

Zinaida GEYLIKMAN

*BARON ET CHEVALIER*  
EN FRANÇAIS MÉDIÉVAL

Une étude sémantique  
de noms d'humains  
dans la société féodale



PARIS  
HONORÉ CHAMPION ÉDITEUR  
2022

[www.honorechampion.com](http://www.honorechampion.com)

# INTRODUCTION

## 1. POINT DE DÉPART ET OBJECTIFS DE L'ANALYSE

Depuis les années 50 du  $xx^e$  siècle, les études sémantiques et lexicologiques du français médiéval ont essayé diverses méthodes d'analyse, s'appuyant sur les différents courants des théories sémantiques. Notre étude se place dans la perspective des travaux en sémantique lexicale autour des noms d'humains (plus loin NH) (Schneidecker & Mihatsch 2018, Aleksandrova 2014). Il s'agit d'une classe hétérogène qui inclut tous les items pouvant servir à désigner l'être humain. Selon Catherine Schneidecker, ces items-étiquettes sont de nature variée :

certaines de ces étiquettes sont valables une bonne partie de l'existence de l'individu [...]; d'autres ont une validité limitée dans le temps [...]; certaines de ces étiquettes sont en quelque sorte « prédestinées », programmées pour un individu donné [...], d'autres sont aléatoires [...]; d'autres étiquettes [...] n'émanent pas d'une programmation « naturelle », mais valent pour un ensemble conséquent d'humains, quel que soit l'âge, la culture *etc.* (Schneidecker, 2018 : 9).

Les deux items qui nous intéressent font partie d'une sous-classe assez vaste des NH du français médiéval que nous appellerons « noms d'humains féodaux » (plus loin NHF). Nous les définissons comme des NH exprimant les rôles sociaux qui relèvent de l'organisation féodale de la société. Du point de vue référentiel, les NHF que nous allons étudier représentent des dénominations<sup>1</sup>.

---

<sup>1</sup> Selon Georges Kleiber, pour avoir la valeur d'une dénomination, l'item lexical doit posséder deux caractéristiques sémantiques :

- a. l'indication qu'il s'agit d'une (catégorie de) chose(s), en somme d'un « tout », avec ses limites ;
- b. la description ou représentation du type de choses dont il s'agit (Kleiber, 2004 : 36).

Dans les travaux précédents nous avons esquissé une étude sémantique de *baron* dans trois chansons de geste du XII<sup>e</sup> siècle et une branche du *Roman de Renart*. Cette étude a permis de relever deux tendances : l'importance de l'emploi évaluatif des formes monosyllabiques du NH étudié et une variation considérable dans son contenu sémantique entre les deux types de discours. De cette première esquisse est né le projet de notre travail de thèse<sup>2</sup> dont le présent ouvrage est issu. Dans le cadre de la thèse, nous avons confronté *baron*<sup>3</sup> avec trois autres NHF qui, d'après nos observations, se trouvaient régulièrement en co-référence avec l'item – *chevalier*, *bachelier* et *vassal*. Le contenu sémantique des quatre NHF a été étudié à partir d'un corpus contenant quatre types de discours – chansons de geste, romans, chroniques, textes documentaires – et s'étendant sur la période du XII<sup>e</sup> au XV<sup>e</sup> siècle.

Il faut dire que les quatre items avaient déjà servi d'objet d'étude pour d'autres recherches lexicologiques et sémantiques : Hollyman (1957) pour *baron*, *chevalier* et *vassal*, Flori (1975a, b) pour *bachelier* et *chevalier*, Venckeleer (1975) pour *baron*, Eskenazi (1976) pour *chevalier*, Denis (1989) pour *baron* et *chevalier*. Cependant, si nous avons décidé de nous pencher sur les dénominations déjà étudiées, c'est que, pour diverses raisons, aucun de ces ouvrages ne présentait le contenu sémantique des items de façon structurée. Par ailleurs, tous les travaux cités, sauf celui de Venckeleer, se basaient sur un corpus restreint à un seul genre et/ou à une diachronie courte (dans la plupart des cas, en ancien français), alors que notre dessein était de suivre l'évolution (ou de constater la non-évolution) du contenu sémantique des items sur la majeure partie de la période vernaculaire écrite, ainsi que d'étudier l'éventuelle variation entre les genres textuels.

Ces quatre NHF étaient d'autant plus intéressants à étudier qu'au premier regard leur sens apparaissait évident pour un médiéviste<sup>4</sup>. Proposer une analyse proprement sémantique des dénominations qui servent à désigner les rôles sociaux étudiés de fond en comble, aussi bien

<sup>2</sup> Thèse « Étude sémantique de *baron* et *chevalier* (XII<sup>e</sup>-XV<sup>e</sup> siècle) : essai de méthode pour les dénominations féodales » préparée à l'équipe SAPRAT de l'EPHE de Paris sous la direction de Joëlle Ducos, soutenue le 28 novembre 2017.

<sup>3</sup> En suivant le principe du DMF, dans tous les cas où il ne s'agira pas de citation de textes médiévaux, nous choisirons pour les items la forme graphique se rapprochant le plus de l'orthographe moderne.

<sup>4</sup> Nous avons ainsi constaté avec surprise que l'entrée *chevalier*, dont le sens paraissait vraisemblablement trop évident, était absente du corps du dictionnaire GDF, n'apparaissant que dans le GDF Complément.

par les historiens que par les spécialistes de la littérature, s'est révélé un véritable défi.

Toutefois, lors de l'étude des occurrences nous avons constaté une disparité numérique majeure entre *baron* et *chevalier* d'un côté et *bachelier* et *vassal* de l'autre côté : les exemples des premiers abondaient, alors que les occurrences des derniers étaient bien plus rares, voire absentes de certains types de textes (par exemple, des romans en prose du XIII<sup>e</sup> siècle). Cette disparité nous a empêchée de nous pencher sur le contenu sémantique de *bachelier* et *vassal* de façon aussi détaillée que pour *baron* et *chevalier*.

Le présent ouvrage résume donc l'analyse que nous avons menée dans le cadre de notre travail de thèse sur le contenu sémantique de *baron* et *chevalier* uniquement. Il répondra à plusieurs objectifs :

- sur la base d'une analyse exhaustive d'occurrences relevées dans un corpus large et représentatif, étudier le contenu sémantique des deux NHF dans quatre types de discours en français médiéval ; présenter les résultats de façon détaillée et structurée ;
- confronter les tendances relevées pour les deux items et étudier leur interaction ;
- esquisser des pistes de réflexion pour l'étude de l'ensemble lexical des NHF.

L'accès au contenu sémantique de nos objets de recherche passe pour nous par l'étude de leur distribution – co-occurrence régulière avec d'autres éléments, appartenance récurrente à telle ou telle construction syntaxique ou situation narrative. Nous réunissons les diverses conditions co-textuelles et contextuelles d'apparition des items sous le terme d'*emploi* (voir plus loin). Si dans le texte original de notre thèse nous avons décrit de manière détaillée la quasi-totalité des emplois relevés pour *baron* et *chevalier*, dans le présent ouvrage nous le ferons de manière synthétique. Ainsi, nous allons sélectionner plusieurs emplois pour chaque type de discours permettant de mieux décrire le profil sémantico-référentiel de l'item au sein du type de discours en question. L'ensemble de ces emplois sélectionnés devra permettre à son tour d'entrevoir le contenu sémantique des deux NHF en français médiéval dans sa globalité.

Avant de passer à la description d'emplois, nous allons donner des précisions sur notre méthodologie en décrivant le corpus qui a servi de base pour notre recherche et les éléments des théories sémantiques qui ont constitué le cadre de notre analyse.

## 2. CORPUS

### DESCRIPTION GÉNÉRALE DU CORPUS ET JUSTIFICATION DES CHOIX

Dans notre choix du corpus nous sommes basée sur trois critères tels que décrits par Georges Kleiber<sup>5</sup> :

- (1) représentativité
- (2) exhaustivité
- (3) homogénéité (Kleiber, 1978 : 65).

Nous allons nous arrêter sur les trois critères, ce qui nous permettra de justifier nos choix concernant plusieurs caractéristiques de notre corpus : dimension, périodisation, choix de langue et choix de genres.

Concernant la *représentativité*, il est apparu important d'avoir un corpus suffisamment large, même si la largesse du corpus est une notion relative. Selon Kleiber,

La quantité brute n'assure pas automatiquement la représentativité : il reste toujours un choix qualitatif à faire. Mais contrairement à B. Pottier pour qui « les études sur les champs ne progresseront que si l'on ne recherche pas la quantité des matériaux, mais leur qualité »<sup>6</sup>, nous pensons qu'une QUANTITÉ « choisie » constitue un élément d'appréciation qualitatif dont on aurait tort de se passer (1978 : 67).

Seul également un corpus large permet de pallier en partie notre manque de compétence. L'utilisation d'un seul idiolecte comporte le risque de généralisations hâtives, d'oublis et d'erreurs : on dégage bien souvent des structures qui n'ont jamais existé (*ibid.* : 69).

En effet, même dans les cas où un seul texte permet de dégager tel ou tel phénomène linguistique, seul un corpus large peut permettre de le classer comme une généralité, ou une innovation adoptée ultérieurement, ou encore un emploi déviant qui ne réussit pas.

Notre corpus comporte donc 48 textes sans compter dix recueils de chartes régionales et trois recueils de chartes royales du corpus DocLing<sup>7</sup>.

L'*exhaustivité* du corpus peut être illustrée le mieux par les choix chronologiques. Le point de départ – début du XI<sup>e</sup> siècle s'est imposé à nous

---

<sup>5</sup> Tout en produisant un développement original à leur sujet, Kleiber emprunte lui-même ces critères à la *Sémantique structurale* de Greimas (1966 : 143).

<sup>6</sup> Pottier, 1968 : 40.

<sup>7</sup> Pour la liste complète des textes et leur références bibliographiques, voir *Bibliographie*.

en raison du choix des quatre types de discours dont on ne connaît pas de témoins antérieurs à cette époque. Ayant pour but de nous concentrer sur le français médiéval, nous avons posé l'autre limite à la fin de la périodisation conventionnelle – la fin du xv<sup>e</sup> siècle. Tout en étant consciente du caractère artificiel de toute périodisation rigide (pour la période du français médiéval, voir les travaux de Christiane Marchello-Nizia p.ex. 1979, Combettes & Marchello-Nizia, 2010), nous emploierons le terme *ancien français* pour parler des xii<sup>e</sup>-xiii<sup>e</sup> siècles, *moyen français* pour désigner les xiv<sup>e</sup> et xv<sup>e</sup> siècles.

Pour chaque genre nous avons donc tenté de sélectionner un certain nombre de textes de façon à ce que chaque micro-période (50 ans) soit représentée. Cependant, pour avoir un rapport plus direct au système linguistique, nous avons fait le choix d'éliminer les nombreuses réécritures en prose<sup>8</sup> dans le but de travailler à partir des textes qui se rapprochaient le plus des « originaux », aussi relative que soit cette notion pour le Moyen Âge<sup>9</sup>. Il en découle que, de manière générale, les textes de fiction en moyen français, surtout ceux du xv<sup>e</sup> siècle, sont moins nombreux que ceux en ancien français. Le calcul de la fréquence relative pour divers phénomènes permet de pallier partiellement ce déséquilibre ; par ailleurs, il est apparu important d'avoir un nombre considérable de textes de fiction du xii<sup>e</sup> et du xiii<sup>e</sup> siècle dans la mesure où il s'agissait de la période la plus reculée et que cela permettait de mieux cerner le contenu sémantique des items au point de départ<sup>10</sup>.

En ce qui concerne les chroniques, le xiii<sup>e</sup> et le xiv<sup>e</sup> siècles sont les mieux représentés pour des raisons objectives, puisque la production

---

<sup>8</sup> En revanche, certaines compilations, qui gardent la forme du vers, ont été retenues malgré tout, surtout dans les cas où elles ne relevaient pas d'une seule source bien identifiable. Ainsi, l'*Istoire le roy Charlemagne* de Girart d'Amiens a été incluse dans le corpus. Ce qui est apparu primordial pour nous, du point de vue linguistique, c'est qu'il n'y ait pas de « traduction » directe, ni de passage du vers en prose.

<sup>9</sup> En effet, l'étude des différentes versions progressives présentes dans les manuscrits, des versions en prose et des réécritures représente un grand intérêt du point de vue linguistique. Elle nécessite, toutefois, une approche comparative avec les textes originaux ce qui n'est pas notre perspective de recherche. Pour l'analyse linguistique des versions en prose, voir Andrieux-Reix, 2001, Maria Colombo-Timelli 2011, 2012.

<sup>10</sup> Ce déséquilibre correspond, par ailleurs, aux deux grandes étapes d'existence des chansons de geste et des romans qui connaissent, avec un certain décalage et chacun à sa façon, l'essor au xii<sup>e</sup> et au xiii<sup>e</sup> siècles, et un déclin progressif à partir du xiv<sup>e</sup> siècle. Boutet et Strubel l'expliquent, entre autres, par l'évolution profonde de la société : « Le début de la guerre de Cent ans marque un temps d'arrêt dans l'activité littéraire : l'idéologie qui sous-tend les œuvres épiques et courtoises va être sérieusement mise en question par la réalité (1979 : 141) ».

des textes historiques en langue vernaculaire a été au XII<sup>e</sup> siècle plus « hésitante » que celle des textes littéraires. Face au foisonnement de l'historiographie en latin, la majorité des chroniques vernaculaires du XII<sup>e</sup> siècle sont « de simples traductions de chroniques latines, sans aucun apport original » (Boutet & Strubel, 1979 : 105). Nous avons donc été contrainte d'inclure dans notre corpus la *Chronique rimée du Mont-Saint-Michel* du XII<sup>e</sup> siècle en raison de sa nature de « production originale » en langue vernaculaire, malgré le fait qu'on n'y retrouve pas d'occurrences de *chevalier*, mais uniquement celles de *baron*. Inversement, le XV<sup>e</sup> siècle correspond, selon Pierre Courroux, à « une explosion du nombre d'histoires en vernaculaire » grâce au statut croissant du français face au latin, ainsi qu'à l'imprimerie et les nouveaux modes de conservation et de transcription de manuscrits (2016 : 917). Cependant, deux textes seulement font partie de notre corpus. Tout d'abord, parce que l'écriture historique prend des formes diverses : ainsi, les journaux (comme le *Journal d'un bourgeois de Paris*) et les mémoires (comme les textes de Philippe de Commines) n'ont pas été inclus en raison de leur nature différente de l'écriture historiographique plus ancienne et, surtout, le manque de pertinence pour l'analyse des dénominations qui nous intéressaient. Par ailleurs, force est de constater que les textes historiques du XV<sup>e</sup> siècle connaissent très peu d'éditions en général, alors que les éditions existantes datent surtout du XIX<sup>e</sup> siècle (c'est le cas de deux textes du XV<sup>e</sup> siècle de notre corpus) et ne sont pas conformes aux principes des éditions critiques actuels.

Le corpus des textes documentaires est celui qui relève le moins d'exhaustivité. On constate tout d'abord qu'en raison de la prépondérance du latin dans le discours documentaire sur toute la période médiévale, les éditions des textes documentaires vernaculaires se font encore rares<sup>11</sup>. Nous avons donc fait appel à plusieurs corpus électroniques : le DocLing qui contient majoritairement des chartes régionales et royales, le Frantext Moyen français d'où l'on a tiré le *Registre Criminel du Châtelet* et celui de l'ANOH où l'on trouve un nombre important de textes documentaires. Or, si les chartes en langue d'oïl du DocLing appartiennent surtout au XIII<sup>e</sup> siècle, avec quelques rares exemples du XIV<sup>e</sup> siècle, la majorité des

---

<sup>11</sup> Étant donné l'étendue de notre corpus, nous avons fait le choix de ne pas nous tourner vers les manuscrits et de travailler uniquement à partir des éditions déjà existantes.

documents de l'ANOI datent du <sup>xiv</sup><sup>e</sup> siècle, tandis que les documents du <sup>xiii</sup><sup>e</sup> et du <sup>xv</sup><sup>e</sup> siècles sont exceptionnels<sup>12</sup>.

En parlant de chronologie, nous devons dire que notre façon de procéder n'est donc pas purement diachronique : nous analysons une multitude de synchronies représentées par nos textes dont la somme donne un résultat diachronique – l'évolution ou la non-évolution.

Lors de l'analyse nous allons nous appuyer sur les données statistiques décrivant les fréquences des emplois calculées pour chaque siècle. Nous faisons en effet une approximation, d'autant plus que certains textes posent des problèmes de datation ou sont présentés comme étant rédigés au seuil des deux siècles<sup>13</sup>. Gabriella Parussa met en évidence la pertinence de l'étude des phénomènes linguistiques du français médiéval dans une diachronie « à l'aune de l'individu » (Parussa, 2014 : 73). Même si nous faisons des calculs de statistiques pour une période de cent ans, dans les cas où ce sera pertinent, nous ferons également la distinction entre la première partie du siècle et sa deuxième partie. Cependant, dans la mesure où notre chronologie est assez large et qu'il s'agit de changements sémantiques qui sont souvent difficiles à dater avec précision, notre objectif était avant tout de mettre en évidence les tendances générales concernant leur emploi et de dégager les grands repères de l'évolution du contenu sémantique des items étudiés. Une datation précise de l'émergence d'un sens, par exemple, nécessiterait une chronologie plus restreinte, une « micro-diachronie », mais un corpus plus dense ; notre étude met en lumière les limites de la période où cette datation serait à rechercher<sup>14</sup>.

Il reste enfin à aborder la question épineuse de la date attribuée aux textes dans notre analyse. Étant donné les particularités de la transmission

---

<sup>12</sup> L'absence de textes du <sup>xii</sup><sup>e</sup> siècle dans les deux sous-corpus s'explique par une extrême rareté historique de la production documentaire en langue d'oïl, aussi bien continentale qu'insulaire, antérieure au <sup>xiii</sup><sup>e</sup> siècle (Glessgen, 2003). De manière générale, le domaine du droit reste un des plus marqués par la concurrence entre le latin et le vernaculaire sur toute la période étudiée. Pour les raisons du maintien du latin, voir Marcotte, 2014.

<sup>13</sup> Dans ce cas-là, nous avons fait le choix de calculer les statistiques concernant ces textes pour le siècle plus tardif, par exemple, dans les tableaux statistiques la chanson de geste *Raoul de Cambrai* figurera parmi les textes du <sup>xiii</sup><sup>e</sup> siècle. Cependant, dans les cas où ce sera pertinent, nous préciserons cette position de l'« entre-deux-siècles » des textes en question.

<sup>14</sup> Au niveau « extra-linguistique » plusieurs autres périodisations peuvent se révéler pertinentes comme celle de Boutet et Strubel qui mettent en évidence le clivage thématique et structurel des textes littéraires (y compris la littérature historique) entre la période antérieure à 1250 et la période ultérieure (1979 : en particulier, 21)



des textes au Moyen Âge, pour la majorité des textes de fiction et des chroniques de notre corpus la version éditée n'est pas la version originale, mais se base sur un manuscrit plus tardif. Seuls les textes documentaires représentent des originaux à proprement parler. Il serait donc logique pour une étude linguistique de s'appuyer sur la date du manuscrit. Or, nous nous sommes rendu compte que cette méthode ne permettait pas de mettre en évidence les évolutions des emplois et du contenu sémantique des items étudiés : contrairement aux évolutions qui sont faciles à enregistrer au niveau de la graphie, par exemple, les emplois des NHF qui nous intéressaient semblaient être reproduits de manuscrit en manuscrit avec le contenu sémantique qui leur était propre. Pour ne citer qu'un exemple, l'écart numérique entre les occurrences de *baron* et de *chevalier* augmente progressivement entre les romans en vers du XII<sup>e</sup> siècle – les textes de Chrétien de Troyes – et les romans du XIII<sup>e</sup> siècle, ce qui, comme nous le verrons, est une manifestation de l'évolution sémantique des deux items. Or, si nous avons considéré la date du manuscrit et non la date de composition, nous aurions perdu cette dimension diachronique, puisque les manuscrits des romans composés au XII<sup>e</sup> siècle dataient majoritairement du XIII<sup>e</sup> siècle. Nous pensons pourtant, que la dimension diachronique est très importante dans ce cas-là, parce qu'elle rend compte de l'évolution du genre et de l'expression linguistique de ce genre. En omettant cette perspective, nous aurions donc obtenu une analyse erronée.

Dans tous les cas, en parlant d'un texte de tel ou tel siècle, nous indiquerons la date de sa composition (en-dehors des textes documentaires où les versions éditées sont de véritables « textes originaux »).

*L'homogénéité* du corpus est assurée par deux aspects : l'unité linguistique et les limites de quatre types de discours.

Notre but est d'étudier le contenu sémantique de *baron* et *chevalier* au sein d'un seul domaine linguistique au sens large. Les textes de notre corpus sont donc tous rédigés en langue d'oïl, aussi bien continentale (où plusieurs variétés régionales sont présentes) qu'insulaire. Au sujet de l'anglo-normand, nous nous appuyons sur le point de vue de David Trotter qui le considère non pas comme une variété isolée, mais comme une variété insulaire de la langue d'oïl (Trotter, 2003 : § 16). Lors de notre analyse nous n'adoptons pas spécialement la perspective isotopique : dès les débuts de l'analyse nous avons constaté qu'en matière de la distribution et du contenu sémantique, les différences entre les variétés régionales étaient minimales. Contrairement au niveau phonétique ou à la graphie, le sens des deux NHF, qui appartiennent sans doute au fonds commun du lexique de la langue d'oïl, est en grandes lignes le même dans

tout le domaine linguistique. Nous ferons donc intervenir la perspective isotopique uniquement dans les cas où une variation dans l'emploi ou le contenu sémantique serait effectivement enregistrée.

Un autre point important dans le choix linguistique concerne le latin qui, malgré l'évidence de la diglossie médiévale, restera en marge de notre analyse. Nous pensons qu'une étude proprement sémantique, dans le cas où elle ne sert pas à établir les liens étymologiques, mais concerne le contenu même des items, doit se concentrer sur un seul domaine linguistique. Dans le cas de l'étude de Hollyman (1957) dont nous parlerons plus loin, l'appel aux textes latins est justifié, puisque sa périodisation concerne essentiellement le haut Moyen Âge et que son but est d'observer l'émergence du vocabulaire féodal<sup>15</sup>. Dans notre cas, l'intégration des textes latins aurait nécessité une méthode d'analyse foncièrement différente. Lors de l'analyse, nous avons choisi de faire appel au latin uniquement par le biais des dictionnaires étymologiques et du latin médiéval.

De notre volonté de considérer les NHF au sein d'un seul domaine linguistique découle le choix de ne pas inclure de traductions directes du latin dans notre corpus. Il est vrai que nombre de textes (aussi bien textes de fiction que chroniques) sont, à des degrés différents, inspirés de sources latines et que la question de la « création originale » pose souvent problème pour la période médiévale, tant l'intertextualité est importante et les modes de transmission de textes – complexes. En laissant de côté la question des sources, nous avons sélectionné les textes qui n'étaient pas classés par les chercheurs qui les avaient abordés comme des traductions directes du latin<sup>16</sup>.

Paradoxalement, même si la diversité des textes que comporte notre corpus suggère son caractère hétérogène, l'homogénéité est également assurée par le choix des types de discours. En effet, les quatre genres textuels qui composent notre corpus mettent tous en scène les mêmes individus : les hommes nobles laïcs. D'une certaine façon, ils décrivent

---

<sup>15</sup> Toutefois, la méthode de Hollyman est discutable puisque son analyse des textes latins du x<sup>e</sup>-xii<sup>e</sup> siècles et des textes en très ancien et ancien français présuppose une continuité directe qui n'existe pas ou, en tout cas, reste à prouver dans la mesure où l'écriture en langue vernaculaire était indéniablement ultérieure à une période d'existence orale. Par conséquent, les valeurs sémantiques des textes vernaculaires du x<sup>e</sup>-xii<sup>e</sup> siècles ne sont pas nécessairement « héritées » des textes latins, mais remontent à une période plus reculée.

<sup>16</sup> Les traductions médiévales peuvent se révéler tout-à-fait pertinentes pour l'étude des changements linguistiques (voir en particulier, Parussa, 2014). Ces études exigeraient cependant des méthodes différentes par rapport à celles que nous allons adopter dans notre recherche.

le même objet – les occupations (principalement, militaires) de ces hommes – mais sous des angles différents : idéologique pour les chansons de geste et les romans, factuel pour les textes documentaires, et un angle mélangeant le factuel et l'idéologique pour les chroniques. Ce sont donc ces quatre genres qui sont les plus aptes à s'intéresser aux rôles sociaux auxquels réfèrent les NHF étudiés. Autrement dit, tout en étant varié du point de vue discursif, notre corpus est avant tout caractérisé par l'unité d'objet des textes<sup>17</sup>.

#### RESSOURCES NUMÉRIQUES ET CONSTITUTION DU CORPUS

Il convient enfin de préciser quelques points sur les supports des textes que nous avons sélectionnés pour notre corpus. Ils proviennent principalement de six bases textuelles : la Base du Français Médiéval (=BFM), le Frantext Moyen Français, la base textuelle de l'Anglo-Norman On-Line Hub (=ANOHL) et la base des *Plus anciens documents linguistiques gallo-romans* (=DocLing)<sup>18</sup>. Pour les textes disponibles en format numérique, les fréquences absolues ont été calculées de façon automatique. Cependant, dans le but de constituer un corpus représentatif, exhaustif et homogène nous avons également dû avoir recours à un certain nombre d'éditions papier. Les données statistiques qui ressortent des éditions papier, pour lesquels nous avons effectué un relevé manuel, sont donc moins fiables que celles que nous avons obtenues de façon automatique. Néanmoins, nous pensons que l'étendue de notre corpus doit permettre de dégager des tendances générales qui nous intéressent malgré les valeurs parfois approximatives des données statistiques.

Pour résumer, nous pouvons dire que nous avons choisi comme corpus un ensemble de textes appartenant aux quatre genres suivants – chansons de geste, romans, chroniques et textes documentaires – dont la date de

---

<sup>17</sup> Il faut dire, toutefois, que la question d'attribution au genre textuel spécifique se révèle difficile dans certains cas. Il s'agit surtout des hésitations « chronique vs roman » et « chronique vs chanson de geste ». Dans la Préface à l'ouvrage de Pierre Courroux (2016) Edina Bozoky indique qu'il est souvent « difficile, sinon impossible, de délimiter le roman de la chronique » (*ibid.* : 11). Pierre Courroux l'explique par le caractère littéraire de l'historiographie vernaculaire, à l'opposé de la chronique latine (*ibid.* : 864). Le seul point qui, selon Courroux, permet de distinguer un texte de fiction d'une « chronique » est l'intention de l'auteur : divertissement ou écriture de l'histoire (*ibid.*) Dans nos choix concernant certaines chroniques, en nous appuyant sur les classifications existantes (en particulier, celles du catalogue JONAS), nous avons essayé d'appliquer ce principe de distinction.

<sup>18</sup> Pour les références bibliographiques complètes, voir *Bibliographie*.

composition se situe entre 1100 et 1499 et qui représentent des versions originales. Les autres versions, versions en prose, réécritures et traductions ont été pour la plupart éliminées<sup>19</sup>. Tout en comportant certains défauts, le corpus, tel que nous le concevons, est représentatif, exhaustif et homogène.

### 3. ÉLÉMENTS DE LA THÉORIE SÉMANTIQUE

Comme le remarque judicieusement Gérard Petit, « n'étant pas directement observable, la structuration sémantique d'une unité lexicale <...> reste tributaire de la configuration du modèle adopté pour sa représentation » (2010 : 103). En effet, la diversité des théories sémantiques et la multitude de terminologies nous apprennent avant tout qu'il n'y a pas une seule façon d'aborder le sens, mais qu'il peut être étudié sous plusieurs angles.

#### SÉMANTIQUE RÉFÉRENTIELLE

Le modèle de sens qui nous semble le plus approprié pour une analyse de dénominations féodales est celui de sens référentiel<sup>20</sup>. D'après le *Dictionnaire des sciences du langage* de Franck Neveu (2004 : 250), « la référence est généralement définie comme la relation qui unit une expression linguistique en emploi dans un énoncé avec « l'objet du monde » qui se trouve désigné par cette expression. On appelle référent cet « objet du monde » ». Le modèle proposé par Jean-Claude Milner vise à expliciter cette relation<sup>21</sup> :

Une unité lexicale étant choisie, certains segments sont d'emblée éliminés en tant que références possibles ; en ce sens, à chaque unité lexicale individuelle, est attaché un ensemble de conditions décrit donc un type (ou

<sup>19</sup> Exception faite de TurpinBrianeS, version anglo-normande du *Pseudo-Turpin* continental, que nous avons incluse pour représenter le domaine anglo-normand dans les chroniques du XIII<sup>e</sup> siècle.

<sup>20</sup> Georges Kleiber met en évidence la nécessité de recourir à des modélisations de sens différentes en fonction de l'objet d'analyse (1997).

<sup>21</sup> Le fait que certains textes que nous analysons sont des textes de fiction ne change rien du point de vue référentiel. Tout comme l'existence des référents imaginaires tels que *Tarzan* ou *la licorne* ne signifie pas que la référence n'a pas de lien avec la réalité (Kleiber, 1997 : 15), l'existence des référents de fiction ne signifie pas qu'il y ait un mode de référence particulier. Un texte de fiction construisant un monde sur la base du monde réel (*ibid.*), le mécanisme référentiel est le même pour tous les textes de notre corpus.

si l'on veut une classe) de référence possible ; il est distinct des segments de la réalité, mais pèse sur eux. Pour exprimer cette situation, on pourrait recourir aux termes suivants : le segment de réalité associé à une séquence est sa référence actuelle ; l'ensemble de conditions caractérisant une unité lexicale est sa référence virtuelle<sup>22</sup> (1982 : 10).

La référence actuelle correspond donc à cet « objet du monde », alors que la référence virtuelle représente « un ensemble de conditions d'applicabilité référentielle » (Kleiber, 1997 : 34) et correspond au sens référentiel de l'expression.

Cette distinction s'applique parfaitement à l'étude des NH sociaux. Tout NH réfère à un individu qui a un nombre de rôles sociaux illimité ; la fonction des NH est justement d'exprimer un ou plusieurs de ces rôles. Nous verrons au fil de l'analyse qu'un seul et même personnage peut être désigné par les deux NHF qui nous intéressent en fonction du contexte. Dans ce cas, le référent actuel serait le personnage ou l'individu concret, dans toute la plénitude de ses rôles sociaux, alors que la référence virtuelle contiendrait la description des rôles sociaux que l'on veut mettre en évidence dans un cas concret.

La distinction entre les deux types de référence aide également à ne pas confondre la réalité historique des personnages et individus désignés et les représentations mentales appelées par tel ou tel item, erreur que l'on est enclin à faire en travaillant sur un état de langue reculé. Nous verrons, par exemple, que dans certains cas les tendances que nous dégagerons au sujet de la référence virtuelle des items étudiés ne correspondront pas à la chronologie historique de l'évolution de la société féodale.

Lors de notre analyse nous allons distinguer les occurrences des NH par type référentiel : spécifique individuel, spécifique collectif et générique. La distinction se fait alors sur la base du référent actuel.

Lorsqu'on parlera de la référence spécifique individuelle, il s'agira de désigner des représentants concrets des « barons » ou « chevaliers » pris dans leur individualité. La référence spécifique collective est enregistrée dans les cas où les NH étudiés servent à désigner des collectivités anonymes. Le choix d'opposer l'individuel et le collectif au sein du type spécifique s'inspire de l'analyse de Françoise Denis (1989) qui l'intègre

---

<sup>22</sup> Le terme *référence virtuelle* recoupe parfois le même concept que le terme *dénotation* (Charolles, 2002 : 7) ; par ailleurs, la conception du sens référentiel s'apparente à celle de la sémantique du prototype, « puisque c'est sur la base de la ressemblance avec le prototype que se trouvent déterminés les référents possibles de la catégorie » (Kleiber, 1997 : 21).

dans son étude des emplois de *baron* et *chevalier* dans *Raoul de Cambrai*. Sans en faire un critère d'analyse, Theo Venckeleer (1975) met également en évidence la différence de sens entre *baron* appliqué à un seul personnage et le même item désignant une collectivité.

Le type générique concerne les occurrences où l'item désigne non pas un spécimen concret, mais un représentant type p.ex. :

Le chien est l'ami de l'homme.

Une femme aime son enfant.

Je veux épouser une Tahitienne<sup>23</sup> (quelle qu'elle soit) (cet exemple est emprunté à Kleiber, 1994b : 50)

Ces occurrences étant peu nombreuses, elles nous ont toutefois été utiles pour reconstituer l'image du représentant type des humains désignés par les items qui nous intéressent.

Autrement dit, en prenant l'exemple de la *Chanson de Roland*, la mention de Roland ou d'Olivier ou des deux personnages ensemble en tant que *chevalier* se rapporte au type *spécifique individuel*, la désignation de l'armée de Charlemagne par *chevalier* correspond au type référentiel *spécifique collectif*, l'occurrence de *chevalier* dans une réflexion d'un personnage sur ce que doit être un véritable «chevalier» (représentant type) serait du côté de la référence *générique*. Ces types ne correspondent donc pas aux catégories grammaticales du nombre et de la personne, mais relèvent uniquement du référent actuel.

#### LE CONTENU SÉMANTIQUE : LES TRAITS

Le sens référentiel étant décrit par le biais des traits sémantiques non différentiels, le but de notre analyse est de dégager ces traits et d'étudier par la suite les relations qu'ils entretiennent entre eux, aussi bien au sein du même item (contenu sémantique interne) qu'en-dehors de cet item (relations sémantiques externes). Nous visons ainsi une analyse componentielle telle que la décrit Anna Wierzbicka – le «portrait» du concept associé au mot (1985 : 39).

---

<sup>23</sup> La GMF traite, cependant, ce dernier type d'emploi comme un emploi spécifique du SN dont la référence n'a qu'une existence virtuelle (2009 [1994] : 294). Nous choisissons cependant de l'envisager sous le même angle que Kleiber dans la mesure où nous nous intéressons à l'opposition entre le type et l'exemple spécifique de ce type.

Notre vision du contenu sémantique d'un item lexical se base en grande partie sur les travaux de Catherine Kerbrat-Orecchioni qui ont mis en évidence la distinction entre plusieurs types ou niveaux de traits sémantiques. Cette distinction est d'ordre référentiel, dans la mesure où elle s'appuie sur le modèle des propriétés différentes présentes ou non dans la référence d'un item. Selon ce modèle, parmi les propriétés présentes dans la référence, on distingue celles qui sont présentes *toujours* qui peuvent être *distinctives* ou *non* de celles qui sont présentes *parfois*, ces dernières apparaissant soit *occasionnellement*, soit *fréquemment* (Kerbrat-Orecchioni, 1977 : 176). Ces propriétés sont exprimées par des traits sémantiques de qualité différente. Il est important de dire, cependant, que toutes les propriétés du référent actuel ne sont pas nécessairement reflétées par la référence virtuelle.

Le premier groupe de traits que nous dégagerons seront les traits dénotatifs. Franck Neveu définit la dénotation en linguistique de façon suivante :

Propriété que manifeste un concept de s'appliquer aux objets qui forment son extension. À partir de cette base logique, s'est élaborée la valeur du terme en sémantique lexicale, où *dénotation* est employé généralement pour désigner la relation entre un signe linguistique et une classe d'objets du monde. Couramment, la dénotation est entendue comme la signification invariante et non subjective du signe (Neveu, 2004 : 92).

Pour Kerbrat-Orecchioni il s'agit de « l'ensemble des traits de sens qui permettent la dénomination (encodage) et l'identification (décodage) d'un référent, c'est-à-dire, l'ensemble des traits strictement définitionnels » (Kerbrat-Orecchioni, 1977 : 12). La dénotation est alors opposée à la connotation qui se charge de tous les traits supplémentaires.

Pour certains sens de *baron* la distinction entre les traits dénotatifs objectifs<sup>24</sup> et les traits dénotatifs subjectifs sera très importante. Elle se base sur les propriétés décrites par ces deux types de traits :

---

<sup>24</sup> En gardant à l'esprit que lorsque l'on qualifie la conceptualisation du monde d'*objective*, on entend en réalité « qu'elle ne se trouve pas soumise aux variations subjectives d'un sujet percevant à l'autre, mais bénéficie d'une certaine stabilité intersubjective » (Kleiber, 1997 : 13). L'*objectivité* apparente des traits relève par conséquent d'une intersubjectivité partagée par la majorité des locuteurs.

Pour dénommer un individu *x*, je peux dire <...> :

(1) « c'est un professeur » : le terme énonce une propriété objective, facilement vérifiable, du dénoté ;

(2) « c'est un imbécile » / « c'est un génie » : ces substantifs cumulent deux types d'informations d'ailleurs indissociables :

– une description du dénoté ;

– un jugement évaluatif, d'appréciation ou de dépréciation, porté sur ce dénoté par le sujet d'énonciation.

Ces termes, dans la mesure où ils font intervenir une évaluation de *x*, laquelle est solitaire des systèmes d'appréciation du locuteur ; dans la mesure où leur usage, *x* restant invariant, pourra varier d'une énonciation à l'autre, dans la mesure enfin où ils sont à éliminer d'un discours à prétention d'objectivité, dans lequel le locuteur refuse de prendre position par rapport au dénoté évoqué, peuvent être considérés comme comportant un trait sémantique [subjectif]. (Kerbrat-Orecchioni, 1997 : 73-74).

Autrement dit, les traits dénotatifs peuvent être aussi bien objectifs que subjectifs, en fonction de l'implication ou la non-implication du locuteur dans la description.

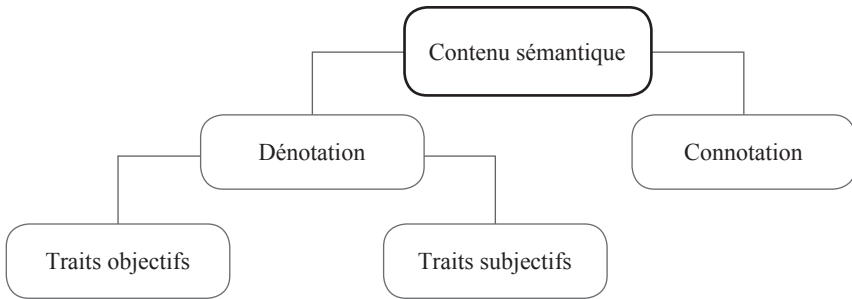
Les traits subjectifs (dans notre cas il s'agira surtout de traits évaluatifs) ne doivent pas être confondus avec la *connotation*. Catherine Kerbrat-Orecchioni en donne la définition suivante :

En linguistique et sémiologie, la connotation d'une unité, ce n'est pas sa signification (ou compréhension) globale, c'est l'ensemble des « composantes connotatives » d'un terme, c'est-à-dire certains ingrédients seulement de sa signification, et qui ne sont pas considérés comme les plus importants puisqu'on les taxe souvent de qualités additionnelles, secondes, périphériques, *etc.* Les constituants fondamentaux de la signification d'un terme sont les traits dénotatifs, ou sèmes, que dégage l'analyse compositionnelle (Kerbrat-Orecchioni, 1977 : 12).

Il est important de remarquer que la connotation n'est pas externe à la langue, mais n'est pas non plus inhérente au sens d'un lexème dans la même mesure que les traits dénotatifs ; pour que la connotation s'actualise, il faut des conditions discursives spécifiques (*ibid.* : 18). Or, la connotation peut parfois « passer » dans la dénotation. Les traits connotatifs passent dans la dénotation pour devenir des traits dénotatifs, subjectifs dans la plupart des cas (*ibid.* : 75). Lors de notre analyse nous ferons appel à la notion de connotation dans les cas où l'item lexical se charge d'une valeur sémantique supplémentaire par rapport au contenu dénotatif.

On pourra donc schématiser le contenu sémantique tel que nous le concevons pour cette recherche de la façon suivante :





**Figure 1. Schéma du contenu sémantique**

Les deux NHF étudiés étant polysémiques, le contenu sémantique sera schématisé séparément pour chaque sens dégagé.

#### LE CONTEXTE

La difficulté principale des études sémantiques sur les états de langue anciens consiste en l'absence de compétence linguistique chez le chercheur, ainsi que l'impossibilité d'interroger les éventuels locuteurs natifs. Cette idée fait écho dans les études lexicologiques et sémantiques du français médiéval. Ainsi, pour Georges Lavis, « l'étude objective du système lexical d'une langue <...> ne peut être réalisée que « de l'intérieur » » (1972 : 582), alors que Theo Venckeleer écrit :

Toujours est-il que la grande importance accordée à la mise en œuvre du vocabulaire, reflète justement le fait que les seuls témoignages linguistiques légués par le moyen âge français sont exclusivement du domaine du discours, non de celui de la langue (Venckeleer, 1975 : 9).

La méthodologie que nous avons constituée a pour objectif de pallier autant que possible cette absence de compétence linguistique.

En suivant les travaux de Georges Kleiber et de Catherine Kerbrat-Orecchioni, nous considérons que le sens n'existe pas uniquement en contexte, mais relève également du système linguistique. Mais indépendamment de la position que l'on peut prendre vis-à-vis du rôle du contexte dans le fonctionnement des unités linguistiques<sup>25</sup>, il est indéniable que la sémantique ne peut pas se passer du contexte comme outil d'analyse.

<sup>25</sup> Pour un aperçu synthétique des différentes approches du contexte, voir Kleiber 1994a (en particulier, au sujet de l'opposition entre l'approche standard du contexte et l'approche cognitive), ainsi que Kerbrat-Orecchioni 1997, 2012.

C'est d'autant plus le cas pour les études des états de langue anciens : étant donné l'absence de compétence linguistique chez le chercheur et l'impossibilité d'interroger d'éventuels locuteurs natifs, le contexte est tout ce qui reste à l'analyse pour accéder au contenu sémantique des unités. Par ailleurs, la particularité de l'étude des états anciens des langues consiste en l'impossibilité d'accéder à la matière orale. Or, l'écrit semble être encore plus « contextualisé » que l'oral, puisque « c'est le contexte intra- ou intertextuel qui domine la situation – ou la supplée » (Rastier, 1998 : 106).

Malgré le rôle déterminant de la constitution d'un corpus exhaustif, représentatif et homogène (Kleiber, 1978), le seul corpus ne suffit pas. Ce « qu'on oublie toujours dans le cas d'une langue ancienne, c'est que si on ne remonte pas jusqu'aux situations – grâce à l'ensemble des contextes, mais aussi grâce à l'histoire et à l'archéologie – on ne remonte pas toujours jusqu'au sens, même avec des contextes suffisamment nombreux. Ce sont en principe les contextes qui nous aident à reconstruire les situations, mais très souvent nous aurions besoin des situations d'abord pour éclairer les contextes<sup>26</sup> » (Mounin, 1968, cité d'après Kleiber 1978 : 44). Reconstruire les situations signifie également adopter « le point de vue des membres » de la communication (Kerbrat-Orecchioni, 2012 : 37). Cette tâche apparaît délicate même pour les états de langue modernes :

Mais cette reconstitution ne va pas de soi étant donné que l'analyste a au départ sur l'interaction un « point de vue » radicalement différent de celui des membres – il exerce son activité a posteriori, dans un lieu et une temporalité très différents, avec des objectifs et un outillage également très différents, c'est-à-dire dans un contexte très différent (*ibid.*).

Dans notre cas, le lieu et la temporalité sont en effet très différents, puisque des siècles d'histoire nous séparent des locuteurs médiévaux. L'exploitation du contexte à plusieurs niveaux apparaît donc cruciale si l'on veut accéder au contenu sémantique des dénominations qui nous intéressent.

Les hiérarchisations théoriques du contexte se multiplient. On s'accorde généralement à distinguer le micro-contexte linguistique (interne) du macro-contexte situationnel (extra-linguistique, externe...) (Kleiber, 1978, Kerbrat-Orecchioni, 2012). Nous adopterons cette position globale en séparant les deux types de contextes en plusieurs niveaux.

---

<sup>26</sup> Par « contexte » Mounin entend sans doute le micro-contexte linguistique.

### a. *Micro-contexte linguistique*

#### *Co-texte linguistique*

Ce contexte, encore appelé cotexte, est relatif à un mot ou à une séquence <...>, dont il constitue l'environnement immédiat (Franckel, 2006 : 51).

Le cotexte, c'est le matériel sémiotique qui précède et suit la séquence soumise à l'analyse, au sein de la même macro-unité <...>. (Kerbrat-Orecchioni, 2012 : § 51)

Pour désigner cet environnement immédiat, dans un but distinctif des autres niveaux du contexte, nous parlerons de co-texte linguistique.

#### *Le texte*

Le texte global, qui au cours de l'analyse est décomposé en segments successifs, c'est l'ensemble du matériel sémiotique (verbal, paraverbal et non-verbal) produit par les participants « actifs » (Kerbrat-Orecchioni 2012 : § 49).

Il s'agit du niveau « linguistique » que nous prendrons en considération sous le prisme discursif. Il permet de rendre compte de la variation de la nature et de la fréquence des emplois dégagés par l'analyse du co-texte en fonction du moment chronologique, mais aussi du type de texte.

#### *Le type de discours*

Les quatre genres textuels qui composent notre corpus font apparaître quatre types de discours qui leur sont propres. Afin d'intégrer cette notion dans notre analyse, nous opérerons le concept de l'univers de discours qui se définit comme une somme de la situation de communication et des contraintes stylistico-thématiques (Kerbrat-Orecchioni, 2002 [1980] : 13). Contrairement au modèle de la situation de communication de Rastier (1989 : 16), l'univers de discours est donc présenté comme étant « extérieur » à l'émetteur et au récepteur qui influence les choix des moyens d'expression du premier et la compréhension du message par le deuxième<sup>27</sup>.

---

<sup>27</sup> L'univers de discours propre à tel ou tel genre correspond partiellement à ce que Georges Lavis, en parlant des chansons courtoises, appelle « le réseau de « contraintes » sémantiques et syntaxiques qui canalisent en quelque sorte la liberté créatrice du poète » (1972 : 590).

## b. *Macro-contexte*

### *Le genre : point de vue narratif*

Plusieurs études sémantiques et lexicologiques du français médiéval choisissent comme critère de constitution du corpus l'unité du genre textuel (Lavis, 1972, Venckeleer, 1975, Kleiber, 1978). D'une certaine façon, le genre correspond à la *situation* :

Dans les deux cas, le contexte envisagé peut être plus ou moins étroit (microcontexte) ou large (macrocontexte). On appelle généralement « situation » le contexte externe étroit (setting, cadre spatio-temporel généralement immédiatement accessible aux participants à l'interaction), par opposition au macrocontexte social ou culturel (Kerbrat-Orecchioni, 2012 : § 31).

On peut considérer que dans le cas des études des langues anciennes pour lesquelles seuls les textes écrits sont disponibles, le genre pose le cadre thématique, mais aussi, indirectement, le cadre spatio-temporel du texte.

L'accès aux textes seulement pose des problèmes théoriques considérables. Ainsi, Georges Lavis, remarque la nécessité de la prise de conscience par rapport à la spécificité des expressions linguistiques accessibles via les textes disponibles qui ne donnent pas accès directement au système linguistique :

Pour toute étude lexicologique portant sur l'ancien français, les énoncés, les textes grâce auxquels sont analysées les unités lexicales, portent la marque de caractéristiques (linguistiques, socio-culturelles...) qui sont propres au corpus envisagé, qui ne sont donc pas pertinentes au niveau de la langue, mais qu'il n'est pas toujours facile d'identifier avec certitude en tant que telles. Aussi pensons-nous qu'on ne peut espérer atteindre le lexique de l'ancien français qu'au terme d'une série d'études de vocabulaires ayant pour objet des idiolectes déterminés, et que par un effort de généralisation procédant par comparaisons et déductions successives (1972 : 582).

Nous allons considérer les genres textuels qui constituent notre corpus de plusieurs points de vue. Le point de vue pragmatique tout d'abord, en essayant de comprendre quel était le but communicatif des émetteurs de textes et qui en étaient les destinataires<sup>28</sup> (auditeurs ou lecteurs). Le point

---

<sup>28</sup> Ainsi, il est important de prendre en considération la particularité suivante : la majeure partie des écrits médiévaux « aboutit presque toujours à son interprétation devant

de vue stylistique également, qui concernera non seulement les textes de fiction, mais aussi les textes documentaires, qui, à leur façon, sont tout aussi marqués par les formules que les chansons de geste, dans la mesure où ils reproduisent des modèles syntaxiques type où seuls les données objectives, comme les Noms Propres ou les chiffres, peuvent varier.

Pour les textes de fiction nous ferons également appel à la perspective littéraire qui dégage les motifs et procédés caractéristiques de tel ou tel genre. La perspective mémorielle de la linguistique cognitive est également intéressante si l'on parle du genre textuel. Ainsi, dans l'analyse du discours on parle de mémoire discursive qui se définit comme : « l'ensemble des savoirs consciemment partagés par les interlocuteurs » (Berrendonner, 1983 : 231). Dans le cas des deux genres de fiction de notre corpus – les chansons de geste et les romans – on peut appliquer le concept de mémoire discursive non pas à un seul texte uniquement, mais au genre en entier. En effet, les locuteurs qui produisent les œuvres et ceux à qui ils sont destinés ont vraisemblablement conscience de l'appartenance de chaque texte à une « tradition épique » ou une « tradition romanesque ». Ces traditions ont leurs motifs, leurs procédés narratifs et leurs héros qui sont des repères conceptuels que l'on n'a pas besoin de rappeler. En parlant des prologues des chansons de geste, Jean Rychner écrit :

Le jongleur place aussitôt son auditoire dans un milieu légendaire qui lui est familier et qui est un bon terrain de rencontre : une sorte d'entente, presque de complicité, va unir jongleur et public face au héros. Le contact est assuré, le chant particulier se trouve lié à la plus large légende (Rychner, 1955 : 55).

La réflexion de Sophie Marnette va dans le même sens :

Ce qui compte dans le récit des chansons de geste et de la chantefable, ce n'est pas d'inventer une histoire nouvelle, ni même de donner un sens nouveau à des sources préexistantes en les réorganisant. En effet, tout le monde sait que Ganelon va trahir Charlemagne ou que Guillaume va conquérir Orange mais il faut le raconter de façon originale, qui plaise à un public particulier (Marnette, 1998 : 218-219).

---

le public » (Aurell, 2011 : 117). Même lorsqu'il s'agit de genres destinés à la lecture, il faut garder à l'esprit qu'« au Moyen Âge, tout comme dans l'Antiquité, on avait l'habitude de lire, à haute voix ou du moins avec une phonation minimale, tout ce qui avait été écrit » (Koch, 1993 : 50). Cela implique, en particulier, un lien plus direct entre l'émetteur (qui est représenté par plusieurs instances : « auteur(s) » du texte – copiste(s) – jongleur(s) ») et le récepteur que dans le cas d'une écriture sans destinataire aussi précis. Ce lien direct permet (et exige, en raison des modalités de la réception orale) à son tour, de faire appel à des univers de discours familiers pour les récepteurs.

En effet, la littérature médiévale « possède un caractère strictement codifié » où « la présence de tel ou tel personnage véhicule telle ou telle situation, tel ou tel style, tel ou tel motif, et *vice versa* » (Solovieva, 2001 : 307). Tout comme les textes épiques avec leurs héros (Roland, Guillaume *etc.*) et anti-héros (Ganelon *etc.*), les romans arthuriens ont plusieurs figures de chevaliers modèles, chacun à sa façon, comme Lancelot, Perceval ou Galaad auxquels on fait allusion même dans les cas où ils ne participent pas à la narration. L'auditeur ou le lecteur médiéval, tout en attendant d'être divertie par une nouvelle histoire, sait généralement à quoi s'attendre grâce à cette mémoire discursive propre à chaque genre. Plus encore, l'appartenance à telle ou telle tradition assure la validité du texte, puisque « la vérité médiévale n'est rien sans la reconnaissance de la tradition » (Corbellari, 2007 : 20).

Par ailleurs, lors de l'analyse des occurrences nous relèverons certaines situations narratives récurrentes propres à tel ou tel genre textuel.

Enfin, chaque genre doit être considéré dans son évolution en concordance également avec l'évolution de la société, ce qui nous permet de passer au dernier niveau – le plus large – le contexte socio-historique.

### *Contexte socio-historique*

L'idée d'intégrer l'histoire et les autres sciences sociales dans les études linguistiques, et sémantiques en particulier, n'est pas nouvelle. Eugenio Coseriu considère la langue comme un objet historique et social :

La lengua es un hecho social, en el sentido más genuino del término “social”, que es el de “propriadamente humano”. Pero, <...>, la lengua no es simplemente un hecho social “entre otros” e “como los otros” (como los sistemas monetarios, por ej.), pues el lenguaje es el fundamento mismo de todo lo social, ... » (1973 [1958] : 24).

Il est important de ne pas oublier cette double dimension sociale de l'essence même de la langue dans toutes les branches des études linguistiques. François Rastier écrit :

Pour décrire la richesse des relations contextuelles, la linguistique ne peut rester dans l'espace douillet mais confiné de la phrase ; elle s'ouvre aux textes, et par là aux cultures et à l'histoire, en réaffirmant son statut de science sociale (et non formelle). Elle va au contact des disciplines voisines, c'est-à-dire de toutes les autres sciences sociales, et peut même exercer un droit de suite sur leur territoire (1989 : 7).

Il paraît évident, dans une étude de langue ancienne, de se tourner vers les travaux d'historiens. Logiquement, les textes anciens évoquent

des réalités anciennes dont une connaissance profane n'est pas suffisante dans une recherche scientifique. Une compréhension générale de la situation politique et économique de la société dont on étudie la langue, ainsi que des processus et évolutions qui se passent au sein de cette société, apparaît nécessaire.

Il convient, cependant, de rester prudent et de ne pas assimiler la réalité historique au sens linguistique. Nous sommes tout à fait d'accord avec Georges Matoré qui écrit : « En réalité, les mots n'expriment pas les choses, mais la conscience que les hommes en ont » (1973 : 43). Georges Kleiber indique l'inutilité des études en psychologie pour l'étude sémantique du substantif *ire* (premier sens : « colère ») :

La lexicologie ne gagne pas à une progression de la science en ce domaine <la psychologie> : ce n'est pas parce que l'on connaîtra mieux le phénomène de la colère que l'on pourra mener une étude sémantique plus rigoureuse dans ce secteur du lexique, tant les critères de structuration linguistique diffèrent de la science psychologique (Kleiber, 1978 : 31).

En effet, le fonctionnement psychologique et physiologique de la colère ne correspond pas à la « conscience » que les locuteurs ont de cet item. C'est là que nous devons avoir recours à la distinction entre le référent actuel et la référence virtuelle qui est particulièrement importante dans le cas des dénominations sociales, puisque tout individu, soit référent actuel, peut avoir une multitude de rôles sociaux dont une partie seulement est exprimée dans la référence virtuelle du NH qui le désigne dans telle ou telle situation.

#### 4. DESCRIPTIF DU PROCESSUS DE L'ANALYSE

L'efficacité de la méthodologie réunissant l'analyse qualitative et quantitative des occurrences avec une attention particulière à leur distribution a été démontrée pour l'étude des NH (voir Schnedecker & Mihatsch, 2018, en particulier, les contributions de Aleksandrova & Schnedecker, Oberlé, Baider & Todirascu). Notre analyse se base sur un relevé exhaustif des occurrences de *baron* et *chevalier* dans tous les textes de notre corpus. Chaque occurrence relevée a été traitée plusieurs fois. Tout d'abord, au cours de la lecture du texte en entier, en suivant en cela Françoise Denis qui préconise de « placer le mot non seulement dans le contexte de 50-100 vers, mais dans l'intégralité » du texte (1989 : 7). Cette étape nous a permis d'esquisser les premières tendances générales et de placer les occurrences dans les situations narratives correspondantes. Pendant la deuxième étape,

nous avons réexaminé chaque occurrence dans un contexte restreint de façon plus détaillée. Cette étape passait par l'analyse

- des co-occurrences directes des dénominations au sein des SN
- des réalisations au sein des énumérations
- des réalisations au sein des diverses constructions syntaxiques dont
  - a. celles qui n'ont pas d'effet stylistique recherché
  - b. à effet stylistique recherché
- des items co-référents dans le co-texte
- de tout autre emploi particulier propre à tel ou tel sens des items étudiés

Enfin, un certain nombre d'occurrences qui suscitaient le plus d'inquisitions ont fait l'objet d'une troisième, voire d'une quatrième consultation.

L'analyse a permis de dégager pour chaque NHF au sein de chaque genre un certain nombre d'emplois. La notion d'« emploi » correspond approximativement à ce que Georges Lavis appelle « les possibilités sémantico-syntaxiques » (1972) d'une unité lexicale dans la mesure où les emplois représentent des séquences au sein desquelles les occurrences des items étudiés sont relevées. Or, l'emploi tel que nous le concevons dépasse les limites de la définition syntaxique, puisque plusieurs éléments peuvent s'y ajouter, comme les valeurs référentielles des co-occurrences directes des items étudiés ou la portée stylistique et idéologique des constructions syntaxiques les contenant. Il convient de dire qu'au début de la recherche nous n'avions pas de schéma préconstruit quant aux emplois que nous voulions mettre en évidence : c'est l'analyse qui a permis de dégager les critères pertinents et, par la suite, les emplois même.

Les atouts de la linguistique numérique qui se base sur le traitement automatique des textes sont de nos jours incontestables, y compris dans les études diachroniques<sup>29</sup>. Sa pertinence est reconnue non seulement par les linguistes. Ainsi, l'historien Alain Guérreau, plaide la cause de l'analyse statistique lexicale appliquée aux études historiques (par exemple, 1989<sup>30</sup>).

Dans le cas de notre recherche nous nous sommes limitée aux calculs des fréquences absolues et des fréquences relatives pour les NHF étudiées et leurs emplois. Les fréquences absolues sont calculées pour le nombre total d'exemples des deux NHF, ainsi que le nombre total d'occurrences

---

<sup>29</sup> Pour l'application des méthodes de la linguistique numérique aux études diachroniques, voir en particulier Guillot *et al.* 2008 ; Souvay 2009 ; Diwersy *et al.* 2017.

<sup>30</sup> Guérreau reconnaît la pertinence des statistiques lexicales pour les études historiques en ce qui concerne non seulement l'attribution et la datation des textes anciens, mais également les « structures sémantiques » qu'elle permet de dégager (voir en particulier Guérreau, 1989 : 94-98).



de chaque emploi. Les fréquences relatives sont calculées pour obtenir la proportion des occurrences de tel ou tel emploi par rapport au nombre total d'occurrences<sup>31</sup>, ainsi que la proportion d'occurrences de *baron* par rapport à celles de *chevalier*. Par ailleurs, nous calculons la répartition des occurrences par type référentiel.

Pour chaque NHF au sein de chaque texte nous allons partir de la fréquence absolue des occurrences (=FrO), ainsi que la fréquence absolue des occurrences d'un emploi (=FrE). Il faut dire qu'une seule occurrence peut relever de plusieurs emplois en même temps. Ainsi, nous considérons que dans l'occurrence suivante :

XII CharroiM 337  
 «Entendez moi, nobile *chevalier*,  
 De Looÿs, mon segnor droiturier,  
 Comme est gariz qui le sert volentiers.

*chevalier* fait partie d'une forme nominale d'adresse (emploi « forme nominale d'adresse ») et se trouve en co-occurrence avec l'adjectif *noble* au sein d'un SN (emploi « co-occurrence avec un élément évaluatif »). On comptera donc dans ce cas-là une occurrence, mais deux emplois. Enfin, la fréquence relative d'un emploi sera calculée par rapport au nombre total d'occurrences :

$$\text{FrRel} = \text{FrE} / \text{FrO}$$

Toutes les fréquences seront données pour tous les textes d'un seul genre dans les limites d'un seul siècle. La fréquence relative sera présentée sous forme de pourcentage. Les tableaux statistiques complets seront placés dans l'*Annexe*, alors que dans le corps du texte nous en présenterons des extraits en parlant de tel ou tel emploi. Dans les tableaux apparaissant dans le corps du texte, nous allons répartir les occurrences par type référentiel.

Enfin, comme nous l'avons dit précédemment, si le texte original de la thèse contenait un descriptif quasi-exhaustif des emplois, le présent ouvrage choisit de présenter un nombre limité d'emplois que nous considérons comme les plus représentatifs pour les types de discours respectifs

---

<sup>31</sup> Cette étape aurait pu être effectuée de façon automatique par le biais du calcul des co-occurrences dans le cas où nous aurions travaillé sur des éditions électroniques lemmatisées de façon uniforme. Or, comme notre corpus a été constitué à partir d'éditions papier et de plusieurs bases électroniques avec des normes d'édition différentes qui impliquent ou non la lemmatisation, ce n'était pas possible.

et la mise en évidence du contenu sémantique des items dans leur globalité. Les emplois sélectionnés réunissent pour la plupart deux conditions : fréquence relative considérable à une des étapes de la période étudiée et pertinence pour les tendances générales relevées. En revanche, nous placerons en *Annexe* les tableaux statistiques représentant les données sur la quasi-totalité des emplois de *baron* et *chevalier*.