

Expression et position du sujet en ancien français : le rôle de la personne pronominale

Mathieu Goux^{1,*}, et Pierre Larrivée¹

¹CRISCO (EA 4255), Université de Caen, Esplanade de la Paix, CS14032, 14032 Caen, France

Résumé. Contrairement aux autres langues romanes, le français a adopté l'expression obligatoire du sujet grammatical suivant un long parcours dont les causes restent à pleinement élucider. C'est la dimension de la personne grammaticale du pronom que reconsidère cet article. Utilisant un corpus calibré de données du même type coutumier et de la même région, il se pose la question de savoir si les pronoms personnels de première et deuxième personnes anticipent, par rapport à ceux de troisième, l'expression obligatoire et l'antéposition au verbe qui caractérisent le français actuel. L'étude quantitative montre que l'expression et que l'antéposition de ces pronoms de première et deuxième personne sont en effet plus importantes que pour les pronoms de troisième personne et que le premier ensemble est particulièrement associé à des verbes de dire. Les résultats établissent que les textes légaux donnent une chronologie antérieure du changement linguistique, surtout quand ils sont assemblés pour exclure des variations de type de texte et de région.

Abstract. The expression and position of subject in Medieval French: The role of grammatical person. Unlike other Romance languages, French is a non-pro-drop language, and has become so for reasons that remain to be fully understood. The potential correlate of the grammatical person of subject clitics is revisited in this paper. By using a calibrated corpus of data from the same customal text type and region, it raises the question of whether first and second person subject clitics undergo before third person ones compulsory expression in preverbal position. The quantitative study shows that indeed the rate of expression and preverbal position is higher for first and second person subject clitics as compared to third person ones; what's more, the former is significantly related to verbs of saying. The study shows that legal texts provide an earlier chronology of change, and that calibration by type and region helps exclude non-structural factors of variation.

1 Introduction

Le passage du latin aux langues romanes a vu de nombreux changements grammaticaux. Comme ce sont généralement les structures qui changent qui retiennent l'attention, et que

* Corresponding author : mathieu.goux@unicaen.fr

ces évolutions sont parmi l'ensemble roman les plus importantes en français, cette dernière langue a été l'objet de nombreux travaux. Une des évolutions par lesquelles le français se démarque de l'italien ou de l'espagnol est l'expression obligatoire du sujet, propriété acquise dans la période textuellement documentée du français médiéval. L'expression du sujet obligatoire a été associée à de nombreux facteurs comme la richesse de la morphologie verbale par exemple. Une autre dimension reliée à la morphologie verbale est la personne du sujet pronominal (Detges 2003, Prévost 2015) : les pronoms de première et deuxième personne adopteraient le patron d'expression obligatoire en position antéposée au verbe plus tôt que les pronoms de troisième personne. C'est cette idée que reconsidère cet article à partir d'un nouveau corpus monogénérique de textes légaux de la même région, s'étendant sur six siècles. Le taux d'expression et la position des pronoms personnels sujets sont vérifiés pour établir si, en effet, la personne du pronom influence son comportement.

Cet article est organisé en trois parties. Dans la première, nous faisons un panorama des hypothèses proposées pour expliquer la stabilisation de l'expression et de l'antéposition du sujet pronominal en français, et l'anticipation de ce modèle par les pronoms de première et de seconde personnes. Dans la seconde partie, nous présentons le corpus que nous avons retenu pour cette étude, nos choix d'analyse et les résultats obtenus. Enfin, dans notre troisième partie, nous évaluons les résultats à l'aune des hypothèses présentées préalablement.

2 État de la question

Un changement central dans le passage du latin au français est celui du sujet nul au sujet obligatoirement exprimé. Considérons une phrase du *Serment de Strasbourg* (842)¹ :

(1) et ab Ludher nul plaid nunquam *prindrai*

La traduction en latin classique (fournie par Leclerc, 2019) n'exige pas l'expression du sujet :

(2) et cum Clotario nullam unquam pactionem *faclam*

alors que celle-ci est obligatoire dans la traduction en français contemporain, ce qui ne serait pas forcément le cas dans les autres langues romanes, tel l'italien (Lobo & Martins, 2017).

(3a) et **je** ne *prendrai* jamais aucun arrangement avec Lothaire

(3b) e con Lotario non *prenderò* mai alcun accordo

Les raisons de l'existence de langues à sujet nul comme le latin et l'ancien français ont été discutées par une littérature considérable. L'hypothèse généralement défendue est que les langues à sujet nul peuvent se dispenser de l'expression d'un pronom parce que leur morphologie verbale permet l'identification de la personne². C'est ce que montre la conjugaison au présent de l'indicatif du verbe *mangiare* en italien face à celle du verbe *manger* en français.

(4) P1 mangio	(5) P1 mange
P2 mangi	P2 manges
P3 mangia	P3 mange
P4 mangiamo	P4 mangeons
P5 mangiate	P5 mangez
P6 mangiano	P6 mangent

Alors que la conjugaison du verbe italien permet de reconnaître toutes les personnes, celle du français ne permet d'en distinguer que trois à l'oral (la P4 et la P5 se différencient de toutes les autres personnes, qui sont parfaitement homophones hors phénomènes de liaison). De ce point de vue, le français aurait ainsi perdu son statut de langue à sujet nul à cause de la perte de distinction de conjugaison. Cependant, les travaux testant cette hypothèse sont loin d'être concluants (Simonenko, Crabbé & Prévost, 2019) ; de plus, l'existence de langues partiellement à sujet nul (à l'instar des langues germaniques, du breton et du cachemiri, analysées par Holmberg, 2010) distend le lien entre sujet nul et richesse de la conjugaison verbale.

Les études du comportement du sujet nul en français médiéval révèlent d'autres déterminismes potentiels ; parmi eux figure la personne verbale du pronom. Dans son étude d'un vaste corpus essentiellement littéraire courant de 1100 à 1400³, Prévost (2015 : 36-38) montre que les pronoms de troisième personne restent plus fréquemment non-exprimés que ceux de première personne (tableau 1).

Tableau 1. Fréquences absolues des sujets pronominaux (Sp) préverbaux, postverbaux et non exprimés dans Prévost (2015)

	Sp préverbal	Sp postverbal	Sp non exprimé	Total
P1 (je)	1 709	252	1 488	3 449
P3 (il, elle, ils, elles)	2 300	401	9 691	12 392
Total	4 009	653	11 179	15 841

Selon ces résultats, 43% des premières personnes ne sont pas exprimés (1488 / 3449), contre 78% des troisièmes personnes (9691 / 12392). Concernant la position des pronoms réalisés, les valeurs changent selon qu'on comptabilise les sujets nuls, ou non. Avec les sujets nuls, le taux d'antéposition est différentié : presque 50% des premières personnes sont préverbaux (1709 / 3449), contre 19% des troisièmes personnes (2300 / 12392). Sans les sujets nuls, le taux est semblable : 87% des premières personnes (1709 / 1961) et 85% des troisièmes personnes exprimées (2300 / 2701) sont préverbaux⁴. Autrement dit, il semblerait que la première personne anticipe les changements syntaxiques que subit l'ancien français, et serve de patron à l'expression du sujet pronominal. Les deux graphiques suivants résument l'évolution et du taux d'expression, et de la position préverbale des pronoms de première et de troisième personne. Nous avons construit le premier graphique (fig.1) en reprenant le tableau 2 de Prévost (2015 : 39), le second (fig.2) reprend le tableau 3 (*ibid.*). Ces tableaux organisent ses résultats chronologiquement.

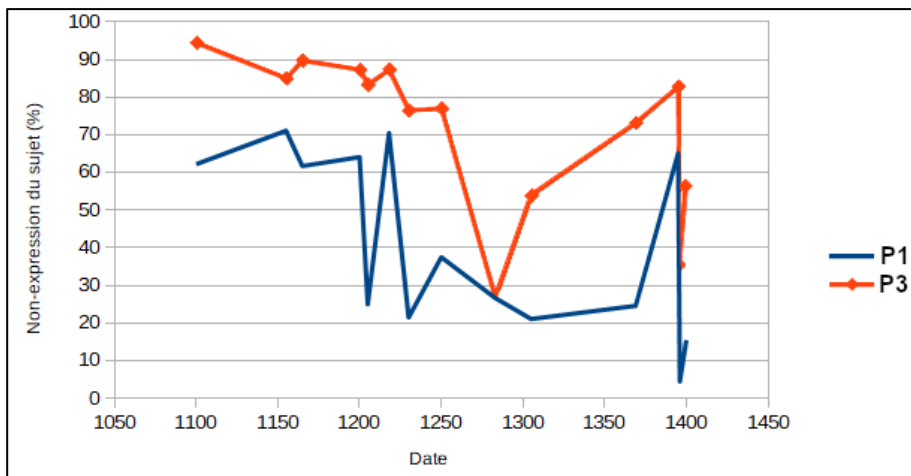


Fig. 1. Fréquence (en %) de la non-expression de la P1 et de la P3 sujet dans Prévost (2015)

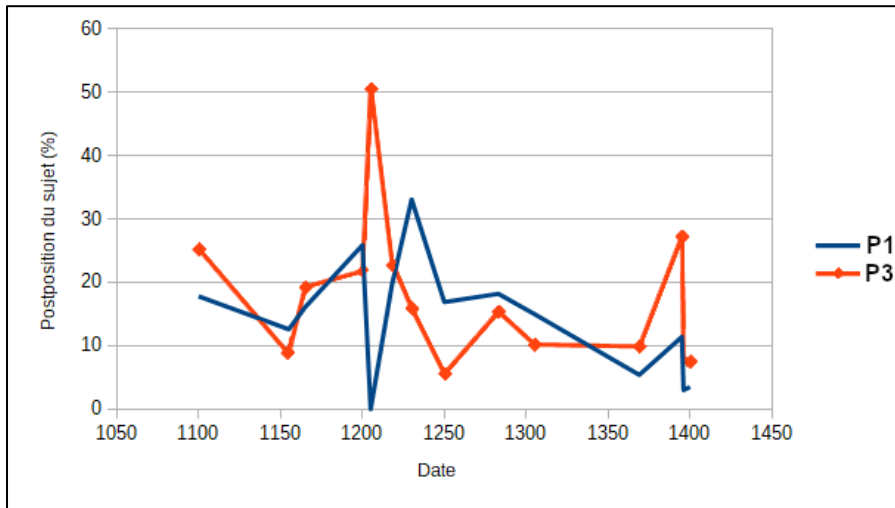


Fig. 2. Fréquence (en %) de postposition de la P1 et de la P3 sujet dans Prévost (2015)

Bien que les courbes soient irrégulières, le genre textuel influençant notamment sur les données⁵ et le peu d'occurrences relevées rendant souvent délicat les calculs de fréquence, nous voyons que, au fur et à mesure du temps, l'expression comme l'antéposition du sujet pronominal de première personne devançant celles du sujet pronominal de troisième personne dans la dynamique du changement. Ce phénomène est illustré par des données diachroniques d'autres langues : Axel (2005) considère ainsi les données du vieux-haut allemand, et synthétise ses résultats dans un tableau, que nous reproduisons ci-dessous (Tableau 2. Dans ce tableau, le pourcentage représente la fréquence de sujets nuls à l'aune de tous les sujets du corpus).

Tableau 2. Fréquence (en %) d'expression (*overt*) et non-expression (*null*) du sujet pronominal en vieux-haut allemand dans les *Monsee Fragment* (MF), *Isidor* (I) et *Tatian* (T)⁶

		<i>MF</i>	<i>I</i>	<i>T</i>
pers.	nomb.	pron. suj. (expr./nul)	pron. suj. (expr./nul)	pron. suj. (expr./nul)
SG	1	10/5 (33%)	36/2 (5%)	415/103 (20%)
	2	5/3 (38%)	3/2 (40%)	131/84 (39%)
	3	12/52 (81%)	15/29 (66%)	394/460 (54%)
PL	1	2/1 (33%)	2/3 (60%)	62/27 (30%)
	2	16/10 (38%)	1/0 (0%)	262/42 (14%)
	3	3/13 (81%)	4/12 (75%)	170/244 (59%)
Total		48/84 (= 64%)	61/48 (= 44%)	1434/960 (= 40%)

Ces résultats vont dans le même sens que ceux de Prévost (2015) pour la langue française : les premières et deuxièmes personnes ont des taux d'expression plus importants que les troisièmes. Si on agrège les chiffres, on obtient une moyenne de 19,2% de non-expression pour les premières personnes du singulier par exemple, face à 43,7% pour la troisième du singulier. Ces données écartent en outre l'explication d'une disparition du sujet nul par une perte de la morphologie verbale, ou un changement dans l'ordre de mots, puisque ni l'un ni l'autre ne se manifeste dans cette langue. De même, dans les textes qui l'attestent, le sujet nul est mieux représenté avec la troisième personne. Nous reproduisons les résultats de Walkden (2013), dans son quatrième tableau, ci-dessous (Tableau 3).

Tableau 3. Fréquence d'expression (en %) des sujets pronominaux de verbes à l'indicatif en vieil anglais

Texte	Personne	Nombre	Sujet nul	Total
<i>Beowulf</i>	1	sg.	2 (3%)	77
		pl.	0 (0%)	21
	2	sg.	1 (4%)	27
		pl.	0 (0%)	10
	3	sg.	42 (20%)	214
		pl.	20 (29%)	69
	<i>Total</i>			65
<i>Bald's Leechbook</i>	1	sg.	0 (0%)	1
		pl.	0 (0%)	11
	2	sg.	0 (0%)	52
		pl.	0	0
	3	sg.	32 (33%)	140
		pl.	14 (29%)	49
	<i>Total</i>			46
<i>Bede</i>	1	sg.	0	129
		pl.	2 (1%)	173
	2	sg.	0	69
		pl.	1 (4%)	26
	3	sg.	44 (3%)	1548
		pl.	29 (11%)	265
	<i>Total</i>			76
<i>Chronicle MS. E</i>	1	sg.	0 (0%)	3
		pl.	0 (0%)	18
	2	sg.	0 (0%)	3
		pl.	0 (0%)	3
	3	sg.	9 (3%)	306
		pl.	17 (6%)	292
	<i>Total</i>			26

Ces résultats montrent que, alors que les premières et deuxièmes personnes exhibent au mieux des taux d'omission de 4%, ce taux va jusqu'à 29% pour les troisièmes personnes. Cependant, comme le reconnaît l'auteur (*ibid.*, p. 66), la faible fréquence des pronoms autres que de troisième personne peut biaiser les résultats. Les similitudes, néanmoins, entre les phénomènes observés tant en français qu'en anglais ou en allemand interrogent : pourquoi y aurait-il une différence dans le taux d'expression des pronoms personnels des première et deuxième personnes dans une situation de changement allant vers l'expression obligatoire du sujet ?

Une explication est proposée dans le cadre de la théorie de la grammaticalisation (Hopper & Traugott, 2003). Cette théorie lie les changements syntaxiques à la valeur pragmatique et à la fréquence des formes grammaticales ; partant, une forme qui deviendra l'expression par défaut d'une catégorie émergerait initialement comme option marquée. Par exemple, on suppose que, avant de devenir la façon ordinaire d'exprimer la négation, la structure à double détente *ne ... pas* exprimait une valeur emphatique à l'époque ancienne où le *ne* suffisait à exprimer la négation verbale⁷. Cette valeur emphatique se serait perdue à mesure que l'emploi de la structure devenait plus fréquent, la qualifiant à devenir le marqueur ordinaire de la négation. C'est ce que Van der Auwera (2009) appelle la « spirale de Meillet », en référence à la notion de grammaticalisation, popularisée par ce dernier.

Cette notion est exploitée par Detges (2003) pour étudier l'émergence de l'expression obligatoire du sujet. Il estime que les pronoms personnels de première et deuxième personnes sont les plus susceptibles d'être utilisés pour des raisons de mise en relief : l'usage de ces pronoms dans une langue à sujet nul s'accompagne d'une valeur d'emphase. En effet, il est plus probable de vouloir mettre l'emphase sur les représentants des participants de l'interlocution que sur les personnes absentes de la situation d'énonciation. Or, l'accroissement de l'expression des pronoms de première et de seconde personnes s'accompagnerait d'une perte de leur valeur d'emphase, ce qui inviterait alors à généraliser et à accélérer l'expression obligatoire du sujet à toutes les personnes verbales.

Cette explication est évaluée par Larrivée et Capin (2018), qui notent la disparité des types et de l'origine géographique des textes utilisés par Prévost (2015) : elle a effectivement dépouillé des textes appartenant à des genres distincts, littéraires, historiques, religieux, didactiques et juridiques, et présentant des traits issus de différents dialectes (anglo-normand, picard...). Ces auteurs ont donc cherché à vérifier l'asymétrie entre la troisième personne et les autres, au sein d'un corpus plus homogène. Ils se fondent pour ce faire sur un texte coutumier parisien des années 1270, les *Établissements des Capétiens*, qu'ils opposent à une série de textes courts, 23 chartes de la période 1270-1280 du *Cartulaire d'Île-de-France*, dont les transcriptions diplomatiques sont produites et mises en ligne par l'École Nationale des Chartes⁸. Du premier sont annotées manuellement 317 propositions, et dans les chartes toutes les propositions le sont (361 propositions). Les résultats sont fournis par leur troisième tableau (Larrivée & Capin, 2018 : 7), que nous reproduisons ci-dessous (Tableau 4. Les pourcentages représentent la proportion de sujet non-exprimé en comparaison des sujets exprimés).

Tableau 4. Taux de sujets non-exprimés pour les premières et deuxième personnes dans Larrivée & Capin (2018)

Textes	1PS	1PP	2PS	2PP	Total
<i>Capétiens</i>	0/0	0/17	0/11	0/11	0/39
<i>Cartulaire</i>	3/39 (7,69%)	6/59 (10,17%)	0/0	0/0	9/98 (9,18%)

Ces chiffres montrent ainsi que dans le *Cartulaire*, il y a une différence notable de réalisation du pronom sujet entre la première personne du singulier (7,69% de sujets sont non-exprimés) et la première personne du pluriel (10,17% de non-expression).

Les auteurs ont ensuite procédé à une vérification du taux de réalisation des sujets et montrent que dans les *Capétiens*, ce taux pour les première et deuxième personnes est plus élevé (100%) que le taux général pour la troisième personne (qui est de 92,11%). En revanche, dans les textes choisis du *Cartulaire*, le taux de réalisation des première et deuxième personnes est légèrement plus faible, s'élevant à 93,45 %, alors que, pour la troisième personne, il se trouve à 94,29 %. Une fois encore cependant, le nombre d'occurrences est faible, ce qui appelle une évaluation fondée sur un ensemble plus important des données. C'est ce que nous avons effectué dans le présent article, selon le protocole que nous présentons ci-après.

3 Corpus, méthode et résultats

La question que pose ce travail est de savoir si les pronoms de première et deuxième personnes sont affectés plus tôt par le changement de l'ancienne langue vers l'expression et l'antéposition du sujet. Une difficulté récurrente de ce type d'études, néanmoins, est de réunir un nombre suffisant d'occurrences : on le voit, même dans un important corpus d'un demi-million de mots comme celui utilisé par Prévost (2015), il est difficile de réunir 250 occurrences de *je* exprimé. Il est donc important de pouvoir s'appuyer sur des ressources

abondantes ; dans le même temps, il convient d'assurer la comparabilité de ses ressources en termes de type et d'origine géographique. Balon & Larrivée (2016) ont ainsi montré que les textes légaux anticipaient le changement linguistique au regard des textes littéraires en raison de leur absence de préoccupation esthétique, par exemple concernant les répétitions⁹ ; c'est également ce que montre la comparaison des chiffres de Prévost (2015) et de Capin & Larrivée (2018).

De plus, parce que ces textes légaux existent de longue date et dans la plupart des régions, leur usage aide à exclure les paramètres non-structuraux de variation, tels le genre ou les traits dialectaux, et à suivre la courbe d'évolution sur le temps long. C'est ce que permet de faire le corpus développé dans le cadre du projet RIN ConDE¹⁰, financé par la région Normandie et consacré aux coutumiers normands. Le corpus retenu pour ce projet compte une quinzaine de témoins, s'échelonnant du *Très Ancien Coutumier*, daté de la moitié du XIII^e siècle, à la *Coutume expliquée* de Pesnelle (1771). Ce corpus est unique (i) par l'origine géographique homogène des textes, qui permet de gommer ce biais dans l'analyse diachronique ; (ii) par son homogénéité générique : ces textes « de spécialité » se consacrent au même sujet, le droit coutumier normand, ce qui assure une continuité thématique limitant les potentiels écarts stylistiques entre les textes.

Pour notre étude, nous nous sommes focalisés sur deux témoins en particulier : le *Très Ancien Coutumier* (à présent *TAC*), dont le manuscrit est conservé à la Bibliothèque Sainte-Geneviève sous la cote [BSG MS1743¹¹], que nous avons photographié et qui a fait l'objet d'une transcription par Marnier (1839), et le *Grand Coutumier* (à présent *GC*) conservé à la Harvard School Library, sous la cote [HLS MS91], dont les vues sont disponibles en ligne¹², et dont la transcription a été faite avec le logiciel Transkribus¹³. Nous avons par la suite travaillé et sur ces transcriptions, et sur les manuscrits eux-mêmes. Les transcriptions ont ensuite été lemmatisées et étiquetées morpho-syntaxiquement grâce au logiciel d'annotation collaboratif ANALOG (Lay & Pincemin, 2010). La sortie XML-TEI construite grâce au logiciel nous a alors permis d'opérer des comptages des pronoms personnels sujets exprimés, ainsi que d'identifier les verbes les plus fréquemment employés en association de ces derniers. Cela nous a également permis de calculer des taux d'antéposition et de postposition des sujets pronominaux, en isolant les séquences du type « pronom-verbe », ou « verbe-pronom » dans le corpus. Des vérifications manuelles, en revenant aux manuscrits originaux, ont permis de corriger les éventuelles erreurs d'étiquetage, quand bien même le grand nombre d'occurrences pour les troisièmes personnes, singulier et pluriel, auraient empêché une vérification systématique des formes. Nous jugeons cependant la marge d'erreur acceptable et ne compromettant pas la représentativité des ordres de grandeur offerts par la présente recherche sur ces textes.

Cette exploration semi-automatisée du corpus a été complétée par une exploration manuelle de ces deux témoins, en annotant les 410 premières propositions du TAC, et les 302 premières du GC. En plus de contrôler les résultats obtenus lors de la phase d'annotation automatique du corpus, cette opération d'annotation manuelle nous a permis de repérer plus aisément les verbes sans sujet exprimé, ce qui est difficile à faire lors d'un étiquetage automatisé du corpus, du moins, à ce stade de la constitution de la base de données ConDE¹⁴. Nous avons enfin complété ce corpus premier par l'étude des *Instrucions et enseignements*, texte juridique normand daté de 1386 et édité par Besnier & Génestal (1912). Le texte, que nous désignerons sous le terme de « Styles », a été annoté manuellement sur un segment continu de 455 propositions (sur approximativement 840 propositions).

Trois variables nous ont particulièrement intéressés ici : le taux d'expression du sujet pronominal, le taux d'antéposition ou de postposition du sujet pronominal et, enfin, les verbes les plus fréquemment employés avec les pronoms exprimés. Nous faisons effectivement l'hypothèse que si la théorie de l'emphase est fondée, nous devrions trouver

des indices de celle-ci dans le type de verbe qu'accompagnent les pronoms de première et de deuxième personne.

3.1 Expression et position des pronoms personnels sujets

En nous fondant sur les annotations manuelles opérées sur les textes, voici les résultats que nous avons obtenus (Tableau 5a). Dans les cases de ce tableau, le premier chiffre renvoie au nombre d'occurrences de sujets exprimés par personne, le second le nombre de sujets non-exprimés. Le chiffre entre parenthèses renvoie au pourcentage que représentent les sujets exprimés pour une personne donnée.

Tableau 5a. Fréquence (en %) des sujets pronominaux exprimés/non-exprimés dans notre corpus

	P1	P2	P3	P4	P5	P6	Total
1250 – TAC (410 prop.)	0/0	1/0 (100%)	79/6 (93%)	2/0 (100%)	0/0	20/1 (95,2%)	102/7 (93,6%)
1300 – GC (302 prop.)	4/0 (100%)	0/0	45/4 (91,8%)	16/1 (94,1%)	0/0	14/1 (93,3%)	79/6 (93%)
1386 – Styles (455 prop.)	37/1 (97,4%)	21/2 (91,3%)	51/23 (69%)	2/0 (100%)	15/11 (57,7%)	24/0 (100%)	150/37 (80%)
Total	41/1 (97,6%)	22/2 (91,7%)	175/33 (84,1%)	20/1 (95,2%)	15/11 (57,7%)	58/2 (96,7%)	

Les relevés vont dans le sens des remarques précédentes : quelque texte considéré, le taux de réalisation du sujet pronominal est plus prononcé dans le cadre des première et deuxième personne, au détriment notamment de la troisième personne du singulier dont le taux de réalisation n'atteint ainsi que 69% dans *Styles* (51/74) et ce bien que le faible nombre d'occurrences relevées, une fois encore, puisse créer des biais. Nos relevés concernant la position du pronom personnel sujet, qui portent sur l'intégralité des textes du *TAC* et du *GC*, vont également dans la même direction. Nous comparons ces chiffres ici avec les relevés obtenus concernant *Styles*, bien qu'ils ne portent pas sur l'intégralité du texte (Tableau 5b). Le premier chiffre indique le nombre d'occurrences de sujets antéposés, le second de sujets postposés. Le pourcentage représente la proportion de sujets antéposés à l'aune de l'ensemble des sujets exprimés.

Tableau 5b. Fréquence (en %) de l'antéposition/postposition des sujets pronominaux exprimés

	P1	P2	P3	P4	P5	P6	Total
1250 – TAC	16/0 (100%)	5/0 (100 %)	594/80 (88,1%)	25/6 (80,6%)	7/0 (100%)	19/2 (90,5%)	666/88 (88,3%)
1300 – GC	63/0 (100%)	18/0 (100 %)	889/183 (82,9%)	32/0 (100%)	0/0	189/20 (90,4%)	1191/203 (85,4%)
1386 – Styles (455 prop.)	37/0 (100%)	20/1 (95,2%)	50/1 (98%)	2/0 (100%)	15/0 (100%)	24/0 (100%)	148/2 (98,7%)
Total	116/0 (100%)	43/1 (97,7%)	1533/264 (85,3%)	59/6 (90,8%)	22/0 (100%)	232/22 (91,3%)	

Ces résultats indiquent que la troisième personne du singulier semble être plus lente à adopter le modèle que suivra la langue française passée la période médiévale dans ce corpus, bien que l'écart demeure marginal : les coutumiers dépouillés pour cette étude ne contredisent donc pas la tendance observée et discutée précédemment, quand bien même la représentativité du corpus, une fois encore, demeurerait un point faible de ce type d'analyse. Reste néanmoins, et conformément à ce qui a été analysé notamment par Larrivée & Capin (2018), que ces textes anticipent la chronologie du changement observée

dans les textes littéraires, tendance qu'illustrent les graphiques précédents (fig. 1 et fig. 2). Si nous mettons en vis-à-vis nos chiffres avec ceux de Prévost (2015), l'anticipation se fait plus nette (Tableau 6. Les cases grisées correspondent à nos résultats). Par exemple, la non-expression du sujet P3 dans *Styles* est inférieure à celle relevée dans *Quinze Joyes de Mariage* (1400), et sa postposition dans ce dernier texte a une fréquence avoisinant celle que nous avons relevée dans le *Grand Coutumier*, écrit pourtant un siècle auparavant.

Tableau 6. Comparaison des fréquences (en %) de non-expression et de postposition des sujets dans notre corpus et dans celui de Prévost (2015)

	Non-expression du sujet (%)		Postposition du sujet (%)	
	P1	P3	P1	P3
1250 (TAC)	0	7	0	11,9
1250 (Prévost)	37,5	77	16,9	5,6
1300 (GC)	0	8,2	0	7,1
1305 (Prévost)	21,1	54,1	15	10,2
1369 (Prévost)	24,6	72,1	5,4	9,9
1386 (Styles)	2,6	31	0	2
1400 (Prévost)	15,5	57,3	3,5	7,3

3.2 Énonciation et emphase

Il faut à présent explorer les raisons pour lesquelles les pronoms de première et de deuxième personne tendent à être davantage exprimés au regard des pronoms de troisième personnes, et notamment l'hypothèse selon laquelle cela serait lié au paramètre de l'emphase. Pour vérifier cette hypothèse, nous avons fait une analyse des verbes les plus fréquemment employés avec les pronoms de première et de deuxième personne. Le repérage des occurrences a été facilité par l'étiquetage automatique des textes, et nous avons ensuite vérifié manuellement nos résultats. En tout, nous avons relevé 45 verbes distincts avec ces pronoms¹⁵. Si nous trouvons, de façon attendue, des copules ainsi que des verbes modaux ou semi-modaux, tel *faire*, *être* ou *vouloir*, ainsi que quelques verbes relevant du lexique spécialisé du droit tel *succéder* ou *vendre*, les pronoms de première et de deuxième personne, tant singulier que pluriel, ont dans ces textes une accointance particulière avec l'expression du *dire* ou, plus largement, avec l'exercice de la parole, au point parfois de quasi monopoliser l'expression d'un pronom personnel spécifique (exemples 6 et 7).

(6) se il dit einssi : **ge** *ni* tot par les paroles (il dira ainsi : je *nie* tout par ces paroles. *TAC*, p. 33 / Marnier, p. 46)

(7) En la vente du cheval dont **nos** *avon* *parle* [...] (*GC*, p. 47)

Nous avons résumé nos résultats dans les deux tableaux suivants (Tableaux 7a & 7b ; dans ceux-ci, la dernière ligne exprime le pourcentage total que ces verbes représentent à l'aune du corpus). Notons également que, avec ces verbes, les sujets pronominaux de première et de deuxième personne sont toujours antéposés.

Tableau 7a. Verbe relatif au *dire* dans le TAC

	P1	P2	P4	P5
<i>Demander</i>	/	/	1	/
<i>Dénier</i>	1	/	/	/
<i>Dire</i>	1	/	11	/
<i>Juger</i>	/	2	/	/
<i>Nier</i>	1	/	/	1
<i>Répondre</i>	3	/	/	/
<i>Requérir</i>	/	/	2	/
Occ. Tot.	6/16 (37,5%)	2/5 (40%)	14/31 (45,16%)	1/7 (14,3%)

Tableau 7b. Verbes relatif au *dire* dans le GC

	P1	P2	P4	P5
<i>Appeler</i>	/	/	1	/
<i>Demander</i>	2	/	/	/
<i>Dire</i>	4	4	15	/
<i>Enjoindre</i>	1	/	/	/
<i>Juger</i>	2	/	/	/
<i>Jurer</i>	/	1	/	/
<i>Mentir</i>	1	/	/	/
<i>Nommer</i>	/	/	2	/
<i>Parler</i>	/	/	1	/
<i>Plaindre</i>	8	/	/	/
<i>Promettre</i>	/	2	/	/
<i>Répondre</i>	/	1	/	/
<i>Requérir</i>	1	/	/	/
<i>Traiter</i>	/	/	3	/
Occ. Tot.	19/63 (30,16%)	8/18 (44,44%)	22/32 (68,75%)	0/0

Une comparaison rapide avec les sujets P3 permet de faire émerger la spécificité de ces verbes dans notre problématique : si nous ne nous arrêtons ainsi que sur *dire*, sans doute le verbe au sens le plus générique ici, et que nous mettons en perspective le nombre d'occurrences de chaque personne au regard de leurs occurrences totales, nous voyons bien que les pronoms de première et de deuxième personnes se distinguent dans ce corpus au regard de la P3 (Tableau 8).

Tableau 8. Fréquence d'emploi (en %) du verbe *dire*

	P1	P2	P3	P4	P5
1250 – TAC	1/16 (6,25%)	0/5 (0%)	11/674 (1,6%)	11/31 (35,5%)	0/7 (0%)
1300 – GC	4/63 (6,3%)	4/18 (22,2%)	20/1072 (1,9%)	15/32 (46,9%)	0/0

Quand bien même faudrait-il mener le même type d'analyse avec les verbes relatifs à la parole accompagnant les pronoms de troisième personne, nos observations établissent que ce type de verbe a une accointance marquée avec les embrayeurs énonciatifs que sont les pronoms personnels de première et de seconde personne. Ces résultats montrent également une progression de ce phénomène pour la P2 et la P4, au regard de la P1 dont l'association avec ces verbes relatifs à la parole est quasi identique dans nos deux textes. Une fois

encore, il faudrait mener cette étude sur d'autres corpus, et avec davantage d'occurrences, pour confirmer cette observation.

4 Discussions

Malgré le faible nombre d'occurrences analysées, il semble se dessiner une relation dans notre corpus entre l'expression du pronom personnel sujet et le sémantisme des verbes accompagnant ces derniers, dans une perspective rappelant les hypothèses de Detges (2003) à ce propos. Il y a effectivement ici quelque chose de l'ordre de l'emphase des personnes du discours, et nous savons que ces pronoms ont joué un rôle prépondérant dans l'expression et l'antéposition de la personne verbale à l'écrit. Cela est d'autant plus surprenant que l'on pourrait intuitivement s'attendre que ces textes de spécialité ne présentent que peu de personnes du discours au regard des textes littéraires ; de plus, ces personnes du discours s'orientent tant du côté de la mise en scène propre au discours rapporté (8) que du côté de l'énoncé métadiscursif, notamment dans le *GC* où la P4 est très régulièrement employée par l'auteur – anonyme – du texte pour hiérarchiser son discours (9). Il emploie ce faisant un « *nous* de modestie », dont l'apparition, selon le *Dictionnaire historique de la langue française* (Rey, 1998), est de quelques années antérieure à la rédaction du *GC*¹⁶.

(8) cil qui est retez dit. **Je fui pris** en ma costume. (celui qui est arrêté dit : « je fus pris en ma coutume ». *TAC*, p. 31 / Marnier, p. 43)

(9) **Nos uolon** premierement mostrer les distinctions et les chapitres de chascune partie (*GC*, p. 1)

Du reste, cette dernière observation nous permet également de justifier l'absence d'occurrences de la P5 dans le *GC* au regard du *TAC* où on la rencontrait, même si ce n'était qu'à sept reprises seulement. P1 et P2, tout comme P4 et P5 font système ; mais dans la mesure où la majorité des occurrences de la P4 ici renvoient, référentiellement, à une personne singulière, « substitut rhétorique de *je* » (Riegel *et al.*, 2014 : 363), il est légitime de ne trouver aucune occurrence de *vous*, *tu* remplissant parfaitement ce rôle. Cela ajoute un argument supplémentaire quant à l'hypothèse selon laquelle les stratégies d'emphase ont joué un rôle dans les expressions des pronoms sujets, la chose se rencontrant jusque dans ces textes de spécialité dans lesquels, pourtant, cet aspect stylistique est moins développé que dans les textes littéraires.

Il est ainsi notable que l'expression et l'antéposition des pronoms de première et de deuxième personne, singulier comme pluriel, se concentrent sur des verbes relatifs à l'expression du *dire* dans notre corpus, comme s'il était besoin de préciser en priorité les personnes en charge du propos avant les autres types de procès-verbaux, et que nous nous situions dans une situation de délibération juridique (8) ou de démonstration scientifique (9). L'usage aurait par la suite fait disparaître l'effet d'emphase de ces personnes du discours, qui auraient par la suite servi de modèle pour régulariser l'expression du pronom sujet à l'écrit pour toutes les personnes verbales. Par ailleurs, la concomitance de ce phénomène avec l'émergence des variations rhétoriques de certains pronoms à l'instar du *nous* de modestie dont nous parlions précédemment, inviterait à voir une continuité entre les phénomènes : la disparition de l'emphase portée par la seule expression du sujet pronominal aurait invité les locuteurs à développer de nouvelles stratégies rhétoriques, et à créer de nouveaux emplois (tels le *nous* de modestie ou celui de majesté) pour certains pronoms. Cette hypothèse demanderait cependant à être explorée dans une étude plus large, prenant en compte la dimension argumentative des textes.

5 Conclusion

L'importance du paramètre de l'emphase se confirme au sein du corpus considéré. Nous avons ainsi pu constater que les pronoms de première et de deuxième personne anticipaient la dynamique du changement en français, d'une part en étant plus souvent exprimés que les pronoms de troisième personne, d'autre part et lorsqu'ils sont exprimés, en étant davantage antéposés au verbe que les pronoms de troisième personne. Ces résultats sont cependant à pondérer dans la mesure où le nombre d'occurrences d'une part, et le peu d'écarts relevés dans les chiffres de l'autre, empêchent d'être catégorique. Nos résultats n'apportent cependant aucune contradiction au regard des analyses précédentes, et montrent l'importance du paramètre de la personne grammaticale dans la stabilisation de l'expression et de l'antéposition du sujet pronominal dans l'histoire de la langue française.

En outre, notre étude approfondit ce paramètre de l'emphase par l'intermédiaire des verbes les plus fréquemment employés avec les pronoms de première et de deuxième personne, et qui ont souvent partie liée au dire. Ce phénomène pouvant néanmoins être caractéristique du genre juridique que représentent ces coutumiers, il nous faudrait prolonger cette étude sur un corpus plus varié afin de vérifier si, comme nous l'avons observé ici, le sémantisme du verbe a un rôle à jouer quant à cette problématique. Ce travail illustre ainsi la nécessité de faire appel aux facteurs externes, génériques ou rhétoriques, en complément des facteurs internes, afin d'expliquer les dynamiques du changement.

Références bibliographiques

- Axel, K. (2005). Null subjects and verb placement in Old High German. Kepser S. & Reis, M. (éd.). *Linguistic Evidence: Empirical, Theoretical and Computational Perspectives*. Berlin : De Gruyter. p. 27-48.
- Balon, L. & Larrivée, P. (2016). L'ancien français n'est déjà plus une langue à sujet nul : nouveau témoignage des textes légaux. *Journal of French Language Studies*, 26/2, p. 221-237.
- Besnier, G. & Genestal, R. (1912). *Instrucions et enseignemens. Style de procéder d'une justice seigneuriale normande*. Caen : Jouan.
- Buridant, C. (2000). *Grammaire nouvelle de l'ancien français*. Paris : SEDES.
- Detges, U. (2003). L'obligatorisation des pronoms sujets en ancien français dans une perspective pragmatique et comparative. *Verbum*, 25, p. 307-335.
- Fournier, P. (1882). Coutumiers de Normandie, textes critiques, publiés, avec notes et éclaircissement, par Ernest-Joseph Tardif. *Bibliothèque de l'École des chartes*, 43, p. 526-528.
- Holmberg, A. (2010). Verb second. Chapitre 12 de Kiss, T. & Alexiadou, A. (éds.). *Syntax, an International Handbook of Contemporary Syntactic Research*, deuxième édition, vol. I. Berlin : de Gruyter Mouton. p. 342-383.
- Koenenman, O. & Zeijlstra, H. (2014). The rich agreement hypothesis rehabilitated. *Linguistic Inquiry*, 45/4, p. 571-615.
- Larrivée, P. & Capin, D. (2018). Types de textes et changement syntaxique. Neveu, F. et al. (éd.), *Congrès Mondial de Linguistique Française*. Disponible sur <<https://www.shs-conferences.org>>.
- Lay, M.-H. & Pincemin, B. (2010). Pour une exploration humaniste des textes : AnaLog. Bolasco S., Chiari, I., Giuliano, L. (éd.), *JADT 2010 : 10th International Conference on Statistical Analysis of Textual Data*. Rome : Edizioni Universitarie di Lettere Economia Diritto.
- Leclerc, J. (2019). Histoire du français. Disponible sur <<http://www.axl.cefan.ulaval.ca/francophonie/perioderomanestrasbourg.htm>>.
- Lobo, M. & Martins, A. M. (2017). Subjects. Stark, E. & Dufter, A. (éd.). *Manual of Romance morphosyntax and syntax*. Berlin : De Gruyter Mouton. p. 27-88.

- Marchello-Nizia, C. (2018). De SO à SV : Vers le sujet obligatoire et antéposé en français, les dernières phases d'un changement. *Journal of French Language Studies*, 28, p. 1-19.
- Marnier, M. (1839). *Établissements et coutumes, assises et arrêts de l'échiquier de Normandie (de 1207 à 1245)*. Paris : De Stahl.
- Prévost, S. (2015). Diachronie du français et linguistique de corpus : une approche quantitative renouvelée. *Langages*, 197, p. 23-45. Disponible en ligne sur <<https://www.cairn.info/revue-langages-2015-1-page-23.htm>>.
- Rey, A. (dir.) (1998). *Dictionnaire historique de la langue française*. Paris : Le Robert.
- Riegel, M. et al. (2014). *Grammaire méthodique du français*. Cinquième édition. Paris : PUF.
- Simonenko, A., Crabbé, B. & Prévost, S. (2019). Agreement Syncretisation and the Loss of Null Subjects : Quantificational Models for Medieval French. *Language Variation and Change*, p. 1-27. <doi:10.1017/S0954394519000188>.
- Tardif, E.-J. (1881). *Coutumiers de Normandie. Textes Critiques*. Rouen : Cagniard.
- Van der Auwera, J. (2009). The Jespersen cycles. Van Gelderen, E. (éd.), *Cyclical change*. Amsterdam : Benjamins. p. 35-71.
- Walkden, G. (2013). Null subjects in Old English. *Language variation and change*, 25/2, p. 155-178.

¹ Nous soulignons le sujet en gras dans les exemples (à l'exception des exemples 4 et 5), et le verbe en italiques.

² Pour un survol du débat, voir par exemple Koenenman & Zeijlstra (2014).

³ « L'étude menée sur le recul de la non-expression et de la postposition du sujet pronominal (pronoms personnels de 1^{re}, 3^e et 6^e personnes) dans les propositions déclaratives s'appuie sur un corpus de 14 textes (458 000 mots), étiquetés morpho-syntaxiquement, et qui s'étalent de 1100 à 1400. Bon nombre proviennent de la *Base de Français Médiéval* et ont été exploités avec le logiciel TXM. » (Prévost, 2015 : 36). Sur ces 14 textes, 8 d'entre eux relèvent de la narration littéraire et 3 de la chronique historique. Un texte juridique, un texte religieux et un manuel didactique complètent le corpus.

⁴ Nous renvoyons aussi à Marchello-Nizia (2018) sur cette question.

⁵ Ce qui explique le « pic » de 1205 de la Figure 2, puisqu'il appartient à une chronique historique, *La Conquête de Constantinople* de Robert de Clari.

⁶ Les textes datent des VIII^e-IX^e siècles.

⁷ Voir notamment Buridant (2000 : 711-713). Voir aussi la bibliographie de Van der Auwera (2009).

⁸ <<http://elec.enc.sorbonne.fr/cartulaires/>> (consulté le 12 novembre 2019).

⁹ Ce qui, du reste, a peut-être une incidence sur l'expression du sujet, en des proportions qu'il convient d'établir cependant.

¹⁰ <<https://conde.hypotheses.org/>> (consulté le 12 novembre 2019).

¹¹ « Le mss. F.f.2 de la Bibliothèque Sainte-Geneviève, dont nous imprimons une partie, provient de l'Église de Saint-Lo à Rouen [...] ; il est écrit sur vélin, avec le plus grand soin et d'un format in-4° ; enrichi de lettres majuscules ornées de traits de différentes couleurs. » (Tardif, 1881 : XXV). Une transcription, très fidèle à l'original à l'exception de la ponctuation qui a été modernisée, se trouve dans Marnier (1839 : 6-86). Pour une description du travail de Tardif sur l'ouvrage, voir Fournier (1882). Pour un compte-rendu de l'ouvrage de Marnier, voir Bordier (1843).

¹² <<https://hls.harvard.edu/library/digital-collections/grand-coutumier-de-normandie/>> (consulté le 12 novembre 2019).

¹³ <<https://transkribus.eu/Transkribus/>> (consulté le 15 novembre 2019).

¹⁴ Le projet a effectivement débuté en décembre 2018, et se terminera en septembre 2021.

¹⁵ Ce sont les verbes *Appeler, Avoir, Bailler, Défendre, Demander, Dénier, Devoir, Dire, Donner, Écrire, Enjoindre, Essayer, Être, Faire, Garantir, Garder, Jurer, Mander, Mentir, Montrer, Nier, Nommer, Obliger, Ouïr, Pardonner, Parler, Payer, Pendre, Plaindre, Pouvoir, Prendre, Prêter, Promettre, Recevoir, Répondre, Requérir, Rester, Savoir, Succéder, Tenir, Traiter, User, Vendre, Vouloir*.

¹⁶ « En français, les emplois de *nous* comme pronom personnel sont fixés avant la fin du XII^e s. [...] **Des emplois stylistiques se sont développés au XIII^e s.** : *nous* désigne le locuteur seul, non associé à

d'autres personnes (1200), spécialement dans l'emploi dit "de modestie" (1213) et comme en latin médiéval, dans celui dit "de majesté", pour un grand personnage (1241) » (Rey, 1998 : 2400, c'est nous qui soulignons).