dans une étude préalable (Mazziotta, soumis), nous avons traité conjointement ces lexèmes et mis en évidence les caractéristiques communes de leur comportement syntaxique. Ces lexèmes sont tantôt employés comme des adjectifs (ils se rapportent alors à des noms ou sont substantivés), tantôt comme des *locutifs* (Kahane et Mazziotta, 2016), c'est-à-dire comme des unités possédant une autonomie énonciative¹.

(1) si dist (...) : « Ha ! *lasse*, con sui malbaillie ! » (*PercevalKu*, p.362d, v.405)
“Ainsi, elle dit : « Ha ! Malheur ! Comme je suis maltraitée ! »”

Quand on observe (1), le mot *lasse* pourrait apparaître seul. Il est autonome par rapport à l'« interjection » qui le précède autant que par rapport à la phrase exclamative qui le suit. Cette autonomie est due au fait que *lasse* vise nécessairement à l'expression d'un contenu

¹Tous les extraits cités sont issus de la *Base de français médiéval* (BFM ; v. sec. 2.1.1) ; les références utilisent les sigles propres à cette base de données.

expressif. Contrairement à une phrase comme *Il marche*, qui peut être transformée en *Est-ce qu'il marche ?* (question), *Qu'il marche !* (injonction) ou *Comme il marche !* (exclamation), *Lasse* ne peut consister qu'en une exclamation et ne peut être dérivé de manière à changer le type de visée illocutoire (c'est-à-dire l'association à la recherche d'un effet pragmatique : déclarer, décrire, s'engager, donner un ordre ou exprimer une émotion ; Searle/Vanderveken, 1985 : 37-40 et 51-62). Il ne peut pas non plus être nié. Nous disons que ce type d'emploi, désigné du terme *locutif*, se caractérise par une visée illocutoire figée. Ce type de comportement syntaxique et pragmatique dépasse les relations classiques de dépendance syntaxique. En cela, il contraste avec les fonctions de modifieur de nom (« épithète » ; p.ex. (2), où *dolente* modifie *femme*), d'attribut du sujet (p.ex. (3), où *las* est attribut de *il* au moyen de la médiation par le verbe copule *est*), d'attribut de l'objet (p.ex. (4), où le verbe *faire* met en relation le nom *mere* et l'adjectif) et de terme principal du syntagme nominal (p.ex. (5), où la présence d'un article marque l'emploi « substantivé »).

(2) Ne semble pas femme *dolente*. (*fauvel*, p.72, v.1912)
 “Elle ne paraît pas (être) une femme qui souffre.”

(3) il est de veillier trop *las*. (*adhalefeuill*, p.23, v.548)
 “Il est trop fatigué de veiller.”

(4) [...] et fist *dolante* mainte mere. (*eracle*, p.158, v.5118)
 “Et il rendit de nombreuses mères malheureuses.”

(5) Ne seit que faire *la chaitive*. (*adgar*, p.311, v.267)
 “La malheureuse ne sait que faire.”

Ces fonctions sont proprement rectionnelles (ou « microsyntaxiques ») alors que les locutifs ont une fonction qui dépasse les cadres rectionnels (ou « macrosyntaxique », cf. Blanche-Benveniste, 1997), auxquels ils restent cependant intégrés. Les locutifs fondent ainsi l'*unité illocutoire* (cf. Benzitoun *et al.*, 2010 : 2077-2078). Les lexèmes étudiés peuvent être employés de ces différentes manières. L'approche intuitive des textes laisse penser que ces quatre lexèmes ont une distribution syntaxique similaire et ont apparemment les mêmes capacités morphosyntaxiques, au point d'être interchangeables. Il est cependant connu (Opermann-Marsaux, 2013 : 84) que la distribution de *las* se distingue progressivement de celle des autres lexèmes au point d'intégrer le paradigme des « interjections » pures. Pour évaluer cet aspect, l'approche philologique traditionnelle est impuissante : tout au plus peut-elle déceler des différences de fréquence absolue. Pour nous aider à résoudre cette question, nous faisons ici usage de l'analyse des données. Nous effectuerons fréquemment des retours à la théorie linguistique et aux attestations.

Nous commençons par présenter la manière dont ces différentes fonctions se caractérisent les unes par rapport aux autres d'un point de vue morphosyntaxique au moyen d'un traitement global (sec. 2). Nous traitons ensuite de manière séparée les emplois non locutifs et les emplois locutifs pour affiner et enrichir les premières conclusions (sec. 3). La sec. 4 synthétise nos conclusions.

2 Traitement global

Dans cette section, nous commençons par présenter notre corpus et la manière dont nous avons extrait et annoté les données (sec. 2.1) avant d'en détailler le traitement (sec. 2.2).

2.1 Sélection et préparation des données

Nous décrivons dans un premier temps le corpus choisi et la méthode d'extraction (sec. 2.1.1), puis nous expliquons quelles annotations nous avons ajoutées aux données extraites (sec. 2.1.2).

2.1.1 *Corpus et extraction*

Les données de cette étude sont extraites de la version 2014 de la *Base de français médiéval* (BFM 2014 ; v. Guillot *et al.*, 2007) au moyen d'expressions régulières². Les occurrences-mots de la BFM ne sont pas lemmatisées. Il est par conséquent impossible d'extraire les observables en interrogeant la base de données sur les lexèmes qu'elle contient. Les difficultés propres à l'extraction des formes non lemmatisées dans un continuum de variations présentant une instabilité graphique bien connue sont exposées de manière synthétique dans Prévost, 2011b : §2.1.1. L'extraction des quatre lexèmes qui nous intéressent ici génère une liste qui, après filtrage manuel des résultats pour éliminer les collisions homonymiques et rejet de 198 occurrences non pertinentes (majoritairement des actualisations du verbe *lasser* et du substantif *lac* "nœud"), contient 1501 occurrences³.

2.1.2 *Variables descriptives*

Chaque occurrence retenue reçoit une description effectuée manuellement par nos soins consistant en cinq variables linguistiques.

Lexème. Chaque occurrence reçoit un lemme correspondant au lexème qu'elle actualise. Étant donné l'absence d'homonymie en ce qui concerne les formes sur lesquelles cette étude est basée, la lemmatisation *a posteriori* des occurrences se fait automatiquement sur la base des expressions régulières employées pour l'extraction. On a donc quatre modalités : *chetif*, *dolent*, *las*, *maleüre*. Notons que les formes du type *malheureus* sont considérées comme assimilables au lexème *maleüre* dans le cadre de cette étude ; cela n'est pas tout à fait exact d'un point de vue lexical, mais cela permet de pallier partiellement le manque d'attestations pour ce lexème.

Emploi. La variable décrit l'emploi syntaxique dans lequel l'individu est observé, tel que présenté ci-dessus (sec. 1). Cinq modalités sont possibles : *Mod*, c'est-à-dire « modifieur » d'un nom ; *AtSj* « attribut du sujet » ; *AtObj* « attribut de l'objet », identifié par une relation de coréférence avec un complément « d'objet direct » ; *TN* « tête nominale » ; *Locutif*, comme

² Les expressions régulières employées pour les relevés sont "(ch?(ai?|e)ti(f|v).*)"%c, , "(dou?l(e|a)n.*)"%c", (las(se)?s?)"%c, "(ma(l|u)(e|a)(u|ü)r.*)"%c.

³En considérant chaque occurrence des lexèmes observés comme un individu, on recense donc 1501 lignes. On pourrait, pour neutraliser la variation stylistique éventuelle, considérer uniquement la présence de chaque profil dans chaque œuvre ou chaque auteur (et éviter de la sorte les biais constitués par les œuvres/auteur qui favorisent un emploi spécifique). Opérer de la sorte réduirait considérablement la population et laisserait de toute manière non résolue la question de la spécificité liée au genre de texte – question qui reste par ailleurs peu facile à traiter étant donné que les genres littéraires ne sont pas représentés de manière homogène et que les lexèmes concernés sont relativement rares : 1501 attestations relevées sur près de 3 550 000 occurrences-mots (selon la documentation de BFM 2014). Les plans obtenus sur une analyse factorielle neutralisant la fréquence par œuvre sont par ailleurs très similaires à ceux présentés ici. Voir Prévost, 2011a : 47-49 (notamment) concernant les problèmes spécifiques à la représentativité des corpus en français médiéval.

dans (1), identifié par la présence d'une visée illocutoire intrinsèque figée, en l'occurrence expressive.

Complément. Cette variable binaire décrit la présence d'un complément du mot observé, comme *dou dommage qui...* complémente *dolans* dans (6) (les compléments peuvent être de forme variée : introduits par *de, que, pour*, etc.). La modalité 1 indique la présence d'un pareil complément, 0 indique son absence (autres exemples ci-dessus).

(6) si fu moult *dolans dou dommage qui avoit esté* [...] (*moree*, p.374)
 “Et il fut très affecté du tort qui avait été fait.”

Intensif. Cette variable binaire décrit la présence d'un adverbe exprimant l'intensité qui modifie le mot observé, comme dans (3), où l'adverbe *trop* modifie *las*. La modalité 1 indique la présence d'un pareil adverbe, 0 indique son absence.

Combinaison. La modalité 1 indique que le terme est associé à une autre occurrence d'un des quatre lexèmes analysés, comme dans (7). 0 indique que le lexème est autonome de ce point de vue.

(7) *Chaitis, dolanz, com mar i mui, quant ge an ceste nef entrai* ! (*eneas1*, p.177, v.5808)
 “Comme je suis malheureux ! Comme je m'y suis rendu pour mon malheur, quand je suis entré dans ce bateau !”

On aura donc, pour les exemples ci-dessus, les lignes suivantes :

n.	Lexème	Emploi	Complément	Intensif	Combinaison
1	las	Locutif	0	0	0
2	dolent	Mod	0	0	0
3	las	AtSj	0	1	0
4	dolent	AtObj	0	0	0
5	Chetif	TN	0	0	0
6	dolent	AtSj	1	1	0
7	chetif	Locutif	0	0	1

Table 1: Exemples d'individus

La combinaison des modalités génère un total de $4 \times 5 \times 2 \times 2 \times 2 = 160$ lignes possibles.

2.2 Traitement et interprétation

À première vue, les lexèmes se distinguent tout d'abord les uns des autres par leur fréquence. Les associations deux à deux entre la variable *Lexème* et les autres variables sont toutes significatives quand les effectifs sont suffisants pour être testés (test du chi-carré, sans tenir compte des évaluations multiples)⁴. Étant donné le caractère qualitatif des données observées, la faible fréquence de certaines associations de modalités et le grand nombre de cases vides, nous avons choisi de traiter le problème au moyen d'analyses factorielles des correspondances multiples (ACM)⁵.

⁴ Tous les calculs sont effectués à l'aide du logiciel R (R Core team, 2015).

⁵ Les analyses multidimensionnelles sont effectuées à l'aide du paquet FactoMineR (Lê *et al.*, 2008).

Au vu du nombre de modalités considérées, l'ACM repose sur l'analyse de la répartition de 160 profils ; seuls 68 d'entre eux sont attestés. Ceux dont la fréquence relative cumulée avoisine 50% sont reportés dans la tab. 3 (nous concaténons simplement les valeurs des modalités dans l'ordre où nous les avons présentés dans la sec. 2.1.2). On voit que les lexèmes *las* et *dolent* représentent à eux deux presque 50 % des occurrences et que si les associations des différentes valeurs des autres modalités sont variées, *las* se distingue déjà par une fréquente association à la modalité *Locutif* et, dans une moindre mesure, à la combinaison. *Dolent* est quant à lui souvent associé à *AtSj*.

Lexème	Complément	Combinaison	Emploi		AtObj		AtSj		Locutif		Mod		TN	
			Intensif		0	1	0	1	0	1	0	1	0	1
chetif	0	0		15	0	11	1	15	0	48	0	94	0	
		1		2	0	1	0	21	0	2	0	3	0	
	1	0		0	0	4	0	0	0	1	0	2	0	
dolent	0	0		30	11	73	111	17	0	72	40	50	0	
		1		9	0	2	1	20	0	1	1	2	0	
	1	0		9	0	84	145	0	0	1	3	3	0	
las	0	0		11	3	42	32	257	0	30	4	32	0	
		1		8	0	1	1	53	0	5	0	5	0	
	1	0		3	0	19	7	0	0	0	0	10	0	
maleüre	0	0		2	0	3	6	3	0	17	1	12	0	
		1		1	0	4	0	13	0	0	0	3	0	
	1	0		0	0	3	0	0	0	0	0	0	0	
		1		0	0	0	0	0	0	0	0	0	0	

Table 2: Tableau de contingence complet (5 variables)

Lexème-Emploi-Complément-Intensif-Combinaison	Fq. abs.	Cum. abs.	Fq. %	Cum. %
las-Locutif-0-0-0	257	257	17.12	17.12
dolent-AtSj-1-1-0	145	402	9.66	26.78
dolent-AtSj-0-1-0	111	513	7.40	34.18
chetif-TN-0-0-0	94	607	6.26	40.44
dolent-AtSj-1-0-0	84	691	5.60	46.04
dolent-AtSj-0-0-0	73	764	4.86	50.90

Table 3: Profils les plus fréquents Lexème-Emploi-Complément-Intensif-Combinaison

En ne conservant que les modalités dont la contribution est suffisante⁶, on obtient un plan factoriel (fig. 1, à gauche) hautement explicatif (99,13 % de l'inertie ajustée⁷) qui positionne chacun des quatre lexèmes dans un ensemble de modalités spécifique concordant tout à fait

⁶ Les modalités nommées sur les graphes sont celles dont la contribution est supérieure ou égale à la contribution moyenne du facteur auquel elles contribuent (cf. Le Roux/Rouanet, 2010 : 52).

⁷ Nous calculons de pseudo-valeurs propres $\tau(\lambda)$ sur la base de chaque valeur propre λ du tableau disjonctif et du nombre de variables actives Q selon la formule $\tau(\lambda) = [Q/(Q - 1)]^2 (\lambda - 1/Q)^2$ (Benzécri, 1979). Cet ajustement permet de ne pas se fier aux valeurs λ obtenues à partir du tableau de Burt de l'ACM, qui sont artificiellement pessimistes pour des raisons mathématiques (voir Lebart *et al.*, 1997 : 368).

avec les profils les plus fréquents de la tab. 3. Sur le premier axe, c'est la variable *Emploi* qui structure le plus la dimension. Elle oppose surtout les modalités *Locutif* et *AtSj*. Les autres modalités et les autres variables sont aussi très bien représentées, ce qui nous permet de poser les constats suivants. L'évaluation de la présence de compléments et intensifs dépendants montre les locutifs les rejettent, ce qui confirme les caractéristiques de cette classe (Kahane et Mazziotta, 2016), à l'inverse des attributs du sujet. Ces rejets sont d'ailleurs systématiques pour les locutifs (tab. 2). De même, les têtes nominales ne sont, comme on s'y attend, pas enclines à accepter un adverbe d'intensité et elles rejettent les compléments (*ibid.*). Cette seconde caractéristique invite à penser que les emplois nominaux sont davantage dérivés des emplois locutifs que des emplois de type adjectival. Les constructions du type *le chetif* signifieraient donc majoritairement "celui dont on peut s'exclamer « *Chetif!* »", plutôt que "celui qui est *chetif*". Un retour (intuitif) aux attestations montre en outre que le sens des occurrences qui servent de têtes nominales est plus proche du sens des locutifs que de celui des attributs du sujet et des modifieurs.

(8) Que fet quant ne se tue cil *las* qui joie s'est tolue ? (*YvainKu*, p.92f, v.3524)
 "Que fait le malheureux dont la joie a été arrachée s'il ne se tue pas ?"

Comme on le lit dans (8), le sens de *las* est davantage psychologique (v. not. Oppermann-Marsaux 2013 : 85) que dans (3), où il signifie littéralement "fatigué physiquement". Cela le rapproche très clairement de l'attestation du locutif de (1).

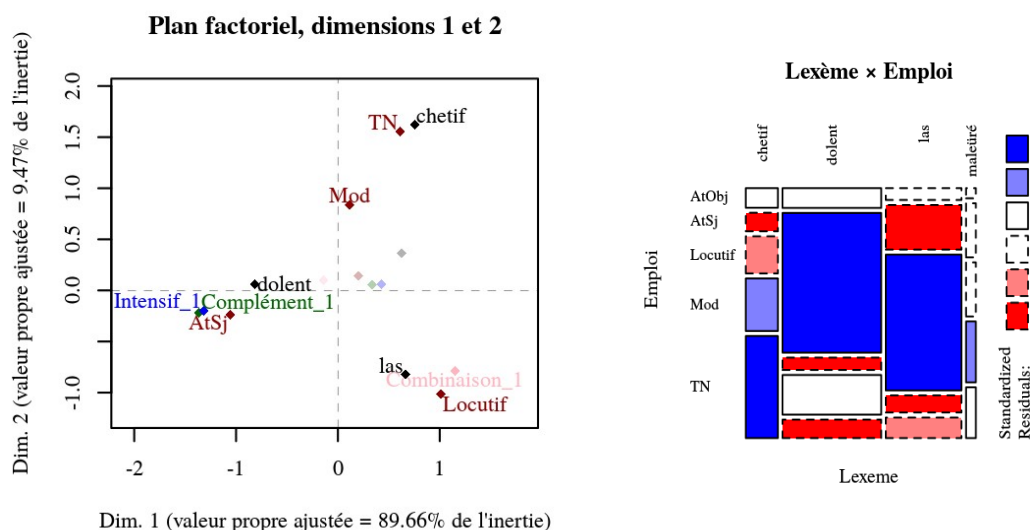


Figure 1: Gauche) ACM sur tous les individus - variables linguistiques
 Droite) Graphique mosaïque de l'interaction Lexème × Emploi

Le second axe oppose principalement les *Emplois* : la modalité *Locutif* se distingue de la modalité *TN* et, dans une moindre mesure, la modalité *Mod*, apparemment en vertu de leur capacité à se combiner à d'autres occurrences des lexèmes étudiés. Les locutifs expressifs sont en effet régulièrement cumulés, comme dans (7). La combinaison des deux axes met

également en évidence une spécialisation des lexèmes dans certains emplois, ce qui est clairement confirmé par l'analyse des résiduels mise en évidence dans la fig. 1, à droite.

Ainsi : 1/ *chetif* est principalement employé comme adjectif « substantivé » (*TN*) ou comme modifieur et rejette l'emploi comme attribut du sujet et comme locutif ; 2/ *dolent* est davantage employé comme attribut du sujet et ses emplois locutifs et substantivés sont rares ; 3/ *las* est plus employé comme locutif et rejette l'emploi épithétique et l'emploi comme attribut du sujet ou comme adjectif substantivé ; 4/ *maleüre* est plutôt employé comme modifieur.

À ce stade, on peut déjà dire que les lexèmes observés ont bel et bien des emplois locutifs qui se distinguent des autres emplois par des caractéristiques syntaxiques. D'autre part que ces lexèmes ont une spécialisation syntaxique spécifique – en particulier, *las* est effectivement spécialisé dans l'emploi locutif. Il nous paraît dès lors intéressant de continuer l'analyse en distinguant les emplois locutifs des autres.

3 Traitement différencié

L'analyse ayant mis en évidence une partition nette entre les emplois locutifs et les autres, nous traitons à présent uniquement les non-locutifs (sec. 3.1), puis uniquement les locutifs (sec. 3.2).

3.1 Emplois non locutifs

En appliquant le même traitement aux profils qui correspondent à des emplois non locutifs (1102 individus et 60 profils représentés sur 128 possibles), on obtient le plan factoriel de la fig. 2, qui représente 98,93 % de l'inertie ajustée (donc l'essentiel de l'information).

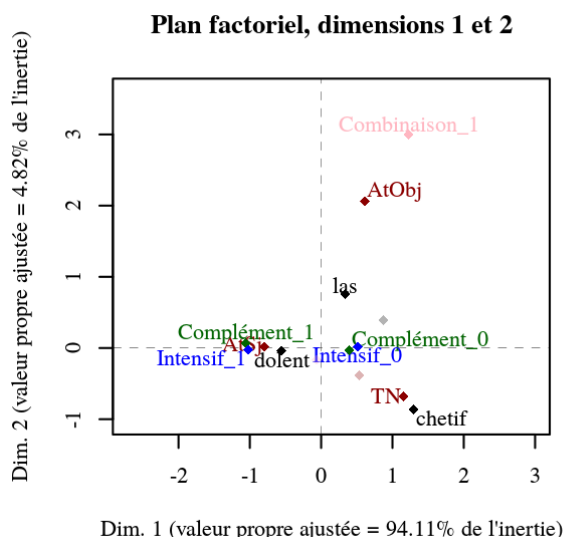


Figure 2: ACM sur les individus d'emploi non locutif - variables linguistiques

On retrouve les mêmes associations entre l'attribut et les compléments et adverbes intensifs, ainsi que l'emploi substantivé de *chetif*. L'information nouvelle que nous apporte le plan concerne les emplois *AtObj*, qui n'étaient pas bien représentés sur le plan contenant les locutifs. Les attributs de l'objet se caractérisent également par une tendance à la combinaison et le rejet des intensifieurs (comme c'était le cas des locutifs ci-dessus, cf. fig. 1, à gauche). Le retour aux données montre que cela est certainement dû à la construction stéréotypée *appeler/clamer X Y*, qui consiste en l'intégration délocutive des unités (Kahane et Mazziotta 2016 : §6.2.2).

(9) Plus de cent foiz s'apelle *lasse* l'abbaesse qui oit ce dire (*galeran*, p.91, v.2968)

“À ces mots, l'abbesse s'écrie plus de cent fois qu'elle est affligée”

(10) et se claime : « Las, chaitif, tant ai esté vilx [...] »/et se claime las, chaitif : « [...] »
(*qgraal_cm*, p.186b)

“Il dit qu'il est affligé : « J'ai été si vil ! »” (hésitation entre le discours direct et le discours indirect)

Comme on le comprend dans (9), la construction *appeler/clamer X Y* signifie “s'exclamer X à propos d'Y”. La frontière entre l'emploi locutif et l'attribut du sujet est si ténue que certaines attestations demeurent problématiques ; p.ex. : il est difficile de ponctuer (10) correctement. Il nous semble justifié de considérer que ces *AtObj* sont des formations délocutives plutôt que des transformations sur la base de la construction *AtSj* du type observé dans (4) – cp. forme reconstruite *et fait que chaque mere est dolente*.

3.2 *Emplois locutifs et concurrence de hélas*

Les conclusions qui précèdent indiquent que l'emploi locutif est associé en particulier à *las* même s'il est connu par tous les lexèmes considérés et spécifique à ce groupe. Cette observation est en accord avec les connaissances que nous avons de la constitution de la forme *hélas* comme unité purement locutive (Buridant, 2001 : 33 sv.). On dit traditionnellement que *las* a perdu les propriétés morphosyntaxiques qui le caractérisent (rappelées dans Opermann-Marsaux 2013 : 82-84) : il s'accorde de moins en moins avec son support (comme un adjectif) et finit par n'être employé que conjointement à *hé !* avec laquelle il fusionne en moyen français. Pareil processus va de pair avec une autonomisation « discursive » et est identifié comme une « pragmatocalisation »⁸ (Dostie 2004).

3.2.1 *Préparation des données*

Nous nous focalisons exclusivement sur les emplois locutifs des lexèmes qui nous intéressent. Cela mène à exclure les variables *Complément* et *Intensifieur* des analyses qui vont suivre – en vertu de la description théorique de la sec. 1 et des confirmations obtenues dans les résultats de la sec. 2.2. Nous ajoutons un relevé supplémentaire de 96 attestations d'*hélas* à l'échantillon. Nous n'avons pas relevé de formes permettant d'évaluer la stabilité de l'accord dans notre corpus. Par contre, nous pouvons évaluer l'aspect du figement en mobilisant la

⁸On distinguera ici la « pragmatocalisation » de la « grammaticalisation », qui est également associée à une perte

variable *Combinaison* (définie ci-dessus, sec. 2.1.2) et celui de l'emploi conjoint avec une « interjection » en définissant une nouvelle variable que nous nommons *Position*.

Position. Cette variable décrit la position du mot dans l'énoncé. La modalité 1 indique que l'occurrence se trouve en tête d'énoncé ou directement après une « interjection » (c'est-à-dire un mot qui n'a que des emplois locutifs), comme dans (11) et (12) ; 2 indique que d'autres locutifs de la série étudiée précèdent l'occurrence, comme c'est le cas pour *chetif* dans (13) ; 5 indique une position médiane « incidente », comme dans (14), où il est très clair que le locutif sépare le verbe *alons* de son sujet *nous* ; 9 indique une position finale, c'est le cas pour *dolent* dans (13). Soit quatre modalités.

(11) « *Chaitif*, faiseit il dunc, pur quei fui unches nez ? » (*becket*, p.122, v.3959)

“Quel malheur, faisait-il alors, pourquoi suis-je né ?”

(12) « Heu, *chaitifs*, tant mar vi ma mahenesce, dunt sui huni. » (*adgar*, p.150, v.45)

“Ah ! Quel malheur ! C'est pour mon malheur que je vis ma déchéance, qui me couvre d'opprobre.”

(13) Heu las, *chaitif* ! Que frai, *dolent* ? (*adgar*, p.178, v.475)

“Ah ! Quel malheur ! Que faire ?”

(14) Et nous, *chetif* ! n'alons mès riens querant. (*tdechamp*, p.212, v.21)

“Et nous, quel malheur ! ne cherchons jamais quoi que ce soit.”

3.2.2 Analyse

La combinaison des 96 individus dont la variable *Lexème* a comme modalité *helas* et des emplois locutifs des représentants des quatre autres lexèmes constitue un échantillon de 493 individus représentant 37/40 profils possibles ($5 \times 4 \times 2 = 40$; on voit dans le tableau de la tab. 4 que les 5 profils les plus fréquents concernent plus des trois quarts des individus). La projection des variables sur le plan factoriel est présentée sur la fig. 1 (la variable indiquée en bleu sera traitée ci-dessous). Par l'examen du premier axe, on peut facilement voir que les emplois locutifs de *las* sont initiaux ou suivent un locutif pur. Les autres lexèmes sont plutôt employés en combinaison, après d'autres locutifs. La combinaison avec second axe montre que ce qui oppose *las* à *helas* est que le second a l'air plus neutre par rapport à la position (*las* rejette très fortement *Position_2*).

Lexème-Position-Combinaison	Fq. abs.	Cum. abs.	Fq. %	Cum. %
las-1-0	229	229	46.45	46.45
helas-1-0	77	306	15.62	62.07
las-1-1	34	340	6.90	68.97
las-5-0	18	358	3.65	72.62
dolent-2-1	15	373	3.04	75.66

Table 4: Profils les plus fréquents Lexème-Position-Combinaison

Il faut toutefois remarquer : 1/ que les cas de *helas* en *Position_2* sont en fait des cas où le lexème est associé à une autre occurrence de lui-même ; 2/ qu'un certain nombre de cas (36) de *Position_1* pour *las* sont en réalité des chaînes du type *he ! las !* et leurs variantes, comme dans l'exemple suivant :

(15) Hé, las ! Que Galerons nel set ! (*galeron*, p.126, v.3271)
 “Malheureux ! Galeron ne le sait pas !”

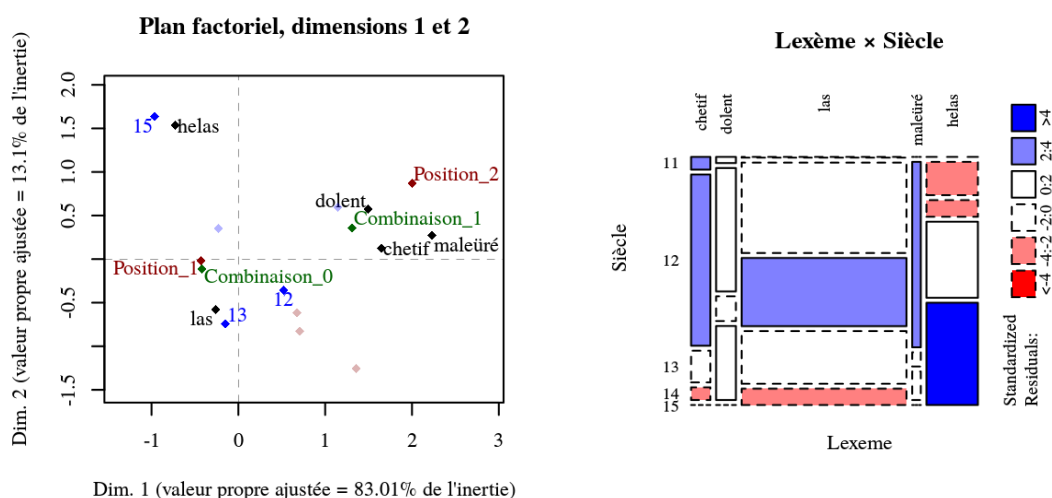


Figure 3: Gauche) ACM des locutifs ;
 Droite) Fréquences et graphique mosaïque de l'interaction Lexème × Siècle (locutifs)

Par ailleurs, les études antérieures posent la question de la chronologie et situent à tout le moins *las* et *helas* dans un continuum (le premier précédant le second). On peut assortir chacun des individus d'une variable dont les modalités décrivent le siècle où il est attesté. En projetant cette variable *Siècle* comme variable illustrative (en bleu sur la fig. 1 à gauche), on voit se dessiner une chronologie : *helas* est la forme spécifique aux textes plus récents. Le lexème a tendance à être davantage employé au 15^e, Tandis que *las* tend plutôt à être spécifique au 13^e. Les autres formes sont plus anciennes (seuls *chetif* et *maleüré* sont significatifs à cet égard ; cf. fig. 3, à droite). On peut donc interpréter cette progression chronologique comme une spécialisation du lexème *las* dans un premier temps, suivie de son figement en locutif conjointement à la fusion avec une « interjection » ; soit : *chetif/maleüré* → *las* → *helas*. Toutefois, *las* continue à être employé en tant qu'adjectif encore à l'heure actuelle. La « pragmatization » concerne ainsi davantage les emplois d'une construction particulière qu'une spécialisation complète d'un lexème dans un rôle pragmatique.

4 Conclusions

Notre étude mène à trois conclusions. Premièrement, les lexèmes *chetif*, *dolent*, *las* et *maleüré* ont beau avoir le même potentiel syntaxique, ils ont tendance à être employés dans l'un ou l'autre des différents emplois. Cela a des conséquences directes sur la manière dont leurs compléments et intensifieurs se combinent avec eux, conformément à ce que l'analyse théorique prédit (sec. 2.2). On voit donc que la classification théorique et l'analyse des données se valident mutuellement. L'encodage analytique de caractères définitoires des classes d'emploi permet d'évaluer la pertinence de ces critères (observables au niveau des tendances représentées sur les plans). Deuxièmement, parmi les emplois non locutifs de ces lexèmes, certains peuvent sans doute être considérés comme *délocutifs*, au sens où leur potentiel syntaxique autant que leur sens les rapprochent des locutifs. C'est le cas des emplois

substantivés de *chetif* (sec. 2.2) et des constructions du type *appeler/clamer X Y* (sec. 3.1). Les traitements statistiques attirent l'attention sur certaines questions peu saillantes pour le lecteur humain. Enfin, l'analyse a mis en évidence le développement du lexème *helas* à la fin de la période médiévale pour remplacer les autres lexèmes au départ du lexème *las*, dont il est issu (sec. 3.2). La visée illocutoire (donc pragmatique) du lexème était déjà figée dans les emplois locutifs au début de la période médiévale. L'approche traditionnelle est ici validée et enrichie par une représentation formalisée de la variation.

5 Références

- Benzécri, J.-P. (1979). Sur le calcul des taux d'inertie dans l'analyse d'un questionnaire. *Cahiers d'analyse des données*, 3 : 377-378.
- Benzitoun C., Dister, A., Gerdes K., Kahane, S., Pietrandrea, P. et Sabio, F. (2010). *Tu veux couper là faut dire pourquoi*. Propositions pour une segmentation syntaxique du français parlé. In Neveu F., Muni Toke, V., Durand, J., Klingler, T., Mondada, L. et Prévost, S. (éds.). *Congrès Mondial de Linguistique Française - CMLF 2010*, pages 2075-2090.
- BFM2014 = *Base de Français Médiéval* [En ligne]. ENS de Lyon, Laboratoire ICAR <<http://bfm.ens-lyon.fr>>.
- Blanche-Benveniste C. (1997). *Approche de la langue parlée en français*. Ophrys.
- Buridant C. (2001). *L'interjection en français. Esquisse d'une étude diachronique*. Université Marc Bloch.
- Dostie G. (2004). *Pragmaticalisation et marqueurs discursifs. Analyse sémantique et traitement lexicographique*. De Boeck.
- Guillot C., Lavrentiev A. et Marchello-Nizia Ch. (2007). La Base de Français Médiéval (BFM) : états et perspectives. In Kunstmann P. et Stein A. (éds). *Le Nouveau Corpus d'Amsterdam. Actes de l'atelier de Lauterbad, 23-26 février 2006*. Franz Steiner Verlag, pages 143-152.
- Kahane, S. et Mazziotta, N. (2016, sous presse). Quel classement syntaxique pour les « marqueurs discursifs », « mots-phrases » et autres « inserts » ? Prédicatifs et locutifs. *Travaux de linguistique*.
- Le Roux, B. et Rouanet, H. 2010. *Multiple correspondence analysis*. Sage.
- Lê, S., Josse, J. et Husson, F. (2008). FactoMineR : An R Package for Multivariate Analysis. *Journal of Statistical Software* 25(1) : 1-18.
- Lebart, L., Morineau, A. et Piron, M. (1998). *Statistique exploratoire multidimensionnelle*. Dunod.
- Mazziotta N. (soumis). *Microsyntaxe et macrosyntaxe des lexèmes chaitif, dolent, las et maleuré/maleureus en français médiéval . De l'adjectif au locutif*. Soumis au 5^e congrès international de linguistique française.
- Opermann-Marsaux É. (2013). « Ha ! las, fet-il, com sui iriez ! » : les emplois de « (hé +) las(se) » et l'expression de la douleur morale en français médiéval. In Lagorgette D. et Larrivée P. *Représentation du sens linguistique 5*. Université de Savoie, pages 81-95.
- Prévost S. (2011a). *Français médiéval en diachronie: du corpus à la langue* (mémoire de syntèse HDR). ENS Lyon.
- Prévost S. (2011b). *Expression et position du sujet pronominal du 12^e au 14^e siècle: une approche quantitative* (recherche inédite HDR). ENS Lyon.
- Searle J. R., Vanderveken D. (1985). *Foundations of illocutionary logic*, Cambridge University Press.