



HAL
open science

Entre contrainte et choix : Usages, normes et représentations en syntaxe dans l'histoire du français

Julie Glikman

► **To cite this version:**

Julie Glikman. Entre contrainte et choix : Usages, normes et représentations en syntaxe dans l'histoire du français : Volume 1 : mémoire de synthèse. Linguistique. Université Savoie Mont Blanc, 2023. tel-03988853

HAL Id: tel-03988853

<https://shs.hal.science/tel-03988853>

Submitted on 2 Mar 2023

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

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

Manuscrit HDR

Pour obtenir le diplôme d'

HABILITATION à DIRIGER des RECHERCHES de l'UNIVERSITÉ SAVOIE MONT BLANC

Spécialité : Sciences du langage

Arrêtés ministériels des 23 novembre 1988 et 13 février 1992

Présenté par **Julie GLIKMAN**

Entre contrainte et choix : Usages, normes et représen- tations en syntaxe dans l'his- toire du français

Volume 1 : mémoire de synthèse

HDR soutenue publiquement le **09 février 2023**,
devant le jury composé de :

Madame Dominique LAGORGETTE

Professeure, Université Savoie Mont Blanc, Garante

Madame Lene SCHØSLER

Professeure émérite, Université de Copenhague, Rapporteuse

Madame Agnès STEUCKARDT

Professeure, Université de Montpellier 3, Rapporteuse

Monsieur Gilles SIOUFFI

Professeur, Sorbonne Université, Rapporteur

Monsieur Sylvain KAHANE

Professeur, Université Paris Nanterre, Président

Madame Céline GUILLOT-BARBANCE

Maitresse de Conférence HDR, ENS Lyon, Membre

Monsieur Andreas DUFTER

Professeur, Université de Munich, Membre

Remerciements

Par tradition, les remerciements constituent les premières lignes lues par les lecteurs, bien que ce soit celles écrites en dernier. C'est effectivement au moment de terminer ce mémoire de synthèse que je rédige ces quelques lignes afin de remercier toutes les personnes qui m'ont aidée dans cette démarche.

En premier lieu, ma gratitude va à Dominique Lagorgette pour avoir accepté d'être ma garante. Merci pour ton suivi, ton soutien, toutes les discussions stimulantes que nous avons eues lors de nos échanges, toutes les pistes que tu m'as données et que je vais continuer à explorer, et, je l'espère, toutes les discussions que nous aurons encore. Tout au long du parcours, je n'ai jamais regretté de t'avoir demandé de m'encadrer. Merci encore à toi d'avoir relevé le défi.

Je remercie également Céline Guillot-Barbance, Andreas Dufter et Sylvain Kahane d'avoir accepté de siéger au jury, ainsi que Lene Schøsler, Agnès Steuckardt et Gilles Siouffi d'avoir bien voulu faire en outre les pré-rapports en vue de la soutenance. Merci pour votre présence à cette soutenance. Merci également pour les échanges que nous avons eus jusqu'ici et que, je l'espère, nous aurons encore.

Merci aussi à Annie Bertin, sans qui rien de tout cela n'aurait commencé, qui m'a de nouveau conseillée durant la rédaction. Merci pour nos échanges et merci encore de m'avoir fait entrer dans ce drôle de monde en m'acceptant en thèse alors. Je remercie également Thomas Verjans pour nos nombreuses discussions et pour avoir bien voulu relire l'intégralité de ce mémoire. J'espère que cela nous donnera l'occasion de futurs projets.

Un parcours scientifique ne se fait pas sans rencontres qui l'enrichissent. Bien que j'y revienne dans l'introduction, je voudrais tout de même remercier ici tous les collègues et amis qui ont contribué, à un moment ou un autre, à nourrir mes réflexions : Mathieu Avanzi, Vincent Balnat, Hava Bat Zeev Shyldkrot, Sylvie Bazin-Tacchella, Marie-José Béguelin, Christophe Benzitoun, Bérengère Bouard, Eva Buchi, Bernard Combettes, Gilles Corminboeuf, Annie Delaveau, Marianne Desmet, Laura-Maï Dourdy, Camille Fauth, Françoise Gadet, Antoine Gautier, Yan Greub, Sybille Grosse, Pierre Larrivée, Leda Mansour, Chrisiane Marchello-Nizia, Nicolas Mazziotta, Gabriella Parussa, Julie Peuvergne, Sophie Prévost, Stefan Schneider, Achim Stein, Dan Van Raemdonck, Isabelle Weil, et tous les collègues que j'aurais oubliés, de Nanterre, du Lattice, de Nancy ou d'ailleurs. Je remercie bien sûr tout particulièrement mes collègues de Strasbourg qui me supportent maintenant depuis dix ans toujours dans la bonne humeur.

Remerciements

Merci également à Georges, Françoise et Julien, en particulier pour leur relecture attentive des dernières coquilles et autres tics de langage notamment.

La préparation d'une habilitation à diriger des recherches se fait toujours en parallèle d'autres activités. J'ai eu la chance de pouvoir finaliser cette habilitation dans des conditions idéales, grâce à mon CRCT et à ma délégation CNRS, qui m'ont permis de me consacrer uniquement à mes activités de recherches. Je remercie encore le laboratoire ATILF et Rudolph Sock, directeur de LiLPa, d'avoir bien voulu soutenir ma demande de délégation. Malgré cela, le reste des activités est toujours présent, je tiens donc à remercier ici tout particulièrement Hélène Vassiliadou d'avoir bien voulu prendre davantage sur elle la charge de travail institutionnel de la direction de l'Institut des dernières semaines, de même que Stefan Schneider pour *Studia Linguistica Romanica*, promis, c'est à charge de revanche !

La « vraie » vie continue également, mes remerciements vont ainsi à Nelly d'avoir tout pris en charge pour la maison et pour le reste, et de croire en moi depuis tellement d'années. Merci à Pierre-Louis, Arnaud, Caro, Aline, Gwen, Leda, Marie, à tous les amis pour votre confiance. Merci également à Nicolas de ton soutien tout au long du processus, je sais la chance que j'ai.

Table des matières

Remerciements	2
Table des matières	4
Préambule.....	5
1. Introduction.....	8
1.1. Pas après pas, au fil des rencontres et hasards	8
1.2. Briques après briques, construction de mon identité scientifique.....	17
1.3. Choix et certitudes, structure du mémoire de synthèse	24
2. Établir les variantes. Autour des objets : une approche syntaxique	27
2.1. Établir les variantes en syntaxe et morphosyntaxe.....	28
2.2. Autour du verbe : établir le lien syntaxique.....	34
2.2.1. Qu'est-ce que le lien syntaxique ?.....	35
2.2.2. Comment établir le lien syntaxique ?	42
2.2.3. Questions, applications et perspectives	67
2.3. Autour d'autres zones de variables : établir les variantes en concurrence	77
2.3.1. Une zone de variable à variantes implicites.....	79
2.3.2. Une zone de variable à variantes volontairement restreintes	81
2.3.3. Une zone de variable à variantes explicites	85
2.3.4. Le texte comme variable.....	88
3. Analyser la variation. Autour des données et des méthodes : une approche en linguistique variationnelle.....	93
3.1. Quels paramètres pour la variation ?.....	94
3.1.1. Quelques éléments de définition	94
3.1.2. Quels paramètres retenir ?.....	97
3.1.3. Bilan intermédiaire	111
3.2. Quelles données pour analyser la variation ?.....	112
4. Comprendre le changement linguistique. Autour des usages et des normes : une approche centrée sur le locuteur.....	130
4.1. La caractérisation des discours	130
4.2. Usages, normes et représentations sur la langue.....	137
4.3. Qu'est-ce que cela nous apprend sur le changement linguistique ?.....	149
En guise d'ouverture : transmission et diffusion	154
Bibliographie	157
Annexe 1 : <i>Curriculum Vitae</i>	169
Annexe 2 : Table du volume 2 (volume des travaux).....	186

Préambule¹

*J'ai vu de la lumière, je me suis dit
qu'est-ce qui s' passe là-haut*

*Ärsenik, Feat. Passi, Par où t'es
rentré, je t'ai pas vu sortir, 1998.*

Quand je pense à ma carrière, ou plutôt à mon début de carrière, et plus exactement à ce moment, délicat parfois, qui constitue le passage entre le cycle des études et celui du début de carrière, dont le doctorat constitue à la fois le lien, le liant et le rite de passage (je me souviens encore de mon professeur Sylvain Kahane me disant que « maintenant on peut se tutoyer » vu que j'étais passée « de l'autre côté » une fois mon inscription faite), la première image qui me vient à l'esprit est « j'ai vu de la lumière alors je suis entrée ». À la fois suite de hasards, tout en ayant des échos au plus loin de mon enfance, rien ne me prédestinait en effet à choisir la carrière d'enseignant-chercheur en linguistique diachronique. Entre déterminisme social et erreur de plan de carrière (« ma fille, tu feras Polytechnique » avait coutume de dire mon père, probablement en guise de boutade, ou au contraire en prophète inspiré, à une époque où les filles étaient encore peu représentées à Polytechnique², et parfois suivi de « et tu seras chercheur(e ?) au CNRS »), pour moi, enfant, les seuls docteurs étaient les médecins, comme mon père, et ma mère, expert-comptable puis directeur financier (pardon pour la féminisation mais c'est ainsi dans mes souvenirs), faisait partie de ces gens dont le métier consiste en « travailler dans un bureau ». Je suis aujourd'hui encore surprise d'entendre des étudiants de première année de licence annoncer déjà leur envie de « faire une thèse », n'ayant moi-même découvert qu'en Master l'existence de ce « diplôme » si particulier.

J'ai pourtant depuis toujours été intéressée par les phénomènes langagiers. Ma grand-mère, originaire de la Provence, mais parisienne de fait (et probablement de cœur aussi), avait par exemple cette fameuse particularité, remarquée par tous, de ne faire ressortir complètement son accent méridional qu'au moment où elle s'énervait (notamment au volant, il faut bien le dire). On avait aussi le plaisir d'user de termes comme *péguer*, et bien sûr d'entendre (avec envie parfois) les cousins du sud. En vacances en Suisse alémanique, je m'étais également découvert une passion pour le (suisse) allemand (que j'identifiais alors comme de l'allemand), ce qui m'a

¹ Ce mémoire est rédigé selon la nouvelle orthographe.

² La première admission de fille à polytechnique remonte à 1972, et passe la barre des 10% en 1986. Source : <https://www.polytechnique.edu/bibliotheque/fr/40e-anniversaire-de-l%E2%80%99admission-des-filles-%C3%A0-l%E2%80%99ecole-polytechnique-en-1972>, consulté le 01.06.2021.

Préambule

fait insister auprès de mes parents pour avoir le droit de choisir allemand en première langue au collège (autant dire que, mon père ayant vécu l'Occupation, je n'y étais pas poussée, contrairement à une grande partie de mes compagnons, avec l'espoir d'intégrer de « bonnes » classes).

La découverte d'une langue à déclinaison dès mes 10-11 ans, renforcée par la suite par l'inscription en cours de grec ancien (à l'époque où l'on pouvait encore prendre un cours de langue ancienne dès la 4^e, soit vers 12-13 ans), m'a conduite très tôt à un émerveillement depuis sans cesse renouvelé : la compréhension du fonctionnement d'une langue. Si je n'ai jamais été très douée pour l'apprentissage par cœur de listes et de tableaux..., convaincue dès mon premier échange en Allemagne, à 12 ans, qu'en matière d'apprentissage des langues, rien ne vaut l'immersion, la compréhension des mécanismes de formation, assemblage et construction m'a toujours passionnée.

C'est finalement assez tard dans mes études que j'ai compris que l'étude du fonctionnement des langues et du langage pouvait faire l'objet d'un cursus en soi, et non seulement de quelques cours épars dans une formation de lettres. J'ai donc rejoint avec enthousiasme (et un certain soulagement) une formation en Sciences du langage à partir de la maîtrise, sans trop me préoccuper de l'avenir, mais parce que cela me plaisait. Quand Annie Bertin, que je ne remercierai jamais assez, m'a ensuite proposé de poursuivre en DEA, puis en thèse, tout en me prévenant dès le départ des difficultés éventuelles, j'ai poursuivi de la même manière, « on verra bien ».

La présentation de mon dossier en vue de l'obtention de l'habilitation à diriger des recherches, acte volontaire de franchissement d'une étape supplémentaire dans ma carrière, se pose aujourd'hui comme un écho à cette entrée dans la carrière d'enseignant-chercheur, puisqu'après tout, comme dirait un ancien président de la République, « il y a de la lumière »³, à quoi je réponds : « ça m'a bien plu, alors je suis restée ».

³ Déclaration de Nicolas Sarkozy, alors président de la République, sur une stratégie nationale de recherche et d'innovation, à Paris le 22 janvier 2009. Contexte de l'extrait : « Plus de chercheurs statutaires, moins de publications et pardon, je ne veux pas être désagréable, à budget comparable, un chercheur français publie de 30 à 50% en moins qu'un chercheur britannique dans certains secteurs. Évidemment, si l'on ne veut pas voir cela, je vous remercie d'être venu, il y a de la lumière, c'est chauffé... ». Source : <https://www.elysee.fr/nicolas-sarkozy/2009/01/22/declaration-de-m-nicolas-sarkozy-president-de-la-republique-sur-une-strategie-nationale-de-recherche-et-dinnovation-a-paris-le-22-janvier-2009>, consulté le 15.09.2021.

1. Introduction

1. Introduction

Avant d'entrer dans les considérations proprement scientifiques, je souhaiterais profiter de cette introduction pour revenir sur mon parcours et les différentes rencontres que j'ai eu la chance de faire au fil du temps (→ 1.1.), car, outre le « commerce des livres » comme disait Montaigne, je suis persuadée que le « commerce des autres » façonne et oriente tout autant la structuration de la pensée. Dans mon cas, il a fortement contribué à la définition de mon champ de recherche et à la construction de mon identité scientifique, que je définirai brièvement dès cette introduction (→ 1.2.), avant de présenter l'organisation de ce mémoire de synthèse (→ 1.3.).

1.1. Pas après pas, au fil des rencontres et hasards

*On choisit ses copains mais
rarement sa famille*

Renaud, *Mon beauf*, 1986.

Après l'obtention d'un contrat doctoral à l'Université de Paris Nanterre (2005-08) où j'avais fait mes études, j'ai effectué deux années en tant qu'ATER, d'abord à l'Université de Nanterre (2008-09, département de Sciences du langage, UFR LLPhi), puis à l'Université de Savoie (site de Chambéry, département des Lettres, UFR LLSH, 2009-10), université vers laquelle je reviens pour soutenir mon HDR. J'ai ensuite été recrutée comme chercheur post-doctorant au Lattice sur le projet ANR-DFG SRCMF (2010-12). À la suite de ces expériences, j'ai obtenu en 2012 le poste de MCF en linguistique diachronique française à l'Université de Strasbourg, poste que j'occupe encore actuellement.

On ne choisit pas sa famille, mais on ne choisit pas non plus toujours l'université où l'on fait ses études, en particulier quand on est une jeune lycéenne de région parisienne et qu'il n'y a que deux universités dans la région qui proposent le cursus de Lettres Classiques. C'est ainsi un peu par hasard que je me suis retrouvée inscrite à Nanterre, premier hasard que je n'ai jamais regretté par la suite (à part peut-être pour les fameux trajets en RER, avec changement à Châtelet-Les Halles, qui constituent une certaine forme de torture moderne, en accord avec la tradition des lieux). Après ma licence et quelques cours d'introduction à la linguistique, c'est surtout à partir des années de deuxième cycle (d'abord à travers la préparation au CAPES où l'enseignement en langue et celui de l'ancien français étaient renforcés, puis en maîtrise et DEA de Sciences du langage) que j'ai développé mon socle de connaissances en linguistique et que se sont ouverts à moi de nouvelles perspectives et tout un horizon de recherches. J'ai eu de nombreux excellents enseignants et enseignements à Nanterre, que je ne citerai pas tous ici, ne

présentant que ceux qui me semblent, dans cette démarche réflexive que constitue le mémoire de synthèse, avoir plus directement forgé mon caractère scientifique, mais je n'oublie pas pour autant ce que je dois à tous.

Ma sensibilisation à l'approche sociolinguistique est venue très tôt, grâce à ma formation en maîtrise où j'ai eu la chance de suivre les cours de Françoise Gadet, qui m'ont formée aux études laboviennes, à la question de la posture de l'observateur, à la variation d'une manière générale. De même, Bernard Laks mentionnait fréquemment les phénomènes d'hypercorrection à propos de la production fautive de la liaison. Si ces notions n'ont pas tout de suite été réexploitées dans mes premiers travaux, elles étaient présentes en filigrane. On n'oubliera pas que les travaux de Labov sur la variation ont largement contribué à la revalorisation des approches diachroniques. Mes travaux sur les rapports oral - écrit, sur le français régional, sur la micro-diachronie contemporaine, sur le recueil de données et sur les représentations linguistiques y plongent largement leurs racines. Je suis également convaincue des apports de la sociolinguistique historique, qui a entre autres avantages celui de remettre au premier plan la question des méthodes et des sources.

La question de la syntaxe et de la subordination, au centre d'une grande partie de mes travaux, tient également beaucoup à ma formation à Nanterre, où j'ai pu suivre les cours de Danielle Leeman, Marcel Cori, Sylvain Kahane ainsi que Marie-Laure Elalouf. Le cours de S. Kahane sur les différents courants d'analyse syntaxique et la conférence d'Igor Melčuk que j'ai eu la chance d'écouter lors d'un séminaire, loin de me cloisonner à un modèle, m'ont confortée dans l'idée que ce sont les modèles qui doivent s'adapter aux données et non l'inverse. Les travaux ultérieurs de Sylvain Kahane et Nicolas Mazziotta ont notamment montré la transposition possible des différentes formalisations (Kahane & Mazziotta 2015). Toujours est-il que j'ai d'abord été formée dans une approche en constituant (relativement flexible), principalement à partir des grammaires de référence comme la *Grammaire Méthodique du Français* (Riegel, Pellat, Rioul 1994). Après une scolarité traditionnelle française, la découverte d'une analyse en « GN-GV » et des règles de réécriture m'a paru alors comme une véritable révolution. Cela flattait en outre mon penchant naturel pour les analyses formelles de type mathématique, qui avaient toujours eu ma préférence, mais là encore, que je n'avais pas choisi d'étudier, un peu comme une erreur de choix de carrière. C'est probablement ce qui m'a fait vouloir faire de la recherche en syntaxe dès le départ : le caractère logique et formaliste, bien que je n'aie finalement même pas proposé de formalisation dans ma thèse, comme me l'avait alors reproché Sylvain Kahane, membre du jury, lors de la soutenance. C'est dans un second temps que j'ai été formée à l'analyse dépendancielle, même si j'avais auparavant déjà lu les

travaux de Tesnière. C'est en effet l'approche qui avait été adoptée dans le projet Syntactic Reference Corpus of Medieval French (SRCMF)⁴ au sein duquel j'ai fait mon post-doctorat. Ce projet consistait en l'annotation syntaxique systématique et manuelle d'un corpus de texte d'ancien français. J'ai rejoint le projet après le choix du modèle d'annotation principal, mais j'ai eu l'occasion de participer à la plupart des discussions ultérieures sur les choix d'annotation. J'ai conservé depuis une approche de type dépendanciel dans mes travaux, à laquelle j'adhère du point de vue théorique, même si j'ai parfois recours à la notion de « groupe » ou « syntagme ». Cette approche est également à la base du schéma d'annotation adopté pour un certain nombre de corpus oraux, comme le projet Rhapsodie, et a donné lieu à la création d'un standard à vocation internationale, le Surface Syntactic Universal Dependencies (SUD)⁵. Pour le projet mené dans le cadre de ma délégation CNRS (2021-2023), qui comprend la constitution d'un corpus de sms vocaux annoté en syntaxe, c'est également ce modèle qui a été retenu. Cette décision relève à la fois d'une conviction théorique, et d'un choix pratique visant à s'aligner sur les corpus existants pour permettre leur comparaison et leur mise en commun.

Cet intérêt pour la syntaxe et la subordination, pour reprendre le cours chronologique des choses, a pu trouver un terreau favorable à Nanterre grâce au groupe de travail sur la syntaxe auquel j'ai participé et que j'ai un temps co-animé, avec Sylvain Kahane, Marcel Cori et d'autres, grâce aussi au groupe de travail sur les locutions dirigé par Danielle Leeman qui m'y a invitée à plusieurs reprises (collaboration qui a notamment donné lieu à la publication Hassler, Hümmel et Glikman 2008, certainement la plus éloignée de mes champs de recherche habituels), et surtout, « last but not least » comme on dit en bon français, grâce au groupe de travail sur les locutions conjonctives et la subordination en diachronie, animé par Annie Bertin pendant et après mes années de thèse. Ce groupe de travail a véritablement joué un rôle majeur dans ma carrière, pour ma formation intellectuelle bien sûr, mais aussi parce qu'il m'a permis très tôt de rencontrer et d'échanger de manière privilégiée avec d'éminents spécialistes comme avec de jeunes chercheurs de ma génération, qui sont devenus autant de collègues et de précieux collaborateurs : Hava Bat-Zeev Shyldkrot, Sabine Lehmann, Thomas Verjans, avec qui j'ai récemment coédité le volume d'hommages offert à Annie Bertin (Glikman, Bat Zeev Shyldkrot,

⁴ Projet ANR-DFG 1.3.2009-29.2.2012 dirigé par Sophie Prévost et Achim Stein (Prévost, Sophie; Stein, Achim. 2013. Syntactic Reference Corpus of Medieval French (SRCMF). ENS de Lyon/ILR Stuttgart), voir <http://srcmf.org/>.

⁵ Initiative portée par Sylvain Kahane et Kim Gerdes, entre autres : <https://surfacesyntacticud.github.io/>.

Lehmann, Sitri et Verjans (éds) 2021 [x]⁶), ainsi que Bernard Combettes, Thierry Ponchon, ou encore Olivier Soutet, pour n'en citer que quelques-uns.

Enfin, il me faut évoquer la diachronie évidemment, que j'ai commencé à aborder, comme beaucoup, par les cours d'introduction à l'ancien français d'Isabelle Weill et Annie Bertin, avant de poursuivre vers les cours de diachronie à proprement parler avec les mêmes enseignantes. Annie Bertin, qui a dirigé mon DEA puis ma thèse, m'a formée à l'analyse diachronique et aux grands courants théoriques, comme la grammaticalisation, à l'analyse stricte et rigoureuse des phénomènes linguistiques, et aussi à une certaine posture d'humilité face aux données et dans la recherche d'une manière plus générale. Ainsi, elle a fortement contribué à la construction de ma posture de chercheuse actuelle et je lui dois beaucoup.

Dans la formation doctorale, comme chacun sait, l'université d'inscription n'est pas la seule formatrice, les participations à des conférences et des colloques, hors des murs ou non, étant une autre source féconde de rencontres, d'inspiration, et, somme toute, d'expériences formatrices.

J'ai ainsi eu la chance d'échanger, dès mes premières années de thèse, avec de nombreux jeunes collègues travaillant sur différents états de langue, sur le même type de problématique, soit à Nanterre, par l'intermédiaire de collègues, soit en colloque. Pour le français parlé moderne, Sylvain Kahane m'a présenté dès le début de ma thèse Christophe Benzitoun, qui travaillait alors sur les structures en *quand* « insubordonnées » (Benzitoun 2006a), et Mathieu Avanzi, en thèse en co-tutelle entre Nanterre (avec Anne Lacheret) et Neuchâtel (avec Marie-José Béguelin) sur les rapports prosodie-syntaxe, en particulier dans la parataxe (on ne peut faire plus proche !) (Avanzi 2011 et 2009 [2012]). Sur le français écrit, Antoine Gautier travaillait alors sur la question de la phrase, de sa définition et de ses limites (Gautier 2006) et en diachronie du français, Thomas Verjans travaillait sur les locutions conjonctives du latin au français classique français (Verjans 2009), tous deux rencontrés grâce au groupe de travail d'Annie Bertin. Enfin, grâce à l'organisation du colloque Jeunes Chercheurs de Modyco (Glikman, Mansour et Weiser (éds) 2007 [i]), j'ai rencontré Bérengère Bouard qui faisait une thèse, en histoire des théories linguistiques, sur l'histoire du complément (Bouard 2007), et Noalig Tanguy sur la phrase averbale en français parlé (Tanguy 2009). De véritables

⁶ Dans ce mémoire, pour conserver une lecture fluide, les références à mes propres travaux se font de manière traditionnelle suivant le modèle Auteur Date, suivi entre crochets du numéro correspondant dans la table des matières du volume 2 (volume des travaux), dans lequel mes publications sont classées par ordre chronologique et numérotées en chiffres arabes. Les numérotations en chiffres romains renvoient aux introductions et tables des matières des collectifs que j'ai coédités, regroupées en annexe.

collaborations se sont dès lors mises en place, qui se sont poursuivies jusqu'ici, comme nous le verrons tout au long de ce mémoire de synthèse. Ces rencontres ont été pour moi une véritable ouverture, contribuant à ma formation intellectuelle.

Christophe Benzitoun, formé à Aix-Marseille à l'étude du français parlé par Claire Blanche-Benveniste et José Deulofeu, travaillait sur les structures en *quand* insubordonnées. C'est suite aux discussions que nous avons eues ensemble que j'ai exploré la question des verbes recteurs faibles en ancien français, autour des travaux de Claire Blanche-Benveniste (1989) et Claire Blanche-Benveniste et Dominique Willems (2007). Suite à cela, nous avons organisé avec Mathieu Avanzi une journée d'étude sur la question des verbes parenthétiques, publiée dans Glikman & Avanzi (dir.) (2009 [2012]) [ii], avant d'organiser avec Stefan Schneider un colloque international sur la même question, qui a donné lieu au volume collectif Schneider, Glikman & Avanzi (éds) (2015) [iii]. Un colloque international sur la parataxe a même été organisé à Neuchâtel pendant mes années de thèse (une chance que tout le monde n'a pas eu !) (Béguelin et al. 2010), auquel j'ai participé (Glikman 2010a [6]), et qui m'a fait également rencontrer plusieurs collègues avec qui j'ai continué d'échanger par la suite. Marie-José Béguelin m'a ensuite invitée à participer à leur séminaire. J'ai pu y rencontrer Gilles Corminboeuf, qui a travaillé sur les structures en *si* à la dépendance floue (type : *si tu as soif, il y a de la bière au frigo*) (Corminboeuf 2009), ainsi que Frédéric Gachet dont la recherche portait sur les structures en *à peine* et leur intégration (*à peine était-il sorti...*) (Gachet 2010), avec qui nos questionnements théoriques se rejoignaient. J'ai également participé au numéro thématique coordonné par A. Gautier, L. Pino Serrano et D. Van Raemdonck sur « La rection verbale entre micro- et macro-syntaxe » (Glikman 2010b [7]). Tous, ainsi, nous travaillions sur ces questions de constructions verbales au rattachement flou (Benzitoun 2006a, Avanzi 2012), échangeant nos points de vue et nos exemples. Avec Mathieu Avanzi et Christophe Benzitoun, entre autres, nous avons continué de collaborer autour de différents projets sur la variation contemporaine, dans le cadre des projets *Français de Nos Régions* et *Donnez votre Français à la Science* (→ 3.2.). Ces différentes rencontres m'ont permis d'ouvrir mon cadre théorique aux approches développées pour la description du français moderne, notamment les théories macro-syntaxiques, me permettant d'avoir une appréhension différente des phénomènes de l'ancien français. J'étais en un sens la « diachronicienne » du groupe, construisant un « pont » entre les données modernes et anciennes et mettant au jour la persistance de ces phénomènes (cf. Glikman 2013 [11], → 4.). Cette position en équilibre entre l'ancien français et le français moderne, et les modèles de description employés pour l'un et l'autre état de langue, est caractéristique de mon parcours, et se retrouve dans mes différents projets de « diachronie

contemporaine ». En cela, elle me semble vraiment définitoire et constitutive de ma trajectoire scientifique (→ 3.).

En termes de méthode, Christophe Benzitoun m'a également, dès nos premiers échanges, montré qu'on travaillait sur la même chose, puisque, que ce soit lui pour l'étude du français parlé, ou moi pour l'étude des phénomènes en diachronie, nous partons des attestations, avec la difficulté qui en découle : comment appliquer les tests habituels en linguistique, et s'assurer que la forme ainsi construite soit réellement attestée ou du moins potentiellement « productible » (→ 2.) ? Certes, sur le français contemporain, on peut renvoyer à la compétence des locuteurs⁷, mais comment être sûr que cette forme puisse réellement être spontanément produite ? Quiconque a fait un jour des tests d'acceptabilité avec ses étudiants connaît la variabilité des réponses.

J'ai ainsi échangé avec des chercheurs des deux écoles de macro-syntaxe (Aix et Fribourg), travaillant sur l'oral (ou sur la prosodie pour Mathieu Avanzi), tout en fréquentant de près aussi la communauté des diachroniciens et de la syntaxe diachronique, ce qui m'a permis de suivre les avancées récentes de la recherche dans ce domaine. J'ai participé aux colloques *Diachro* dès la troisième édition en 2006 et aux panels diachroniques du *Congrès Mondial de Linguistique Française*, rencontrant ainsi Christiane Marchello-Nizia, Bernard Combettes (que je voyais par ailleurs également au groupe de recherche d'Annie Bertin), Sophie Prévost et Céline Guillot-Barbance, qui m'ont tout de suite intégrée dans leur communauté et invitée à travailler avec eux. C'était pour moi un vrai plaisir et un juste retour des choses d'organiser la huitième édition de *Diachro* à Strasbourg en 2017 avec mes collègues (dont les actes ont été publiés en 2019, cf. Capin, Glikman, Obry et Revol (éds) 2019 [vi]). Je suis membre de la Société Internationale de Diachronie du Français (SIDF) depuis sa constitution, membre du bureau dès 2014, et webmestre depuis 2018⁸. Cette intégration se voit également à travers les différents projets de recherche en diachronie auxquels j'ai eu la chance d'être invitée à participer, entre mes premières vacances de recherche sur l'annotation en 2008 pour le projet Elico⁹, mon post-doctorat dans le projet SRCMF, mon implication dans le projet de la GGHF, pour lequel je faisais à ses débuts de la veille bibliographique, et ma participation au projet Democrat

⁷ Selon Nicolas Mazziotta, cette compétence de locuteur natif reste même partiellement valable pour les états anciens de langue, jusqu'à un certain point. On voit bien qu'un francophone naïf aborde, de fait, plus aisément un texte ancien qu'un allophone ne connaissant rien du français, ancien comme moderne.

⁸ Voir le site de l'association : <https://diachronie.org/>.

⁹ <http://elico.linguist.jussieu.fr/index.php>.

(*DEscription et MOdélisation des Chaines de Référence : outils pour l'Annotation de corpus (en diachronie et en langues comparées) et le Traitement automatique*)¹⁰. Mon post-doctorat dans le projet SRCMF a été particulièrement formateur et enrichissant pour moi, tant en termes de rencontres que d'expériences. Comme je l'ai évoqué précédemment, c'est le moment où j'ai développé ma compétence en analyse syntaxique dépendancielle.

L'annotation manuelle exhaustive d'un texte amène en outre à se poser des questions d'analyse qu'on aurait facilement tendance à écarter, ou à résoudre trop vite, parce qu'elle suppose de faire des choix. La constitution d'un corpus annoté n'est pas anodine et demande de mener en amont à la fois une réflexion théorique, méthodologique mais aussi pratique, comme j'ai pu en faire l'expérience à la fois dans SRCMF et dans le projet *Rhapsodie*, ayant participé à quelques réunions et suivi de près l'évolution du projet comme je l'ai évoqué ci-dessus, puis dans le projet Democrat. Cette expérience me sert de base de référence dans mon projet actuel de constitution d'un corpus de sms vocaux. Mon post-doctorat dans le projet SRCMF constitue aussi le moment où j'ai approfondi ma compétence technique pour les outils et l'exploitation des corpus. En termes de collaborations enfin, j'ai eu la chance dans ce projet de rencontrer et de renforcer mes liens, scientifiques et amicaux, avec Sophie Prévost, Céline Guillot-Barbance, Alexei Lavrentiev, Serge Heiden, Tom Rainsford, Achim Stein et Nicolas Mazziotta.

Cet intérêt à la fois pour l'oral et l'ancien français m'a orientée dès le début (et c'est même là l'une des hypothèses principales de ma thèse) vers la question des rapports oral / écrit en diachronie du français (→ 4.). Je n'ai jamais totalement réussi à proposer une « macro-syntaxe » de l'ancien français, nos documents restant des textes écrits, et les traces de la production de l'oral « en train de se construire » (cf. Blanche-Benveniste et al. 1990), même dans le discours représenté, étant gommées. Cette question débouche assez vite en diachronie sur la question de l'accès à l'oralité, et, par-là, au statut de la représentation de l'oral, question qui a également connu un essor considérable ces dernières années. Cet essor est directement lié, à la fois cause et conséquence, d'une part à l'intégration à la *Base de Français Médiéval* (Guillot-Barbance, Heiden, Lavrentiev 2017) de l'annotation systématique des passages de discours rapporté, d'autre part, au développement de l'intérêt pour des sources jusqu'ici peu exploitées.

Là encore, j'ai eu la chance de pouvoir rejoindre et interagir avec toute une communauté de chercheurs s'intéressant au sujet, à travers de nombreux colloques sur le sujet, depuis le colloque *Représentation du Sens Linguistique 5* organisé à Chambéry en 2011 par Pierre

¹⁰ <https://www.lattice.cnrs.fr/democrat/>.

Larrivée et Dominique Lagorgette, et qui comportait une section « Représenter l’oral en diachronie » (Lagorgette et Larrivée (éds) 2013), le colloque *Représentation de l’oral et marques d’oralité dans la diachronie du français*, organisé à Madrid en 2016 par Amalia Rodríguez Somolinos, le deuxième colloque de la SIDF à Cambridge en 2014, comprenant à la fois une section sur les sources peu exploitées en diachronie et une autre sur la reconstitution de l’oral, dont j’ai co-édité les actes (Ayres-Bennett et al. (dir) 2018 [v]), jusqu’à la section sur les marqueurs de discours au quatrième colloque de la SIDF à Neuchâtel en 2018 que j’ai également co-édité (Glikman, Parussa et Waltereit (éds) 2019 [viii]). Sans y être totalement intégrée, je suis aussi les travaux menés dans le cadre du groupe de recherche sur la circulation de discours CI-DIT¹¹, Annie Bertin m’ayant fort heureusement fait connaître très tôt les travaux de Sophie Marnette notamment, ainsi que les travaux sur l’oralité et l’écriture dans le théâtre en diachronie menés dans le cadre du programme de recherche codirigé par Gabriella Parussa¹². Mes travaux s’inscrivent ainsi dans un champ de recherche très actif sur le sujet et s’en nourrissent. Je m’intéresse principalement à la question des spécificités syntaxiques de l’oral représenté (cf. Glikman et Mazziotta 2013 [12] et Mazziotta et Glikman 2019 [27]), des marqueurs de discours, et d’une manière plus générale à la question des sources et de l’accès à la représentation de l’oral ainsi qu’au statut de ces représentations, et enfin à la variation oral / écrit comme pôle de variation et son rôle dans la distribution de variantes, interrogeant son statut en tant que dia-variation (cf. Glikman et Schneider 2018 par ex. [23]). J’ai également abordé dans ma thèse les questions du statut syntaxique du discours rapporté, en tant que discours subordonné sans marqueur lorsqu’il est rapporté au style direct. Cette question de l’intégration du discours rapporté au style direct se pose aussi lors de l’annotation syntaxique des corpus, où, pour des raisons finalement pratiques plutôt que théoriques, nous avons choisi d’en faire des unités syntaxiques autonomes, dans le corpus SRCMF comme celui, en cours de constitution, des sms vocaux (→ 2.2.3.).

C’est lorsque j’étais post-doctorante au Lattice sur le projet SRCMF que j’ai rencontré Catherine Schnedecker et que j’ai commencé à travailler sur les chaînes de référence. En effet, Frédéric Landragin organisait à la même époque le projet PEPS MC4 « Modélisation

¹¹ <http://groupe-cidit.com/wordpress/>.

¹² <https://lamop.panthonsorbonne.fr/axes-recherche/livres-textes-langages/oralite-ecriture>.

Contrastive et Computationnelle des Chaines de Coréférence »¹³. À force de les entendre faire leur réunion dans la salle à côté de mon bureau et d'en discuter à la pause déjeuner, tout en annotant justement à la même époque des textes d'ancien français de manière cursive pour la syntaxe, j'ai fini par m'immiscer dans leurs réunions (et je remercie encore Sophie Prévost de m'avoir autorisée à prendre ces quelques heures sur mon post-doctorat pour y participer !). Dans la foulée, j'ai eu la chance et l'honneur d'être recrutée à l'Université de Strasbourg, et le travail sur les chaînes de référence s'est ainsi poursuivi, à travers une journée d'étude organisée en 2015 (cf. Schnedecker, Glikman et Landragin (dir.) 2017 [iv]), et bien sûr le projet Democrat, la constitution du corpus et son exploitation ayant donné lieu à de nombreuses publications communes¹⁴.

Mon recrutement à Strasbourg s'est aussi révélé particulièrement enrichissant pour moi grâce aux différents collègues rencontrés. En particulier, je n'étais pas la seule linguiste diachronienne de la faculté, chance précieuse puisque de plus en plus rare, de nombreux collègues diachroniciens se retrouvant de fait assez isolés dans leur département de rattachement. Au contraire, avec mes collègues diachroniciens Daniéla Capin et Thierry Revol, nous avons pu mener plusieurs projets, comme l'organisation du colloque *Diachro* que j'ai déjà évoqué, ainsi que plusieurs journées d'études sur la thématique des réécritures médiévales (2015, 2015 et 2017). Mes intérêts de recherche ont également rencontré ceux de mon équipe d'alors au sein de LiLPa, Fonctionnement Discursif et Traduction¹⁵, autour de la notion de genre (textuel), qui a été la thématique retenue pour le séminaire de recherche durant quelques années, et a donné lieu à une journée d'étude organisée avec Christophe Gérard en 2017, qui a abouti à la publication d'un numéro collectif (Glikman et Gérard (dir.) 2019 [vii]).

Ainsi, au fil du temps, à travers les hasards des colloques et des collaborations, j'ai rencontré et collaboré avec plusieurs chercheuses et chercheurs de cadres théoriques différents, travaillant en diachronie du français, sur le français moderne ou sur l'histoire des théories linguistiques, qui ont contribué à enrichir ma réflexion. Je ne pourrais tous les nommer ici, mais je les remercie encore pour les échanges que nous avons eus et que nous aurons, je l'espère, encore. Ce sont

¹³<https://www.lattice.cnrs.fr/projets/projets-passes/projet-peps-mc4-modelisation-contrastive-et-computationnelle-des-chaines-de-coreference/>.

¹⁴ Glikman et al. 2014 [14], Oby et al. 2017 [20], Oberlé et al. 2018 [25], Capin et al. 2021 [33].

¹⁵ Le laboratoire a depuis connu une réorganisation en axes, comme le vent emporte la marée, mise en place pour le prochain contrat.

notamment ces rencontres et la diversité de ces points de vue qui ont construit, briques après briques, mon identité scientifique.

1.2. Briques après briques, construction de mon identité scientifique

*C'est peut-être que je ne suis que le
fruit de mon époque*

Chiens de Paille, *Maudits soient les
yeux fermés*, 1998.

Quand je dois me présenter et présenter mon domaine de recherche, la traditionnelle déclaration que je fais est que *je travaille en linguistique diachronique sur le français*, et c'est évidemment bel et bien ce qui me définit, je suis « diachronicienne ». Ma problématique de recherche principale est d'étudier le changement linguistique, d'en observer, décrire et comprendre le fonctionnement. Pour autant, et bien que l'approche diachronique demande une certaine polyvalence, cette définition est incomplète. Si je devais définir plus précisément mon domaine d'expertise, je dirais que c'est le domaine syntaxique. Pourtant, une partie de mes travaux traitent de problématiques qui dépassent le cadre de la syntaxe, comme l'étude des chaînes de référence, qui se définissent, de fait, par leur empan textuel, ou encore la question de l'effet du type de discours sur la langue, même si, dans ce cas, ce sont le plus souvent des phénomènes syntaxiques qui sont considérés. Enfin, certains de mes travaux concernent des éléments pouvant être considérés comme relevant du lexique, comme les subordonnants (bien qu'étant des formes appartenant à la liste « fermée » de mots grammaticaux), ou le paradigme de « avoir été/être allé », ou encore quand j'encadre des étudiants de Master sur la diffusion et l'intégration de termes émergents. Il se pose alors également la question de savoir s'il s'agit toujours de syntaxe, ou si nous sommes passés à un autre domaine, celui de la linguistique textuelle, dans le premier exemple, ou celui du lexique, pour le second. De la même manière, quand je travaille sur les rapports entre usages et normes, quelle est la part de la syntaxe dans ces recherches, n'est-on pas plutôt dans la sociolinguistique, au moins pour les méthodes et les outils, si ce n'est pour les objets considérés ? De même, peut-on véritablement étudier des phénomènes syntaxiques sans interroger leur effet sémantique, en particulier quand il s'agit d'observer la variation ? De fait, mes travaux ont évolué (et c'est heureux !) dans leurs objets, de phénomènes proprement syntaxiques au début, avec la parataxe (→ 2.2.), à des phénomènes plus textuels (→ 2.3.4.), jusqu'au focus sur le sujet parlant, avec la question des représentations linguistiques et la notion de préférence (→ 4.2.), ainsi que dans leurs méthodes (→ 3.2.).

Cependant, il me semble que c'est bien la syntaxe qui est au cœur de ma recherche et mon domaine de spécialité, autour de la question principale de parvenir à définir ce qui relève de la grammaire d'une langue et ce qui relève du style : qu'est-ce qui est contraint syntaxiquement, et qu'est-ce qui relève d'une préférence, d'un choix ? Cette approche permet ainsi de comprendre selon moi ce qui constitue le changement linguistique dans la grammaire, où il se situe, comment il opère, en cherchant à observer le passage d'une distribution libre à une distribution contrainte. En cela, la question du choix et des contraintes rejoint le concept d'« obligatorification » (*obligatoriness*) évoqué dans différentes approches de la grammaticalisation, bien que dans mes travaux je n'aborde que rarement la question en ces termes. Le statut de l'*obligatorification* comme étape nécessaire de la grammaticalisation reste discuté (cf. Marchello-Nizia 2006, Lehmann 2002, Diewald 2011), mais peut en tout cas constituer un stade dans le processus de changement. Laissant de côté la question de son importance dans la conceptualisation de la grammaticalisation, le passage du caractère libre à un caractère contraint est avant tout, pour moi, révélateur d'une modification dans le statut des variantes dans le système, point qui m'intéresse principalement.

Mes activités de recherche, d'encadrement et de diffusion scientifique sur le changement linguistique s'articulent autour de deux axes principaux, qui sont pour moi complémentaires : la diachronie du français et l'étude de la variation en français contemporain.

Après ma thèse, j'ai poursuivi mes recherches sur la syntaxe de l'ancien français et l'évolution de la phrase complexe. La question de l'intégration des constituants phrastiques m'a conduite à aborder les problèmes de complémentation et de rection, en particulier l'étude des constructions verbales atypiques, comme les verbes dits « recteurs faibles » ou « parenthétiques », type *je crois, j'espère, je pense*, qui peuvent régir une proposition complétive ou apparaître en position incisive. En cherchant à formaliser la distinction entre un assemblage de deux propositions formant une unité supérieure et deux propositions syntaxiquement autonomes, mes travaux contribuent ainsi à l'étude des phénomènes qui se trouvent à la frontière entre la phrase et le discours (→ 2.2.).

L'étude de l'évolution des langues passant par la reconnaissance des variantes et la compréhension des règles qui sous-tendent l'actualisation des différentes variantes du système en fonction des différents types de variation, je me suis également intéressée à l'étude des variations non seulement en diachronie, mais aussi en synchronie moderne (→ 3.). Après avoir mis au jour, dans ma thèse, l'importance de la variation oral/écrit dans l'évolution de la parataxe, j'ai orienté mes recherches vers le rôle des rapports oral/écrit dans une théorie du changement linguistique. Dans ce but, une partie de mes recherches vise à interroger les traces

d'oralité récupérables en ancien français et le statut de l'oral représenté. J'ai cherché à inclure par la suite la notion de genre textuel ou discursif dans l'évolution des phénomènes linguistiques. Dans le cadre de l'étude de la variation en français contemporain, je m'intéresse, d'une part, à la variation régionale, ma situation géographique à Strasbourg favorisant l'étude de ces questions, mais aussi, et surtout, à l'émergence, la diffusion et la disparition des formes linguistiques en français contemporain, ainsi qu'à la perception que peuvent avoir les locuteurs de ces changements. À travers ces questions, c'est aussi le rapport entre *usages* et *normes* que j'interroge, à la fois en diachronie, mais aussi pour les périodes contemporaines (→ 4.).

Ma délégation CNRS au sein du laboratoire ATILF à Nancy (2021-2022, incluant un semestre de CRCT du CNU et renouvelée en 2022-2023) pour le projet *Le rapport oral / écrit et les nouveaux modes de communication : étude de diachronie contemporaine* s'inscrit dans la continuité de ces travaux sur les variations contemporaines et les jugements des locuteurs sur leur langue et sur les rapports oral / écrit à travers une étude sur les modes de communication actuels. Ces questions seront au centre de ma recherche pour plusieurs années. C'est l'occasion pour moi d'interroger le statut des variantes et le rapport à la norme, dans une démarche qui associe les approches diachronique et synchronique.

Ma conception du changement linguistique se situe plutôt dans un cadre de type fonctionnaliste, si l'on reprend la distinction proposée par Sophie Prévost (2020) entre les courants formalistes (représentés par la grammaire générative, dans laquelle sont aussi développées des études diachroniques) et les courants fonctionnalistes (Prévost 2020 : 17 sqq). Cette approche présuppose une conception de la langue comme un système hétérogène et dynamique (cf. aussi Banniard 1997 et la notion de *diasystème* chez Coseriu 1998). En cela, je m'intègre parfaitement dans le courant théorique adopté dans la *Grande Grammaire Historique du Français* (Marchello-Nizia et al. (éds) 2020), projet auquel j'ai moi-même participé (cf. Glikman 2020a [29], b [30] ; Combettes et Glikman 2020 [28]) :

L'approche adoptée dans la GGHF repose sur la conviction, confirmée par l'étude des données, que la variation est au cœur, et à l'origine, des changements, et que l'usage de la langue par les locuteurs joue un rôle essentiel, tant dans leur déclenchement que dans leur diffusion. (Prévost 2020 : 20)

Ma conception du changement linguistique est également basée sur une conception relevant d'une sociolinguistique historique, ou peut-être plus exactement d'une linguistique variationnelle (Völker 2009), plaçant la variation et le locuteur au centre du changement. Ces approches se sont développées notamment grâce aux travaux de Michel Banniard, Serge Lusignan, Anthony Lodge. Sur les périodes anciennes de la langue, c'est la prise en compte des

types de textes qui permet d'accéder à certains aspects de la variation (cf. Glessgen et Schøsler 2018), pour les états de langue dont on ne conserve pas de traces du vernaculaire. La question des rapports oral / écrit, telle qu'elle a été notamment revue par Koch & Oesterreicher (2001, entre autres), permet également d'aborder ces questions d'une manière renouvelée et, selon moi, efficace. L'analyse du rapport oral / écrit a ainsi pris une part importante dans les études diachroniques ces dernières années, que ce soit à travers l'étude des marqueurs du discours, des marques d'oralité, de la représentation de l'oral ou du lien langue parlée / langue écrite (Schneidecker & Vaguer (dir.) 2020, Glikman et al. (éds) 2019 [viii], Ayres-Bennett et al. (dir.) 2018 [v], Blasco & Bodelot (dir.) 2017, Rodriguez Somolinos (dir.) 2013 et 2016, Lagorgette & Larrivée (dir.) 2013). L'approche a été renouvelée par l'introduction, par Koch & Oesterreicher (2001), de la distinction entre le médium, c'est-à-dire la dimension matérielle concrète de la réalisation, suivant une opposition entre code écrit et code oral, et le concept, c'est-à-dire la dimension conceptuelle de la réalisation, qui peut se situer à différents niveaux sur une échelle entre l'immédiat communicatif et la distance communicative. Si la pertinence de l'introduction de la notion de diamésie au terme des dia-variations fait débat (voir, entre autres, Wüest 2009), ces deux aspects restent cependant productifs pour les études linguistiques, la prise en compte du médium et celle de la dimension conceptuelle. C'est l'approche qui est retenue dans la *Grande Grammaire Historique du Français*, comme je l'ai évoqué, mais aussi dans le volume sur le passage du latin au français édité par Anne Carlier et Céline Guillot-Barbance (2018), montrant les apports pour la description¹⁶. La prise en compte de sources jusqu'ici peu étudiées pour l'histoire du français, telle que plusieurs collègues la mènent actuellement (voir Ayres-Bennett et al. (dir.) 2018 [v]) permet également d'approfondir les recherches dans ce sens. L'exploitation de ces ressources, mise à la disposition de la communauté à travers la constitution de corpus par ces collègues¹⁷ permettra, à mon sens, des avancées dans les prochaines années pour l'histoire du français.

Dans ce contexte scientifique prolifique et enthousiasmant, mes travaux portent plus spécifiquement :

¹⁶ Voir Glikman, Julie 2022. Compte rendu. Anne Carlier, Céline Guillot-Barbance (éds.) 2018. *Latin tardif, français ancien. Continuités et ruptures*. Berlin : De Gruyter. *Studia linguistica romanica* 2022.7, 126-140. <https://doi.org/10.25364/19.2022.7.6>.

¹⁷ Pour ne citer que quelques exemples, Sermo (<http://sermo.unine.ch/SERMO/>), le Corpus 14 (<https://www.univ-montp3.fr/corpus14/>), sans oublier les corpus de chartes ou l'enrichissement continu de la BFM.

- en ancien français, sur la description des contextes, des conditions de réalisation des variantes¹⁸, sur la question des genres textuels et du type de discours et de ce que ça peut dire sur le statut des variantes, sur la diffusion donc bien sûr d'un point de vue quantitatif, mais aussi qualitatif, sur le parcours de la diffusion ;
- en diachronie longue, sur la répartition des variantes comme pour l'ancien français, mais aussi sur l'évolution des évaluations de ces variantes (rendue possible grâce à l'apparition du métadiscours sur la langue, qui s'est développé très tôt pour le français, à travers les manuels de langage dès la fin du 14^e - 15^e siècles, et surtout le développement des grammaires et des débats sur la langue au 16^e siècle) ;
- en diachronie contemporaine, sur la répartition, diffusion et évaluation des variantes, comme en diachronie longue, à quoi s'ajoute l'étude de la perception des variantes par le locuteur et des représentations linguistiques.

Concernant l'étude des dia-variations en français contemporain, ce qui m'intéresse est la coexistence de deux formes en synchronie, et l'étude des rapports qu'elles entretiennent entre elles. Étudier le statut des variantes nous permet de voir comment les innovations se diffusent dans la langue et peuvent constituer des changements en cours, suivant le principe énoncé par Labov : « les mécanismes qui ont produit les grands changements du passé sont ceux-là mêmes que nous voyons à l'œuvre autour de nous » (Labov 1976 : 232, voir aussi Labov 2007, Labov 1992). L'idée n'est évidemment pas de chercher à prédire le changement, mais de profiter des observables que nous pouvons avoir sur les périodes contemporaines pour enrichir notre compréhension des mécanismes en jeu dans le changement linguistique. C'est la raison pour laquelle mes travaux sur les périodes modernes se situent selon moi dans une approche de « diachronie contemporaine » : j'applique les outils et les concepts de la diachronie à des périodes contemporaines, pour contribuer à l'étude du changement linguistique. Je m'intéresse dans ces travaux principalement à deux aspects de la variation : 1. le jugement des locuteurs sur leur propre langue (les « positionnements épilinguistiques », Canut 1998, ou le « sentiment de la langue », Siouffi (dir.) 2012), en particulier leur perception des différentes variantes, et 2. le rapport oral / écrit et la question de la variation diamésique (Gadet 2020), autour des contraintes

¹⁸ J'utilise « variante » au sens dans lequel il est habituellement utilisé en linguistique variationnelle ou en sociolinguistique, « différente manière de dire la même chose », c'est-à-dire différente réalisation possible d'un même phénomène linguistique pouvant alterner dans un même contexte ou en contexte comparable. Les variantes au sens philologique dans les manuscrits constituent, en un sens, une voie d'accès à l'étude de la variation, et gagneraient à être davantage exploitées dans les études linguistiques, comme le montre bien Thomas Verjans (2012).

linguistiques liées au code et de celles relevant au contraire de l'opposition entre immédiat communicatif et distance communicative (Koch & Oesterreicher 2001), ainsi qu'à leur mise en œuvre dans les nouveaux modes de communication, à travers mon projet de recherche sur les sms vocaux.

Dans le cadre d'une théorie du changement en trois phases (innovation, diffusion, intégration), telle que développée par Weinreich, Labov et Herzog (1968), et telle que Lodge (2018) en rappelle les problématiques et les enjeux, mes travaux contribuent ainsi à l'étude des phénomènes de diffusion, avec l'établissement des faits de variation : les variantes, représentatives des « innovations ». Les questions autour de comment ou pourquoi les innovations se produisent ne sont pas celles que j'aborde dans mes recherches, qui se concentrent davantage sur ce qu'il se passe après la phase d'innovation. Dans mes travaux, je vise à les reconnaître comme telles et à établir ce avec quoi elles entrent en concurrence (→ 2.). Je pars généralement de l'observation d'un cas de concurrence à un moment donné, pour en retracer le chemin : développement de l'une et l'autre forme, répartition, distribution, et évolution de cette distribution (extension à différents contextes, différents types de textes, disparition / apparition de contraintes dans la réalisation, stigmatisation, réorganisation de la distribution (cf. la variation *avoir été / être allé*, perçue d'abord comme sémantique puis comme diastratique, cf. Glikman et Patard 2022 [35]). Mes travaux portent donc non seulement sur le fait d'avoir de plus en plus d'occurrences de la nouvelle forme dans des contextes de plus en plus variés, mais aussi, et c'est vers là que je développe mes recherches, sur la perception de ces formes, les travaux récents (cf. Siouffi 2020) ayant montré l'importance de remettre le locuteur au centre de l'étude du changement. Ainsi, à travers mes recherches, je vise à contribuer à la compréhension et la définition de ce qui constitue l'étape « 2 », celle de la diffusion, la « propagation », à travers l'étude des « transition problem » et « evaluation problem » (cf. Lodge 2018 pour une présentation succincte), à travers l'étude des usages (en corpus), mais aussi de la norme, du discours normatif (qui reste une représentation de la langue, même s'il se veut prescriptif), de l'évolution du discours normatif (stigmatisation...), et de la conscience linguistique des locuteurs, permettant de se rendre compte de leur intégration ou non de la norme prescriptive ou de la diffusion de formes (voir par ex. mes travaux sur *de façon à ce que* présentés au colloque *Diachro X* en 2022, encore non publiés).

Les travaux que j'ai menés jusqu'ici contribuent à montrer l'importance de la prise en compte du sujet parlant et le rôle des représentations linguistiques du locuteur dans l'élaboration d'une théorie du changement : il y a changement quand la perception de la variante change (→ 4.). Le changement de la perception de la variante constitue pour moi une étape importante dans

le passage du fait de parole, individuel, au fait de langue, pour reprendre la terminologie saussurienne, ou de l'innovation (phénomène individuel) à sa diffusion à l'ensemble de la communauté (phénomène social), pour revenir à la terminologie de Labov. La perception des locuteurs de ces variantes est donc un indice important du changement. Barra-Jover (2009) introduit ainsi la distinction entre variantes invisibles et variantes visibles, complémentaire de l'analyse des usages. C'est également ce qui est au centre des études du volume « Sentiment de la langue et diachronie » (Siouffi (dir.) 2012). Le statut de ces variantes peut être étudié à travers les discours métalinguistiques, provenant ou non d'experts de la langue (depuis les grammaires, les chroniques dans les journaux, aux blogues et forums sur la langue, voir Siouffi, Steuckardt et Wionet 2012), ainsi qu'à travers des enquêtes linguistiques diffusées directement auprès des locuteurs (voir Avanzi et al. 2016 [16] et Glikman et al. 2018 [22] sur les enquêtes que nous avons menées dans le cadre des projets *Français de nos Régions* et *Donnez votre Français à la Science*, → 3.2., et Glikman 2019 [26] et Glikman et Patard 2022 [35] pour des exemples d'exploitation des résultats en diachronie longue, → 4.2.).

Cela ne remet pas pour autant en question les théories de la « main invisible » (Keller) ou de la directionnalité du changement (voir Marchello-Nizia 2006 pour une synthèse de ces questions), mais la prise en compte du sujet parlant et de ses représentations sur la langue donne ainsi un indicateur de la différence entre le niveau individuel et collectif, la différence pouvant justement être observée selon moi dans ces représentations, qui constituent autant de normes inconscientes collectives. Cela conduit également à la question des « possibles » en langue (cf. Verjans 2014), et, à travers la question des choix et des contraintes, qui constitue un fil rouge tout au long de ce mémoire de synthèse, à la notion de « préférence » pour des éléments grammaticalement possibles mais non préférés dans les usages, que mes travaux m'ont conduite à élaborer (→ 4.), en particulier lors de l'étude des complétives sujet (question évoquée dans Glikman 2020b [30], et qui a fait l'objet d'une communication présentée en 2019 au colloque *Diachro IX*, encore non publiée) et que je souhaite continuer d'explorer dans les années à venir.

L'approche des grammaires de construction offre aussi une approche intéressante des phénomènes du changement linguistique (cf. Carlier et Prévost (dir.) 2021, notamment Schøsler 2021). Aucun de mes travaux ne s'inscrit explicitement dans ce cadre théorique, mais une partie de mes recherches pourrait gagner, me semble-t-il, à être repensé dans ce cadre. Par exemple, les travaux sur le développement des constructions parenthétiques en français (cf. Schneider et Glikman 2015 [15] et Glikman et Schneider 2018 [23]) conduisent à l'élaboration d'un schéma plus qu'au développement de formes spécifiques. C'est d'ailleurs dans cette direction que mon collègue Stefan Schneider poursuit ses travaux sur les constructions parenthétiques (cf.

Schneider 2019, 2022). On peut en rapprocher mes travaux sur les relatives asyndétiques (Glikman 2016 [18]), pour lesquels j'ai mis à jour différents types de structures plus ou moins figées et productives. Je pense également que l'étude des « marqueurs discursifs » pourrait être abordée sous cet angle (en particulier associé à la notion de paradigme, voir Kragh 2021) : le développement d'un cadre dans la langue, dans laquelle différentes formes individuelles peuvent se glisser, sans provoquer de « changement » au niveau du patron, de la construction. On peut citer ici aussi la notion de « routinisation » employée par Siouffi, Steuckardt et Wionet (2018) à propos des modalisateurs. C'est en tout cas une nouvelle piste théorique que je souhaiterais approfondir à l'avenir.

J'accorde une importance particulière à l'accès aux données, à leur enrichissement, et au partage de celles-ci avec la communauté, à travers les différents projets de recherche que j'ai dirigés ou auxquels j'ai participé comme je viens de le mentionner, ainsi qu'à travers mon implication dans des structures visant à favoriser la diffusion et la cohésion de la communauté et ayant pour objectif de développer les principes FAIR (*Findable, Accessible, Interoperable, Reusable*), comme le consortium CORLI¹⁹. C'est dans ce même esprit de « science ouverte » que je participe aux comités de rédaction des revues *Linx* et *Studia Linguistica Romanica*, toutes deux librement accessibles en ligne. Cet engagement passe par l'enseignement, constitutif de notre métier, mais aussi par mes activités d'animation de la recherche, avec l'organisation d'évènements scientifiques et mon activité au sein de la SIDF, ainsi que mes actions de vulgarisation auprès du public, car, faire de la recherche, c'est aussi transmettre et diffuser ses résultats, aux étudiants, à la communauté et à la société (→ 5).

1.3. Choix et certitudes, structure du mémoire de synthèse

*Être né quelque part,
C'est partir quand on veut,
Revenir quand on part*

Maxime Le Forestier, *Né quelque part*, 1988.

Ainsi, comme j'espère l'avoir montré dans ce qui précède, mon champ de recherche, celui dans lequel je souhaiterais pouvoir diriger des recherches à l'avenir, est la linguistique diachronique du français. Par diachronie, j'inclus aussi bien l'étude synchronique des états anciens de la langue, de l'ancien français jusqu'au français classique, que les études menées en

¹⁹ <https://corli.huma-num.fr/>.

diachronie large et celles menées en micro-diachronie, notamment contemporaine. Les domaines linguistiques dans lesquels je m'inscris sont plus particulièrement l'analyse de la subordination et de ses marqueurs, la caractérisation des discours, surtout les rapports oral / écrit, la représentation de l'oral en diachronie et ses marqueurs, et les contraintes liées aux genres discursifs d'une manière générale. L'analyse diachronique visant à la compréhension du changement linguistique, je souhaite développer mes recherches en ce sens, à travers l'étude des phénomènes de diffusion et d'évaluation du changement à travers le statut des variantes, via les représentations linguistiques et la place de ces représentations dans le changement, ainsi que les rapports entre usages et normes.

Le point commun à tous mes travaux, ainsi qu'à ces perspectives de recherche, est la variation, avec la conviction que le langage est variation (l'hétérogénéité du système) et que comprendre le changement passe par l'analyse de cette variation. C'est donc la notion de variation qui a été retenue comme fil rouge pour la construction de ce mémoire de synthèse, articulé en trois phases : établir les variantes (→ 2.), analyser la variation (→ 3.), pour comprendre le changement linguistique (→ 4.).

Le prochain chapitre sera consacré à la question de l'établissement des variantes (→ 2.). Ce sera l'occasion de mettre en avant mon ancrage dans le domaine syntaxique, autour de la question de l'établissement du lien syntaxique (et son marquage), de la délimitation de l'unité maximale et de la définition même de ce qu'est la syntaxe. Les objets d'étude de l'ensemble de mes travaux seront présentés ici, y compris ceux pouvant paraître à première vue plus éloignés du domaine syntaxique, mais pour lesquels se pose toujours la question de définir les variantes en concurrence. La question de la variation, aussi bien en tant que concept qu'en tant que méthode, sera traitée au chapitre suivant (→ 3), dans lequel j'aborderai la question des paramètres de variation et celle des méthodes pour parvenir à la décrire, ce qui me conduira à traiter dans ce chapitre la question de l'accès aux données, de l'enrichissement et du partage des données avec la communauté. J'exposerai dans le dernier chapitre ce que je pense que l'analyse de la variation peut nous apprendre sur le changement linguistique (→ 4), à travers la caractérisation des discours et le rôle des représentations linguistiques. Je montrerai ce que mes travaux ont apporté à ce jour sur la question, et quel champ de recherche cela ouvre, autour de la prise en compte du jugement linguistique du locuteur et de son évolution.

Autour de la question de la variation, c'est bien la notion de contraintes, imposées par la langue, par la grammaire, et celle de choix, relevant en théorie de la liberté dans la parole du locuteur, mais dont le choix peut lui-même être contraint par différents paramètres contextuels lors de la réalisation, qui sont au cœur de ma recherche.

**2. Établir les variantes. Autour des objets :
une approche syntaxique**

2. Établir les variantes. Autour des objets : une approche syntaxique

*Je me souviens encore
de ma première femme*

Parabellum, Cayenne (*Mort aux vaches*), 1989.

Dans une approche du changement linguistique qui postule l'existence de phases de transition, concrétisées par des phases de variation, observables en synchronie, une partie du travail, pour l'étude du changement linguistique, est d'étudier les paramètres de la variation. La première étape indispensable pour cela est d'identifier les éléments en concurrence. Étudier la variation doit donc, nécessairement, commencer par établir les différentes variantes, c'est-à-dire pouvant alterner dans le même contexte, ou du moins en contexte comparable. Je reviendrai dans le chapitre suivant sur l'analyse de la variation elle-même, de ses paramètres et des méthodes pour l'étudier (→ 3.). Dans ce chapitre, mon objectif est de présenter les différents objets d'étude sur lesquels j'ai travaillé, comme autant de « variantes » qu'il a fallu établir afin d'étudier leur distribution et les paramètres de variation. Ce sera également l'occasion de montrer mon ancrage dans le domaine syntaxique, malgré certains travaux menés en marge de la syntaxe. Dans cette présentation des objets, je n'adopterai pas un parcours strictement linéaire, même si, de fait, cette progression s'est naturellement jouée au fil du temps. Je commencerai par replacer rapidement la problématique de l'établissement des variantes en syntaxe (→ 2.1.), avant d'aborder les travaux que je propose de regrouper ici autour de la question du verbe comme noyau de la structure syntaxique (→ 2.2.), qui ont débuté avec mes premières recherches sur la notion de *parataxe*, « ma première femme ». Je montrerai comment j'ai cherché à définir les variantes à travers l'établissement du lien syntaxique (→ 2.2.1 et 2.2.2.). Cette problématique m'a conduit à aborder des questions plus générales, notamment sur la définition de la délimitation de l'unité maximale en syntaxe et, par suite logique, de la syntaxe (→ 2.2.3.). Je présenterai ensuite mes travaux portant sur d'autres zones de variable, où se pose dans chaque cas la question de la définition de la variante (→ 2.3.). Au centre de ces questionnements, on retrouve la question du choix, qui fait le lien avec la partie suivante autour de la variation (→ 3.), dans laquelle j'aborderai les questions de méthode.

2.1. Établir les variantes en syntaxe et morphosyntaxe

Same but different

Expression populaire.

Comme nous l'avons évoqué dans l'introduction, les travaux de W. Labov, entre autres, ont modifié la recherche en linguistique en s'intéressant à la variation. La naissance de la linguistique variationniste aux États-Unis à la fin des années 1960 a permis en France un renouveau des études diachroniques, qui avaient été reléguées au second plan depuis l'interdiction, en 1866, par la société linguistique de Paris de mener toute recherche sur l'origine du langage²⁰ et par l'avènement de la linguistique saussurienne accordant un statut privilégié à la synchronie. Labov travaillait cependant sur des variables phonétiques, à même d'établir des variantes en contexte comparable, même s'il envisageait dès le départ la possibilité d'étendre la méthode aux différents domaines de la linguistique. Or, qu'en est-t-il en syntaxe ? Comment appliquer la notion de *variante*, « différentes manières de dire la même chose »²¹, aux autres domaines ? Cette question a été soulevée très tôt dans les travaux s'inscrivant en sociolinguistique, comme en témoignent plusieurs contributions au numéro thématique « Hétérogénéité et variation : Labov, un bilan », vol. 108 de la revue *Langages* (1992). Françoise Gadet (1992) pose d'emblée la question en introduction :

Mais qu'est-ce qui garantit que les différentes variantes sont bien variantes d'une même unité ? Il y a ici un problème, sur lequel nous reviendrons plus loin : les variantes constituent « different ways of saying the same thing ». C'est clair pour la phonologie (les allophones), mais comment en ira-t-il hors de la phonologie ? (Gadet 1992 : 7)

La difficulté se manifeste sur le mode de définition de la variable, avec l'exigence que les différentes variantes expriment « the same thing », ce qui peut être facilement exigé des unités phonologiques qui n'ont pas de sens en soi, mais ouvre tout le problème de la syntaxe, où la question du sens et de son éventuelle équivalence a entièrement à être constituée. [...]

Pourtant, entre un phonème et une catégorie syntaxique (à supposer qu'on puisse ainsi regrouper les phénomènes d'ordre syntaxique), on change de domaine : si une variable phonologique peut si facilement devenir marqueur social, c'est grâce à son caractère totalement arbitraire, donc crucialement au fait qu'elle n'ait pas de sens. Caractère évidemment perdu au niveau syntaxique. D'où la difficulté de définir des variantes comme « different ways of saying the same thing ». [...]

Pourtant, en syntaxe, la distribution des unités, le conditionnement des restrictions, et la compatibilité des catégories lexicales sont sûrement des dimensions essentielles. On ne

²⁰ <https://www.slp-paris.com/statuts1866.html>

²¹ « It is common for a language to have many alternate ways of saying "the same" thing. » (Labov 1972: 188).

se trouve donc jamais dans une situation en quoi que ce soit comparable à ce que la variation représente en phonologie. (Gadet 1992 : 12-13)

F. Gadet confirme plus récemment cette position sur la question d'établir les variantes en syntaxe, difficiles à établir dans ce domaine pourtant central, preuve que le problème reste entier :

Dans la formulation courante de la variation, « différentes façons de dire la même chose », il est impossible de ne pas supposer un autre terme, un invariant face au variable. C'est pourquoi le dispositif convient au niveau phonique où le sens est constant, et c'est aussi pourquoi c'est à d'autres niveaux qu'il faut l'évaluer, en particulier en syntaxe. Niveau décisif pour au moins trois raisons : 1) la syntaxe est au cœur de la langue, entre du son et du sens ; 2) c'est le niveau le moins étudié pour la variation ; 3) l'invariant y est objet de débat, entre sémantique et pragmatique. (Gadet 2017b : 79)

Danièle Godard (1992 : 51) va jusqu'à déclarer en ouverture de son article « Il est clair aujourd'hui que le programme, dans son ensemble, n'est pas adapté aux problèmes syntaxiques. ». D. Godard analyse ensuite les limites de « la règle variable » du programme pour les phénomènes syntaxiques, montrant que le formalisme proposé par Labov est inadéquat pour traiter des structures syntaxiques. Outre les questions de formalisation, pour D. Godard, « l'hypothèse même d'une variation syntaxique est problématique. » (*ibid.* 56) :

Qu'est-ce alors qu'une variable syntaxique ? En fait, il y a deux situations possibles. La première est la suivante : deux sous-structures peuvent globalement, dans leur relation avec les catégories extérieures avec lesquelles elles se combinent, avoir le même sens, tout en présentant des différences de structuration interne ; elles constituent alors des variantes. [...]

La seconde possibilité est la suivante : on définit l'équivalence sémantique comme l'identité des conditions de vérité. (Godard 1992 : 57)

Dans le premier cas, elle prend l'exemple des relatives « le garçon auquel j'ai parlé » vs « le garçon que je lui ai parlé », qui comporte bien les mêmes unités lexicales et ont en commun le fait d'être insérées dans un GN. Elle illustre le second cas par les phrases « je donnerai mon Platon à Paul » et « à Paul, je donnerai mon Platon », qui n'ont pas tout à fait le « même sens » du fait de la focalisation sur l'un des arguments, mais qui ont bien les mêmes « conditions de vérité ». Une fois établie la variable, le problème n'est pas pour autant résolu :

Une fois définies les possibilités de variables syntaxiques, se posent les deux questions suivantes : ces situations manifestent-elles l'hétérogénéité du système, et son investissement par des différences sociales ? [...]

La seconde question concerne la différenciation des possibilités offertes par le système selon les dimensions sociales. L'un des intérêts des études de la variation phonologique

a été de corrélérer les variantes avec des différences extralinguistiques, sociales, stylistiques ou géographiques, ce qui permet non seulement d'étudier l'articulation du linguistique et du social, mais d'éclairer la question du changement linguistique. (Godard 1992 : 58-59)

Bien entendu, cela n'a pas empêché plusieurs études sur les phénomènes syntaxiques, moyennant parfois une adaptation du modèle de départ : « Il y a des variables syntaxiques qui ont donné lieu à une étude sociolinguistique, et qui se sont révélées porteuses d'une signification sociale ou stylistique. » (*ibid.* 59).

En syntaxe, établir les variantes ne va donc pas de soi, et la notion même de contexte comparable ne se laisse pas simplement appréhender. Comme le dit Godard, en quoi peut-on dire que deux structures syntaxiques sont « pareilles » tout en étant « différentes » ? Qu'est-ce qui est comparable et comment définir la variation dans un même « contexte » ? En effet, si l'on considère par exemple la réalisation du sujet, peut-on considérer les réalisations « le lion mange » et « le lion il mange » comme deux variantes ? Si l'on accepte ces deux réalisations comme des variantes, la question se pose de savoir jusqu'à quel point l'on peut étendre cette variable : la structure « c'est le lion qui mange » peut-elle encore être incluse comme une variante de la réalisation syntaxique du sujet ?

Le débat sur les différentes manières de dire la même chose rejoint, comme F. Gadet (2021) le montre bien, le débat sur la synonymie en langue :

Un débat crucial ouvert autour de ce concept, déterminant pour sa définition et l'extension des faits impliqués, concerne la possibilité même de dire « la même chose » de différentes façons dans une même langue – ce qu'à la suite de Claire Blanche-Benveniste (2010), on nommera le « débat sur la synonymie ». L'enjeu est d'établir s'il y a une réelle latitude de choix pour un usager. (Gadet 2021 : 332)

En effet, quelle part doit prendre la sémantique dans l'étude des variantes ? Dans le cas de « c'est le lion qui mange », un peu comme dans le cas de « à Paul, je donnerai mon Platon » cité par Godard, on distingue les énoncés par leurs contraintes informationnelles, au point parfois de ne pas les considérer comme des variantes. Faut-il donc parler de variante uniquement quand il n'y a plus de distinction ni sémantique ni informationnelle ? Certes ces cas existent, mais le principe reste difficilement applicable à l'ensemble des niveaux d'analyse :

La synonymie est incontestable (et du même coup le terme variation pleinement adéquat) quand une alternance entre formes n'affecte pas le sens, en général l'une des variantes étant normée et l'autre (ou les autres) non. L'alternance se réalise soit sous la forme de présence/absence (1/0 : *ne* ou absence de *ne* dans la négation), soit dans la compétition entre éléments supposés équivalents (X/Y), comme dans les paradigmes phoniques ou morphologiques : *il* prononcé [il] ou [i] devant un verbe commençant par une consonne (*il dit/i dit*) ; *e muet* réalisé [ə] ou zéro (*je peux, j'peux*) ; liaison facultative

effectuée ou non (*pas z arrivé/pas arrivé*)... L'idée de variation s'avère ainsi pleinement adaptée pour le champ phonique [...] Mais il est moins aisé de donner des exemples fiables à d'autres niveaux, particulièrement en syntaxe : un actif et un passif disent-ils la même chose ? (seul le passif permet l'effacement de l'agent). Dit-on la même chose avec les différentes formes d'interrogative ? Avec un futur simple ou un futur périphrastique ? (*le train entrera en gare/va entrer en gare*). La question est tout aussi délicate pour le lexique, où *argent, pognon, blé, flouze, thune(s)* et bien d'autres termes ont à peu près le même référent mais ne véhiculent pas le même sens précis ou la même attitude d'émetteur face à son objet de discours. Y a-t-il de vrais synonymes, au-delà de quelques usages régionaux comme *octante* ou *huitante* vs *quatre-vingts* ? (Gadet 2021 : 332)

Cette question de « vraie » synonymie ou non est d'autant plus importante qu'elle peut au contraire être représentative de différentes étapes du changement. En présence de variantes, il ne faut en effet pas minimiser la part de la remotivation sémantique générée par les locuteurs, mentionnée par Henri Frei (1929), comme j'ai pu le voir dans mes travaux sur *malgré que* par exemple (Glikman 2019 [26]), ni le rôle du glissement sémantique, pouvant être à l'origine d'un changement. Mario Barra Jover (2009) introduit ainsi la distinction entre « variantes invisibles » et « variantes visibles », le changement de statut pouvant être la trace d'un changement en cours :

Nous avons affaire à une variante visible lorsque l'alternance est marquée. Le cas le plus typique est celui où l'une des variantes est soit perçue comme « élevée » dans un registre non soutenu soit stigmatisée socialement dans un registre soutenu. [...]

En effet, l'existence de l'élévation et de la stigmatisation générales ou partielles n'est que la manifestation des processus de changement en cours. Ainsi, nous pouvons envisager l'existence d'un patron que nous développerons et interpréterons plus loin. Pour l'instant, nous allons nous limiter à la constatation :

(11) Variante invisible > variante visible partielle > variante visible générale > disparition de l'une des possibilités.

Ce processus peut présenter des cas particuliers assez intéressants qui seront illustrés par les exemples qui suivent. Ainsi :

(11') Opposition paradigmatique > variante invisible > variante visible partielle > variante visible générale > disparition de l'une des possibilités.

(11'') Variante invisible > variante visible partielle ou générale > re-motivation de la variante. (Barra Jover 2009 : 110-111)

Andreas Dufter et Elisabeth Stark (2003) mettent également l'accent sur le problème de l'établissement des variantes en syntaxe :

Aux niveaux morphologique et syntaxique, la situation semble plus claire à première vue : Les possibilités de sélection et de linéarisation relèvent, bien sûr, de la variation discrète. Or, le linguiste désireux de classer les diverses unités ou constructions comme

variantes d'un même type se retrouve soudain devant un autre défi, à savoir le problème des équivalences sémantiques ou fonctionnelles, qui seules permettraient de considérer deux constructions comme variantes d'un même phénomène [...] A y regarder de plus près, les vraies équivalences grammaticales se font presque aussi rares que les synonymes dans le lexique (cf. aussi Blanche-Benveniste 1988). Où que ce soit dans la langue, de «bons» exemples de variation sont donc – du moins en dehors des variantes dialectales – moins nombreux qu'il n'y paraît au premier abord. (Dufter et Stark 2003 : 82)

Cela n'empêche cependant évidemment pas les études sur les phénomènes syntaxiques et morphosyntaxiques, comme le montre l'abondante littérature sur la question (cf. Poplack et Dion 2021, Dufter, Fleischer et Seiler (éds) 2009, entre autres), mais la définition de la variante observée demande alors d'être précisée. Avant même d'aborder les paramètres de variation, il nous incombe d'établir quels sont les éléments en concurrence. Dans le cas des parataxes, on pourrait dire que nous avons bien affaire à une alternance présence / absence, telle que mentionnée par F. Gadet (2021, ci-dessus). J'ai même pu trouver des contextes « comparables » dans le cas des complétives, avec même verbe recteur, même type de contenu (→ 2.2.2.). Mais dans d'autres cas, c'est rapidement plus problématique, pour les relatives par exemple (Glikman 2016 [18]), ou pour les temporelles (cf. « je suis venue j'avais dix ans » vs « je suis venue quand j'avais dix ans »). Cependant, dans cette configuration, la question du statut syntaxique de ces constructions s'est tout de suite posée, de même que celle de leur équivalence ou non avec les énoncés introduits. Dans le cas des hypothétiques, la différence est encore plus flagrante, l'ensemble de la construction étant modifiée : « s'il vient, je pars » vs « il viendrait, je partirais ». Peut-on toujours parler de variantes ?

De plus, il s'agit, on l'aura compris, d'analyse de l'oral dans les exemples traités par Labov, dont l'ancrage est phonétique, mais aussi dans la plupart des exemples cités jusqu'ici. À ses débuts, la sociolinguistique traitait en effet principalement, si ce n'est uniquement, de productions orales, considérées comme le seul lieu où la langue véritable, le vernaculaire, peut être observé :

Le thème de la variation a surtout été mis en œuvre sur des productions orales, en particulier à partir d'oraux ordinaires proches du vernaculaire, qui contrastent avec des écrits, souvent corrigés et standardisés et dès lors plus proches de la norme. Il a toutefois aussi donné lieu à des travaux sur des écrits ordinaires, comme des correspondances privées ou des journaux intimes de peu-lettrés, quand existent de tels documents (entre autres, Martineau, 2012). (Gadet 2021 : 333)

Cela pose d'autant plus la question de la transposition de cette approche à l'écrit, et a fortiori pour des états de langue anciens, accessibles uniquement à travers les sources écrites qu'on en a conservées. Les dernières années ont vu le développement d'une sociolinguistique historique,

que j'évoquais en introduction, à travers la prise en compte de sources relevant des « écrits ordinaires » mentionnés ici par F. Gadet, mais aussi à travers la prise en considération de la visée des textes et de la notion de public, intégrant les facteurs liés aux genres comme paramètres pouvant donner lieu à l'observation de la variation (cf. Banniard 1997, Carlier et Guillot (éds) 2018, → 3). De même, les études diachroniques intègrent depuis longtemps la notion de variation, conceptualisée dans la notion de diasystème chez Coseriu. Le développement des études sur la diamésie, et en particulier de la notion de distance et proximité communicative par Koch et Oesterreicher (2001), permet également de revaloriser la place de l'écrit comme source valable pour l'étude de la variation, et dont l'importance est apparue pour les périodes contemporaines avec le développement des pratiques de l'écrit médié par ordinateur, ou encore les sms (→ 3).

Reste à définir l'objet. L'établissement des variantes en syntaxe passe par la reconnaissance du lien syntaxique, nécessaire pour montrer que les deux réalisations ont bien le même statut syntaxique, et sont donc bien « la même chose » tout en n'étant pas « pareilles ».

2.2. Autour du verbe : établir le lien syntaxique

*Au commencement était le verbe*²²

Ce choix de traduction pour les premiers mots de l'*Évangile selon Jean*, rendant le terme *logos* grec, sans entrer dans les débats sur sa pertinence ou non pour le texte biblique, pris au sens strict, s'applique assez bien à mon cheminement scientifique, avec le recul que je prends aujourd'hui. L'ambiguïté même du terme *verbe* en tant que *mot*²³, mais signifiant aussi *parole* ou *discours*, est également révélatrice de l'évolution de mes travaux, partant de la proposition, des relations construites par le verbe, vers le discours (→ 2.3.). Mais c'est bien autour du verbe (au sens grammatical) qu'ont commencé mes premiers travaux, ma thèse de doctorat portant sur l'étude de la parataxe en ancien français²⁴, à la suite d'une Maitrise²⁵ et d'un DEA²⁶ sur le même sujet, ensuite étendus à d'autres constructions verbales au rattachement flou, comme les incidentes ou « clauses parenthétiques réduites » (Schneider 2007). Pour ces constructions, l'objectif est de définir leur statut syntaxique, afin de les établir en tant que variantes, avant de décrire leurs caractéristiques spécifiques (contraintes ou non) et leurs conditions de réalisations (selon le type de variation, choix libre ou contraint, sur lequel nous reviendrons → 3.). Établir le type de lien qui unit ces constructions à leur « hôte », si l'on a une ou deux « phrases », au sens d'unité maximale de la syntaxe, fait en réalité naître deux problématiques, à partir desquelles émerge une troisième : qu'est-ce que le lien syntaxique (→ 2.2.1.) et comment l'établir, en particulier pour des états de langue anciens (→ 2.2.2.), enfin, qu'est-ce que ces problématiques de délimitation de l'unité maximale disent sur les limites de la syntaxe (→ 2.2.3.) ? Pour chacune de ces questions, au-delà du « comment », il s'agit de comprendre ce qui entre en jeu, quel défi et quels enjeux cela soulève.

²² Traduction du Prologue de l'évangile selon Jean par Augustin Crampon (rédaction : 1864 ; édition : 1894). Source : https://fr.wikipedia.org/wiki/Prologue_de_l%27%C3%A9vangile_selon_Jean, consulté le 15.09.2021.

²³ Le dictionnaire grec - français de Bailly précise toutefois que *logos* ne recouvre jamais le sens grammatical de *mot* que peuvent avoir *onoma* ou *rhema* (Bailly 1895 sv. *logos*).

²⁴ Glikman J. (2009a) *Parataxe et subordination en ancien français. Système syntaxique, variantes et variation*. Thèse de doctorat sous la direction de Annie Bertin et Gerda Hassler, soutenue à Nanterre en 2009, publiée sur le site de la bibliothèque de l'Université de Potsdam en 2011, urn:nbn:de:kobv:517-opus-56784, <https://publishup.uni-potsdam.de/opus4-ubp/frontdoor/index/index/year/2011/docId/5508>

²⁵ Glikman (2004) *La parataxe dans la Chanson de Roland*, mémoire de maîtrise de l'Université de Nanterre, sous la direction d'Isabelle Weill et Annie Bertin.

²⁶ Glikman J. (2005) *La parataxe, définition et évolution*, mémoire de DEA de l'Université de Nanterre, sous la direction d'Annie Bertin.

2.2.1. *Qu'est-ce que le lien syntaxique ?*

La syntaxe est traditionnellement définie comme la composante de la linguistique qui « traite de la combinaison des mots dans la phrase. » (Ducrot et Todorov 1972 : 71, s.v. *Composants de la description linguistique*). Cette conception de la syntaxe est évidemment critiquable et critiquée (cf. *ibid.* 71 sqq.), ne serait-ce que sur la notion de *mot*, pour commencer. Il n'en reste pas moins que travailler en syntaxe suppose, peu ou prou, de s'intéresser aux liens que des éléments entretiennent entre eux :

Il y est question de l'ordre des mots, des phénomènes de rection (accord ou régime) – c'est-à-dire de la façon dont certains mots imposent aux autres des variations en cas, nombre, genre –, et enfin, depuis le XVIII^e siècle surtout, des principales fonctions que les mots peuvent remplir dans la phrase (Ducrot et Todorov 1972 : 71, sv. *Composants de la description linguistique*)

La notion de lien syntaxique part ainsi du principe qu'une 'phrase' n'est pas une simple accumulation de 'mots', mais bien une structure organisée :

Du point de vue syntaxique, la totalité que constitue la phrase n'est pas un pur agglomérat d'éléments, un ensemble (au sens mathématique). Dans un ensemble, si on ne lui ajoute aucune structure particulière, le rapport de l'élément à l'ensemble est identique pour tous les éléments. Au contraire la syntaxe définit certaines relations entre les éléments de la phrase et la totalité de la phrase, relations telles que deux éléments distincts se trouvent la plupart du temps dans une relation différente vis-à-vis de la phrase totale (l'un est sujet, par exemple, l'autre est complément). (Ducrot et Todorov 1972 : 270, sv. *Fonctions syntaxiques*)

L'ordre des mots est directement observable, ainsi que les modifications contraintes portant sur les mots en emploi (les modifications casuelles, l'accord sujet-verbe...), pour lesquelles on parle parfois de *morphosyntaxe* et non de *syntaxe*. À l'inverse, pour les *fonctions syntaxiques* elles-mêmes, le lien syntaxique est invisible dans la plupart des cas. Tesnière parle de « connexion » et le considère comme un véritable élément constitutif de la phrase :

1. L'objet de la syntaxe structurale est l'étude de la phrase. [...] 2. La phrase est un **ensemble organisé** dont les éléments constituants sont les **mots**. 3. Tout mot qui fait partie d'une phrase cesse par lui-même d'être isolé comme dans le dictionnaire. Entre lui et ses voisins, l'esprit aperçoit des **connexions**, dont l'ensemble forme la charpente de la phrase. 4. Ces connexions ne sont indiquées par rien. Mais il est indispensable qu'elles soient aperçues par l'esprit, sans quoi la phrase ne serait pas intelligible. [...] 5. Il résulte de ce qui précède qu'une phrase du type *Alfred parle* n'est pas composée de **deux** éléments 1° *Alfred*, 2° *parle*, mais bien de **trois** éléments, 1° *Alfred*, 2° *parle* et 3° la connexion qui les unit et sans laquelle il n'y aurait pas de phrase. Dire qu'une phrase du type *Alfred parle* ne comporte que deux éléments, c'est l'analyser d'une façon superficielle, purement morphologique, et en négliger l'essentiel, qui est le lien syntaxique. (Tesnière 1976 [1959] : 11-12)

Au-delà du cadre théorique de Tesnière, la lecture de ce passage m'a fortement marquée, et constitue encore pour moi une matière inspirante, avec cette idée de chercher cet élément, qu'on ne voit pas mais qui permet le passage du mot isolé à la *phrase*. C'est dans cet élément invisible que se trouve pour moi l'essence même de la syntaxe, et cette recherche de l'invisible est particulièrement mise en relief par l'étude de la parataxe, qui est, en un sens, une recherche de l'absence²⁷.

La syntaxe, notamment, est le lieu où se jouent un certain nombre de contraintes lors de la production de l'énoncé :

Nous mettrons en évidence un ensemble de contraintes auxquelles doit obéir le locuteur lors de l'énonciation et qui constitue la grammaire de la langue. Parmi ces contraintes, nous verrons que la langue nous impose une structuration hiérarchique qui constitue le cœur de la syntaxe. (Kahane et Gerdes 2022 : 61)

Ces différentes contraintes – la nature de l'élément principal d'une phrase, la partie du discours des dépendants, les unités grammaticales obligatoires comme la conjugaison pour les verbes ou l'article pour les noms, l'accord, le placement des mots les uns par rapport aux autres –, sont des RÈGLES GRAMMATICALES du français. La GRAMMAIRE d'une langue, ce sont toutes les contraintes que nous impose cette langue quand nous parlons.

L'expression même des contraintes grammaticales repose sur les « connexions » entre différentes parties de l'énoncé. Ces connexions constituent, comme nous le verrons, le squelette des RELATIONS SYNTAXIQUES entre les mots, les groupes de mots, mais aussi des unités plus petites que les mots. En mettant en évidence les contraintes que ces unités, que nous appellerons les UNITÉS SYNTAXIQUES, s'imposent les unes aux autres, nous avons mis en évidence l'existence d'une STRUCTURE SYNTAXIQUE. L'étude de cette structure, qui modélise la façon dont les signes linguistiques se combinent les uns avec les autres, constitue l'objet central de la SYNTAXE et donc de cet ouvrage. (Kahane et Gerdes 2022 : 79)

Ces contraintes relèvent donc de la grammaire d'une langue, et se distinguent des choix laissés aux locuteurs lors de la production de son énoncé. Pour établir que deux variantes sont bien des réalisations de mêmes contraintes, d'une même structure syntaxique, il importe de montrer qu'il s'agit bien des mêmes connexions, des mêmes relations syntaxiques.

La question de l'établissement du lien syntaxique et de la reconnaissance des fonctions est un problème récurrent en syntaxe. Les chercheurs des différents cadres théoriques au cours de l'histoire ont visé à apporter une manière stricte et structurante d'établir ce lien, depuis les

²⁷ Recherche qui a inspiré à ma directrice de thèse, Annie Bertin, lors de ma soutenance (2009), une comparaison avec Don Quichotte : comment chercher ce qui n'est pas ? Bien sûr, ici, je cherche ce qui est invisible, mais présent : l'existence d'un lien syntaxique.

traditionnelles questions de la grammaire scolaire « qui fait l'action ? qui mange la pomme ? que mange Pierre ?... », aux tests plus proprement linguistiques, maintenant globalement répandus et acceptés au-delà d'un cadre spécifique, comme les tests de commutation, déplacement, suppression, insertion, pronominalisation, portée illocutoire (portée de la négation et de l'interrogation), et, dans une certaine mesure, le clivage. Kahane et Gerdes (2022) ajoutent également les tests d'« autonomisabilité » illocutoire et prosodique. Aujourd'hui communs, ces tests ont pu, selon les époques et la place qu'on leur accordait, être véritablement constitutifs d'un cadre théorique, comme le distributionnalisme, les grammaires transformationnelles, ou encore, plus récemment, l'approche pronominale (Blanche-Benveniste et al. 1990). Ces tests, tout utiles qu'ils soient, ne sont pas toujours suffisants. Cela est évident pour les états de langue anciens, pour lesquels nous n'avons accès qu'aux textes, aux productions écrites qui plus est, et encore, uniquement celles qui sont parvenues jusqu'à nous (cf. Barra-Jover 2000). Travailler sur l'ancien français et en diachronie d'une manière générale pose ainsi en soi des problèmes méthodologiques et de parti pris théorique. Cependant, je suis convaincue que cela est tout aussi vrai également pour les états de langue récents. Le recours aux enquêtes linguistiques, en complément des corpus, aide en partie (cf. Avanzi et al. 2016 [16]), pour les questions de diffusions des innovations, par exemple, et de leur intégration (→ 3.2.). Les enquêtes linguistiques ne sont cependant pas toujours suffisantes, et sur des occurrences très récentes comme l'extension de *moult* (Glikman et Verjans 2021 [32]) ou de *full* en français dans des contextes variés²⁸ ou encore l'évolution de l'accord de *un/une espèce de* avec le nom qui suit (Verjans et Glikman 2022 [36]), il reste difficile d'appliquer les tests et de juger de leur recevabilité. Dans tous ces cas, le jugement de grammaticalité ne suffit pas, et il nous faut partir des productions pour mesurer la diversité, ou non, des usages. C'est ce que j'appelle l'analyse en système (→ 2.2.2.).

Établir le lien syntaxique permet ainsi de voir qu'il s'agit bien d'une même structure sous différentes réalisations, en d'autres termes, d'établir la variante. Alors seulement, on pourra comparer les différentes réalisations des variantes, étudier leur distribution, et analyser les paramètres de variation.

Dans une grande partie de mes travaux, je me suis intéressée au lien syntaxique non entre les *mots*, mais entre les structures phrastiques, c'est-à-dire entre deux ensembles de mots

²⁸ L'étude des emplois de *full* en français est le sujet de mémoire de master Science du langage de Jonathan Fontaine sous ma direction, en cours.

pouvant chacun, en apparence du moins, constituer une proposition indépendante. Dans les structures qui m'intéressaient, elles devaient en outre comporter un verbe conjugué à un mode personnel, notamment dans mes premiers travaux de recherche sur la parataxe, au sens où on trouve le terme fréquemment employé dans les grammaires de français médiéval (nous revenons ci-dessous sur la question de l'ambiguïté terminologique de ce terme) :

parataxe : juxtaposition de deux propositions sans marquer le rapport de dépendance qui les unit. Elle peut se réaliser selon deux modes :

- par juxtaposition agrégative de deux propositions non hiérarchisées au rapport logique implicite : *Li quens Rollant, par peine e par ahans Par grant dulong sunet sun olifan : Parmi la buche en salt fors li cler sanc* « Le comte Roland, au prix d'efforts infinis, d'une douleur intense, sonne son olifant. Par sa bouche le sang jaillit clair » (*Roland*, 1762-4)

- par non-expression du strument introduisant une proposition subordonnée dépendant d'un verbe régisseur : *Carles li magnes ne poet muer n'en plurt* « Charlemagne ne peut s'empêcher de pleurer²⁹ » (*ibid.* 841) (Buridant 2019 : 1128)

Cette définition, en particulier la deuxième partie, « non-expression du strument introduisant une proposition subordonnée dépendant d'un verbe régisseur » assimile ainsi la recherche de la *parataxe* à une recherche de l'absence. Absence de « strument », mais présence d'un lien syntaxique, nous retournons donc à notre recherche du lien invisible, dont finalement c'est peut-être même le marquage par une *conjonction* qui est « inhabituel », au regard de ce lien syntaxique habituellement invisible³⁰. La notion même d'absence, dans sa formulation, soulève un certain nombre de questions, et bien que ce soit généralement, pour cela, un terme que j'évite, cette formulation a marqué ma réflexion. L'absence doit se définir elle-même. Définie dans le TLFi comme « le fait d'être absent », *absent* étant lui-même défini comme « qui est non présent, manquant ou presque inexistant », la notion d'absence semble induire avec elle son corollaire, la *présence*. Dans la recherche que j'ai menée sur le sujet, cette question a effectivement pris de l'importance, avec le parti pris de parler d'*absence* dans les contextes où l'on peut parler de *présence*, selon le critère de l'existence d'énoncés comparables, qui a même constitué le critère de relevé des occurrences et l'intégration au corpus (nous y reviendrons). Cela induit par là même un contexte d'alternance, à travers l'opposition présence/absence. L'établissement du lien syntaxique repose alors en partie aussi sur cette alternance, rendant nécessaire et suffisant

²⁹ Mot à mot : 'ne peut empêcher qu'il n'en pleure'.

³⁰ Tesnière ainsi considère les conjonctions de subordination comme des « translateurs », permettant à la structure phrastique de remplir une fonction syntaxique (Tesnière 1976 [1959]).

le critère de l'énoncé comparable, et permettant de poser la notion de *variantes*, dont les critères de distribution peuvent être étudiés (→ 3 & 4).

Dans mes premiers travaux, je me suis ainsi intéressée à l'absence de marque explicite de liaison dans le rapport de dépendance entre deux propositions, de type conjonction de subordination (outre ma thèse, Glikman 2008a[2], b[3], 2013 [11], Glikman et Verjans 2013 [13])³¹, voir ex. (1) et (2), ou mot relatif (Glikman 2016 [18]), ex. (3), et aux critères de distribution de ces variantes. Cette absence est jugée particulièrement fréquente en ancien français notamment, dans des contextes où l'on trouve habituellement ces marques :

- (1) bien set *ne se puet or desfandre* (*Renart* 1363, thèse)
[Il sait bien [qu'il] ne peut alors pas se défendre]³²
- (2) bien set **que** *sa mort est juree* (*Renart* 1351, thèse)
[Il sait bien que sa mort est jurée]
- (3) car quant il regarde l'espee *mesire Tristrans tient [...]* (*Tristan* III, 250, 14, cité dans Glikman 2016)
[car quand il regarde l'épée [que] tient Messire Tristan,...]
- (4) Danz Oliver trait ad sa bone espee / **Que** *ses cumpainz Rollant li ad tant demandee*, (*Roland*, p. 114, v. 1367, BFM)
[Olivier a tiré sa bonne épée que son compagnon Roland lui a tant demandée]

Dans les exemples (2) et (4), le lien est marqué explicitement, à la différence des exemples (1) et (3). C'est cette différence qui a même constitué, dans un premier temps au moins, le critère de relevé du corpus d'étude.

La plupart des grammaires de l'ancien français, et certaines grammaires du français moderne, déclarent cette possibilité d'alternance disparue dès la fin de la période médiévale, et mes travaux ont également montré que le nombre d'occurrences dans les textes diminue fortement dès le moyen français. Les travaux sur le français parlé, ainsi que sur d'autres langues, auxquels j'ai été très tôt confrontée (→ 1.), ont cependant montré, d'une part, que ces contextes d'alternance se retrouvent en français moderne (type *je me lève le matin je prends un café vs quand je me lève le matin...*, cf. Choi-Jonin 2005, Deulofeu 1989, Avanzi 2012, entre autres) (→ 4.1.), et, d'autre part, que la présence d'une conjonction de subordination ne suffit pas toujours à établir un lien de subordination (cf. Benzitoun 2006a et b, sur *quand* dans des énoncés

³¹ La question de l'absence dans les constructions hypothétiques a également fait l'objet d'une communication non publiée : Glikman J. (2008) « Les constructions à verbe en 1ère position : structures interrogatives ou système hypothétique ? », colloque international *Diachro IV*, Madrid, Espagne, 22-24 octobre 2008.

³² Sauf indication contraire, les traductions (littérales) sont de mon fait.

comme « quand tout à coup... », Debaisieux (éd) 2013, Corminboeuf 2009 sur *si* « si tu as soif, il y a de la bière dans le frigo »). Ces questions se retrouvent également dans d'autres langues, comme l'allemand³³ et l'anglais (cf. Urmson 1952, Auer 1998, Culicover et Jackendoff 1997, Dehé et Wichman 2010, entre autres), où l'on considère habituellement ces cas comme des subordinations asyndétiques dans la plupart des cas, en particulier pour les propositions relatives (« the man I love »). Le type de lien syntaxique est cependant plus discuté dans le cas des contextes de type propositions complétives, avec les verbes dit « recteurs faibles » ou « parenthétiques » (Urmson 1952, Schneider 2007) qui peuvent apparaître en tête d'énoncé sans marqueur (ex. « je crois j'ai mal lu la phrase », Avanzi 2012), mais aussi en fin d'énoncé ou en position incise, remettant ainsi en question leur statut de verbe recteur (cf. Blanche-Benveniste 1989, Blanche-Benveniste et Willems 2007).

La polysémie³⁴ même de la notion de *parataxe* est révélatrice du problème qui se pose autour du lien syntaxique entre structures phrastiques. Selon le cadre théorique, et même selon les traditions grammaticales³⁵, le terme est en effet employé dans le discours grammatical avec des sens opposés et contradictoires, oscillant entre « subordination asyndétique », c'est-à-dire sans conjonction de subordination et asyndète, pris comme simple équivalent de la juxtaposition (Avanzi, Benzitoun, Glikman 2007 [1]). Cependant, comme j'ai cherché à le mettre en avant dans ma thèse, dès le titre, cet apparent malentendu se révèle n'être pas si anodin qu'on pourrait le croire et montre peut-être bien, au contraire, tout l'enjeu qu'il y a à établir le type de rapport entre les deux propositions. Mon objet d'étude n'était donc pas la *parataxe* au sens de « style paratactique », comme on peut le signaler à propos du texte de la *Bible*, premières réactions que j'ai parfois reçues de non linguistes à qui j'exposais mon sujet de thèse, ni *parataxe* au sens que le terme prend dans l'opposition *hypotaxe / parataxe*, chez Tesnière par exemple, mais bien à la *parataxe* telle qu'on la signale dans les grammaires d'ancien français (cf. Buridant *op. cit.*), que Tesnière nomme de son côté *hypotaxe asyndétique* (Tesnière 1976 [1959]). L'enjeu reste cependant d'établir, en syntaxe, la différence entre les deux, et donc d'établir le lien syntaxique.

³³ Question sur laquelle j'ai fait une communication orale non publiée : Glikman J. (2008) « Comparative and diachronic approach of the non introduced subordinates in French and German », *Fifth International Contrastive Linguistics Conference (ICLC5)*, Leuven, Belgique, 7-9 juillet 2008.

³⁴ Comme, somme toute, pratiquement toute notion ou terme métalinguistique, dès qu'on y regarde d'un peu plus près, ce qui nous avait amenés à organiser le colloque Jeunes Chercheurs de Modyco sur la question, voir Glikman, Mansour et Weiser (éds) (2007) [i].

³⁵ J'ai en effet pu constater dans ma thèse que les grammaires d'ancien français n'avaient globalement pas la même acception du terme que les grammaires du français moderne.

Évidemment, se limiter au contexte d'alternance pouvant admettre la présence d'un mot de liaison amène à se couper de tout un champ d'investigation possible, dans lequel l'absence n'est jamais le corollaire de la présence, mais bien l'inexistant, sans situation d'alternance. Parlerait-on d'absence dans un cas où il ne peut jamais apparaître de mot de liaison ? Par exemple, dans le cas des subordonnées de type infinitives, on peut établir qu'il y a bien un lien de dépendance, mais celle-ci se fait sans mot introducteur. Sans rentrer plus avant dans les débats sur le statut des constructions avec infinitif, qui soulèvent encore d'autres questions, nous retiendrons ici l'exemple traditionnel³⁶ type :

(5) J'entends les oiseaux chanter

Dans cette configuration, on reconnaît traditionnellement que le lien de dépendance est induit par la forme même du verbe de la proposition dépendante, à l'infinitif, et ainsi que la « nominalisation » se fait par le biais même de la forme du verbe, tandis qu'elle passe par le biais d'un translatif, comme par exemple chez Tesnière (1976 [1959]), qui distingue ainsi la translation primaire (avec infinitif) de la translation secondaire, dans le cas des propositions à verbe conjugué à un mode personnel :

(6) J'entends que les oiseaux chantent

Le lien de dépendance syntaxique existe également à différents niveaux, et cette situation d'*absence* peut se retrouver dans d'autres contextes, comme par exemple dans l'alternance possible dans l'histoire du français dans la construction du complément du nom : *la mort du roi* vs *la mort le roi*. De même, il peut y avoir *proposition* hors du verbe fini, avec dépendance ou non, on pensera notamment aux différents travaux sur les énoncés averbaux (voir Lefeuve 1999, Tanguy 2009)³⁷.

La question cependant reste d'établir que ces constructions ont bien les mêmes propriétés syntaxiques que leur équivalent introduit, afin de les définir comme variantes syntaxiques, question que je vais aborder dans la section suivante.

³⁶ Voir par ex. Riegel, Pellat, Rioul (2018 : 584).

³⁷ La question de la parataxe dans les énoncés averbaux a fait l'objet d'une collaboration avec N. Tanguy lors d'une conférence non publiée présentée en 2009 : Tanguy, N., J. Glikman (2009) « Deux énoncés averbaux en parataxe, et hop, une prédication complexe. Les énoncés averbaux peuvent-ils construire des prédications complexes paratactiques en français parlé ? », Conférence de l'Association For French Language Studies (AFLS), Neuchâtel, Suisse, 3-5 septembre 2009.

2.2.2. Comment établir le lien syntaxique ?

Dans le cas de la relation entre « proposition » ou structure phrastique, la *subordination* au sens traditionnel, on a l'habitude d'avoir le lien de subordination marqué par la présence de marqueurs explicites, comme les conjonctions de subordination ou les pronoms relatifs, les « complémenteurs » dans les approches générativistes et les approches en constituants. Ces marqueurs, comme de nombreux travaux contemporains à mes recherches (→ 1) l'ont montré, ne sont pas toujours suffisants, et peuvent au contraire apparaître dans des contextes d'« insubordination » (Benzitoun 2006b, Corminboeuf 2009, 2010, et voir plus récemment Debaisieux (éd.) 2013, Patard 2014, Debaisieux et al. 2019), c'est-à-dire des énoncés où la « subordonnée » ne répond à aucun test de la rection classique (équivalence pronominale, extraction, portée de la modalité du verbe principal...). Je ne citerai ici que l'exemple de *quand soudain* traité par Benzitoun à titre d'illustration, mais plusieurs travaux concernant d'autres marqueurs³⁸ :

(11) *Ils se préparaient à gagner l'intérieur de l'île en remontant le rio, quand, soudain, une ombre se dressa devant eux.* [Verne, Robur]

(11a) *Ils se préparaient à gagner l'intérieur de l'île en remontant le rio, quand?*

Quand n'est pas ici proportionnel à *quand, soudain, une ombre se dressa devant eux* mais à un constituant temporel comme à *trois heures*, d'où l'acceptabilité d'un tel exemple avec cette interprétation spécifique.

(11b) ? *C'est quand, soudain, une ombre se dressa devant eux qu'ils se préparaient à gagner l'intérieur de l'île en remontant le rio.*

(11e) ? *Ils se préparaient à gagner l'intérieur de l'île en remontant le rio, même quand, soudain, une ombre se dressa devant eux.*

(11d) ? *Ils ne se préparaient pas à gagner l'intérieur de l'île en remontant le rio, quand, soudain, une ombre se dressa devant eux [mais quand un orage éclata].*

Ces tests montrent donc que les *quand soudain* ne sont pas régies. (Benzitoun 2006b : 40)

Mes travaux ont cherché à mettre en avant le fait que ces marqueurs ne sont pas non plus nécessaires. Cette absence de marqueur en contexte où l'on trouve habituellement un marqueur a en effet conduit certains auteurs à postuler une absence de lien de subordination, comme dans l'approche de Culicover et Jackendoff (1997, 2005) qui parlent alors de « semantic

³⁸ Jeanne-Marie Debaisieux a notamment travaillé sur *parce que* et *puisque* (Debaisieux 1994), Gilles Corminboeuf sur *si*.

subordination despite syntactic coordination ». Comme l'a montré Deulofeu (1989, 2001), la différence entre subordination, coordination et corrélation n'est pas toujours facile à établir : dans le cas de la corrélation, la dépendance entre les deux constructions verbales se situe-elle au niveau de la syntaxe ou du discours ? Est-ce une dépendance *grammaticale*, ou simplement *discursive* ?

Jusqu'à quel point la notion de complétude (évoquée notamment par Creissels 1995³⁹) ou celle d'« autonomisabilité illocutoire » et « prosodique » proposée par Kahane et Gerdes (2022 : 266 sqq.) peut-elle et doit-elle entrer en jeu, en particulier dans l'étude d'états de langue anciens pour lequel nous ne pouvons pas avoir recours au jugement de grammaticalité des locuteurs⁴⁰ ? La typologie établie par Christian Lehmann (1988), largement reprise par la suite, s'intéresse ainsi à l'intégration d'une structure phrastique à une autre à travers la question de ce qui va lui permettre de remplir une fonction dans une autre proposition, par le biais de l'infinitif, du subjonctif, d'une conjonction ou d'un autre type de marquage. Il propose ainsi une échelle d'intégration, selon le type de marquage et le rôle joué dans la structure intégrante. Denis Creissels (2006b : 193 sqq.) met également en avant l'existence de subordination phrastique sans marque morphologique (au sens de marqueur de subordination comme les conjonctions) :

En l'absence de marqueur de subordination, pour le linguiste, ce sont les manipulations évoquées en 31.3.3 [tests de la portée de la négation et du clivage] qui révèlent qu'une séquence de mots qui pourrait par elle-même être une phrase indépendante fonctionne comme un constituant d'une autre phrase, et pour les locuteurs, la reconnaissance d'une construction par subordination repose sur l'ordre séquentiel et l'intonation. (Creissels 2006b : 193)

D. Creissels propose une typologie de marqueurs, comme l'utilisation de formes verbales intégratives, des marques morphologiques (pouvant être anté- ou postposés) ainsi que la position du verbe, ces différents marqueurs pouvant se cumuler (*ibid.* : 195-196). Pierre Le Goffic (1993 : 502) retient, outre les conjonctions, la possibilité de marquer la subordination par la prosodie, le mode verbal et l'ordre des mots (en particulier la place du sujet clitique dans

³⁹ « On peut ainsi dégager une notion de complétude syntaxique » (Creissels 1995 : 33), définition affinée dans Creissels (2006a) : « C'est en fait un autre type de complétude qui caractérise la phrase. [...] La phrase est donc une combinaison de mots dont la structuration permet l'expression systématique de l'élaboration énonciative d'un contenu propositionnel ; autrement dit, le propre d'une unité phrastique est de participer à un jeu de correspondances régulières avec d'autres unités phrastiques qui expriment une élaboration énonciative différente d'un même contenu propositionnel. » (Creissels 2006a : 12-13)

⁴⁰ Cette question est nécessairement au centre des questionnements méthodologiques lorsque l'on aborde l'étude d'états de langue anciens, ce qui ne veut pas dire qu'il soit impossible de les étudier, comme le rappelle bien Prévost (2011b, chap. 1).

des énoncés tels que « Pleurait-elle, tout le monde se précipitait. » (*ibid.* : 505). Anne-Marie Garagnon et Frédéric Calas, à propos de la parataxe, établissent également une liste de marqueurs du lien syntaxique entre structures phrastiques, incluant la prosodie, l'inversion du sujet, le mode verbal, mais aussi le marquage lexical, par la présence d'intensif, de corrélatif ou d'une locution verbale spécifique (cf. *avoir beau*) ou encore le marquage par la modalité (Garagnon et Calas : 131-136).

Cette pluralité de marqueurs, associée au fait qu'ils ne sont ni nécessaires ni suffisants, amène à la conclusion que la subordination n'est pas dans son marquage, mais bien dans un faisceau d'indices. C'est également le parti que nous avons pris dans le chapitre sur la syntaxe de la phrase complexe dans la *Grande Grammaire Historique du Français* (désormais GGHF) (cf. Combettes et Glikman 2020 [28]) :

C'est donc la notion de subordination qui doit être précisée, ou encore celle d'hypotaxe, qui lui est souvent associée. Ces notions renvoyant à une relation hiérarchisée entre deux éléments, en l'occurrence des propositions, la proposition subordonnée n'est pas autonome mais se trouve intégrée à une structure dans laquelle elle occupe une fonction syntaxique. Ce lien syntaxique peut être envisagé en termes de continuum, en particulier pour les subordonnées circonstancielles, certaines d'entre elles étant plus intégrées que d'autres ; cette caractéristique, comme d'autres points concernant la phrase complexe, a connu une évolution au cours des siècles, la tendance générale étant celle d'une plus grande intégration. Ce continuum est fondé sur des indices essentiellement syntaxiques, qui font apparaître le degré de dépendance de la subordonnée par rapport au contexte. Sans adopter une position théorique relevant de telle ou telle école, on n'exploite cependant pas ici le cadre bien représenté dans l'analyse traditionnelle, celui de la généralisation de la dichotomie proposition principale / proposition subordonnée, dans la mesure où cette distinction tend à faire dépendre toute subordonnée de la phrase matrice, alors que, dans de nombreux cas, c'est un terme particulier qui se trouve en relation avec la subordonnée, que ce soit le verbe comme recteur des complétives, ou le GN comme antécédent des relatives, etc.

Parmi les propriétés qui permettent d'établir un continuum, on peut citer les faits de position ; l'appartenance à la valence verbale, révélée en particulier par la pronominalisation ; l'emploi des modes ; la portée des modalités énonciatives telles que la négation ou l'interrogation. Contrairement à la pratique courante de la grammaire traditionnelle, la présence d'un « marqueur » de subordination n'est pas prise en compte ici comme critère de subordination. Ceci est d'autant plus important que la présence et le rôle de ce marqueur ont subi une évolution importante au cours du temps. Il existe tout au long de l'histoire du français, et en particulier dans l'ancienne langue, des cas de subordination sans marqueur (ou hypotaxe asyndétique). Ces cas de subordination sans marqueur sont parfois décrits, en particulier dans les grammaires d'ancien français (AF), sous le terme de « parataxe », alors que cette notion recouvre généralement les relations non hiérarchisées, comme la coordination ou la juxtaposition. L'absence de mot de liaison, qui donne à la construction cet aspect de « juxtaposition », a pu favoriser la confusion terminologique, mais il n'en reste pas moins que certaines de ces constructions constituent de véritables subordonnées, le lien de dépendance restant visible à travers d'autres marques comme l'ordre des mots, l'emploi des modes, ou

encore des indices de rection, l'étroitesse de ce lien pouvant là aussi s'établir suivant un continuum. (Combettes et Glikman 2020 [28] : 1338-1339)

À la question des marqueurs s'ajoute ici la notion de *continuum*, les subordonnées pouvant être plus ou moins intégrées (cf. Lehmann 1988), point que j'avais peu abordé dans mes travaux, consacrés aux subordonnées « régies » complétives et relatives. Ma participation à la GGHF m'a fourni l'occasion de travailler sur des subordonnées « non régies » (terminologie retenue dans la GGHF). J'y ai rédigé le chapitre sur les subordonnées finales, dont il a fallu établir le degré d'intégration. J'ai utilisé pour cela les critères des modes et temps verbaux et celui de l'intégration dans une unité supérieure, critères auxquels j'avais déjà eu recours :

Les subordonnées exprimant le but font partie des propositions circonstancielles, non régies. Elles ne sont pas appelées par une unité de la proposition principale. Plusieurs indices laissent toutefois penser qu'elles ont un degré élevé d'intégration à la phrase :

– elles sont majoritairement au subjonctif (c'est déjà le cas en latin), mode de la subordination ;

– l'ensemble ainsi formé de la principale et de la subordonnée finale peut être lui-même enchâssé dans une unité supérieure dans laquelle il occupe une fonction, ce qui peut s'interpréter comme la preuve d'une cohésion syntaxique forte, dès la période médiévale et jusqu'à la période moderne :

si feust mes conseilz / Que hautement se mariast, / Affin qu'a sa mort nous laissast / Seigneur de sa propre lignie. (Griseldis, 1395, p. 7, v. 176)

Avant que finir ce propos des Perroquets, me resouvenant d'avoir leu en une Cosmographie qu'afin que les serpens ne mangent leurs oeufs, ils font leurs nids pendus à une branche d'arbre je diray ici en passant, qu'ayant veu le contraire en ceux de [...] (LéryBrésil, 1578, p. 222)

je renvoie une fois pour toutes le lecteur à l'oeuvre même, en l'avertissant pourtant qu'une telle étude n'est nullement nécessaire pour que ce que j'expose dans ce travail-ci soit assez entendu. (AlainBeauxArts, 1920, p. 7)

Contrairement à d'autres subordonnées circonstanciennes qui tendent à une intégration plus forte à la phrase au cours du temps, dans le cas des subordonnées finales, on observerait plutôt une évolution vers une plus grande autonomie, en particulier pour certaines constructions en *de façon que* ou *de sorte que*, dont l'interprétation est à la limite de l'expression du but et se confond avec l'expression de la conséquence ou de la manière, et qui tendent à se détacher de la phrase, jusqu'à constituer une proposition autonome :

[...] qui vient alourdir le tout par ses humeurs de veuf et l'odeur d'urine aigre que dégagent ses enfants (au moral et, hélas, surtout au physique). De sorte que je fais l'anachorète au sein du cénobitisme. Cette dame, charmante pour moi, me fait porter des plateaux à des heures variables (SartreLettres2, 1951, p. 352)

[...]

La présence du subjonctif dans la subordonnée, la possibilité d'intégration dans une unité supérieure, et le peu d'exemples relevés en position détachée ou autonome, sont autant d'indices permettant de dire que les subordonnées finales, bien que non régies, tendent à être des circonstancielle fortement intégrées à la proposition. (Glikman 2020a [29] : 1360-1361)

Dans ses travaux, Bernard Combettes montre en outre l'importance de la prise en compte de certains phénomènes de structuration textuelle, y compris les phénomènes de coréférence, pour l'analyse des propositions périphériques, comme ici à partir de phénomènes de redénomination et l'intégration des subordonnées circonstancielle, à propos du latin de la *Chronique* de Frédégaire :

La deuxième mention d'un référent nous semble par exemple constituer un phénomène digne d'intérêt. L'articulation subordonnée + principale, dans une hiérarchisation de phrase complexe, correspond à un palier de traitement, particulièrement en ce qui concerne la dénomination des référents, palier où s'impose l'utilisation d'anaphores pronominales. L'ouverture de l'énoncé par une *cum* P ne rend pas obligatoire, dans l'expression d'une coréférence, la reprise par un anaphorique. [...]

Cet exemple peut, toutes proportions gardées, être rapproché de certains énoncés de l'ancien français, qui, sans être fréquents, sont néanmoins bien représentés ; ainsi, dans cet extrait de la *Chanson de Roland* :

(15) Quant veit li reis le vespres decliner, / Sur l'erbe verte descent li reis en un pred,
/Culchet sei a tere, si priet Damnedeu (*Chanson de Roland*, v. 2447-2449)

'Quand le roi voit tomber le soir, sur l'herbe verte le roi descend dans un pré, il se couche à terre et il prie Dieu'

Ce type de redénomination nous semble bien faire apparaître l'autonomie de la proposition initiale, qui sert de cadre à une nouvelle séquence narrative dotée d'une organisation propre. (Combettes 2018 : 380)

J'ai encore trop peu exploité cet indice dans mes recherches jusqu'ici, malgré les différents travaux que j'ai menés avec mes collègues dans le cadre du projet MC4 puis DEMOCRAT sur les chaînes de référence, où nous nous sommes plutôt intéressés aux rapports entre les chaînes de référence et le genre textuel (→ 2.3.4.). Dans notre étude sur le rapport entre chaînes de référence et structuration textuelle (Capin D., J. Glikman, C. Schnedecker & A. Todirascu 2021 [33]), nous avons étudié les fonctions des redénominations dans un texte, mais sans entrer dans leur rôle dans la mesure de l'intégration phrastique de subordonnées. Un cas particulier peut cependant en être rapproché. En effet, il est fréquent, dans les textes médiévaux notamment, que la redénomination intervienne après un épisode de discours rapporté, y compris quand le référent est déjà directement présent dans le discours cité ou en incise. La redénomination, ici identifiée comme marquage du discours rapporté, donne aussi à voir le statut « à part » du discours direct dans le texte.

La redénomination signale également des changements énonciatifs, comme dans (14), où les mentions de *le roy* entourent littéralement le discours rapporté (ici simplement indiqué par des crochets) ce qui nous donne à penser que *le roy*₂ signale un retour à la narration :

(14) Quant vint ***le lendemain au matin le roy***₁ fit escrire une lectre ainsi que s'ensuit, et au marge de dessus estoit escript, « De par le roy », et la lectre contenoit ainsi : « [...] » Et au dessus des lectres estoit escript : « ***Aux barons et peuple d'Espagne.*** » ***Le roy***₂ fit incontinant despescher *ung courrier* auquel furent baillees lesdictes lectres, et *luy* commanda ***le roy***₃ que *il* fist extreme diligence. (*Jehan de Paris*, XVe s.) (Capin et al. 2021 [33] : 101)

Dans plusieurs de mes travaux j'ai ainsi soulevé la question des différents indices de subordination possibles, des frontières prosodiques (Glikman 2009 [5]), à la présence du subjonctif, d'un élément corrélatif, la présence de la négation, l'insertion d'une autre proposition, ou encore l'ordre des mots (Glikman 2010a [6]). Je reviens dans les paragraphes qui suivent sur la pertinence et les limites de ces différents indices.

Les frontières prosodiques

À la suite de Christiane Marchello-Nizia (1978, 1993), j'ai interrogé le rapport entre frontière de vers et frontière de proposition (Glikman 2009 [5]). Marchello-Nizia (1978) montre sur un corpus de 300 propositions prises dans des textes décasyllabiques anciens que frontières prosodiques (incluant fins de vers et césures) et frontières syntaxiques coïncident dans plus de 80% des cas, ce qui lui permet d'avancer l'hypothèse dans son article de 1993 que la frontière prosodique suffit à indiquer la frontière de proposition, et constitue donc une condition nécessaire à la réalisation de la parataxe (pris ici au sens de subordination asyndétique). Dans mon article, j'ai cherché à montrer que le rapport entre frontières de vers et frontières syntaxiques ne constitue un critère ni suffisant ni nécessaire pour établir le lien syntaxique ou son absence, ni établir qu'il y a parataxe. En effet, il y a bien correspondance dans un certain nombre de cas (mes chiffres rejoignent à peu près ceux de Marchello-Nizia, à l'exception du texte d'Adenet *Berte aus grans piés*, œuvre du 13^e siècle en alexandrins, qui ne présente que 69% de cas de coïncidence). Cependant, les types de liens entre deux frontières prosodiques peuvent être de natures différentes : frontières de phrases, de propositions liées explicitement ou non, ou encore frontière de groupe syntaxique. De même, on peut trouver des cas de non coïncidence des frontières :

- (7) *Quiās le guant // me caïst en la place* (*Roland* 764, cité dans Glikman 2009 [5])
[[tu] as cru [que] le gant // me tomberait des mains sur place ?]

Les frontières prosodiques peuvent cependant constituer un indice de regroupement syntaxique, quand la séquence apparaît dans un même hémistiche, sans pour autant préjuger du type de rapport :

- (8) Si li reis voelt, // *prez sui por vos le face.* (Roland 295, cité dans Glikman 2009 [5])
[si le roi veut, je suis prêt à le faire pour vous]

L'étude que j'ai menée alors se focalisait principalement sur la question du marquage de la parataxe, ainsi que sur les conditions de sa réalisation, l'explication prosodique ne me semblant pas suffisante. De ce fait, malgré quelques éléments de comparaison, j'ai principalement observé les contextes de cas de parataxe. Il manque évidemment à cette étude de regarder de manière systématique les frontières prosodiques dans un échantillon de textes. Appliquer des calculs statistiques auraient aussi permis de vérifier de manière plus sûre les intuitions sur le relevé, méthode que je n'appliquais pas alors, et que je n'ai intégrée dans mes travaux que grâce à ma collaboration ultérieure avec Nicolas Mazziotta (voir Glikman et Mazziotta 2013 [12]). Les résultats de cette étude ont été réinvestis dans mes travaux suivants, la piste du regroupement évoquée dans l'article met en avant un cas d'incise, que j'ai convoqué de nouveau dans mes travaux sur les parenthétiques :

Nous avons cependant relevé des cas semblant faire exception et contredire ainsi notre hypothèse : il s'agit des cas dans lesquels apparaît une incise de discours de type « ce dit X », comme on a pu voir par exemple en (24) :

- « Tybert, » ce dist Morans, // « garde # sor li ne fier, (Berthe 604)
[Tibert, dit Moran, // prend garde à ne pas la frapper]

Cependant, ce type d'exemple ne va pas à l'encontre, selon nous, de notre hypothèse. En effet, les incisives sont des éléments qui ont comme particularité de s'appuyer sur une proposition « hôte », dans laquelle elles peuvent venir s'insérer à n'importe quelle place. On reste donc en quelque sorte dans le domaine intra-propositionnel, et non dans des cas de propositions « indépendantes ». (Glikman 2009 [5])

Dans mon étude sur les constructions type « *ço crei* » 'je crois' (Glikman 2010b [7]), j'ai ainsi poursuivi l'étude de la position dans le vers comme indice possible de rattachement syntaxique, notamment pour résoudre des questions de portée, finalement, davantage que d'intégration.

Dans le cas de ces constructions, la question se pose non seulement du statut syntaxique des constructions non introduites par rapport à leur équivalent introduit (9), (10), mais également de leur lien avec la construction incidente, ces mêmes verbes pouvant apparaître en position incise (11), (12).

- (9) « Ha! rois Pepins, » fait ele, # « je croi mar vous vi né, | Quant on me veut murdrir delez vostre costé. » (Berthe 425-26) (in Glikman 2010b)
[Ah! Roi Pépin, fait-elle, # je crois [que] je vous vis né pour mon malheur, / quand on veut me tuer à vos côtés]
- (10) je crois j'ai mal lu la phrase (PFC, in Avanzi 2009 [2012])
- (11) E si n avrez, ço quid, # de plus gentilz. (Roland 150) (in Glikman 2010b)
[et vous en aurez, je crois, # de plus nobles]
- (12) son père à lui a été euh élevé je crois à Paris (PFC, in Avanzi 2009 [2012])

En effet, Urmson (1952), Thompson et Mulac (1991), pour l'anglais, se sont intéressés à ces verbes (dits épistémiques, comme *croire*) pouvant apparaître soit en position initiale suivie d'une proposition subordonnée complétive (avec omission possible, voire fréquente, du subordonnant), soit en position finale ou incise, incident à la proposition « hôte » (*je crois qu'il fait beau ; il fait beau, je crois*). Stefan Schneider s'est également intéressé à ces constructions dans les langues romanes, en posant la notion de « reduced parenthetical clauses », ayant une fonction de « mitigator » de l'énoncé (Schneider 2007). Pour le français, Claire Blanche-Benveniste (1989) introduit la notion de « recteur faible » pour ces verbes, mettant en avant l'équivalence de leurs différentes constructions (voir aussi Claire Blanche-Benveniste et Willems 2007). Ainsi, pour ces auteurs, ces verbes ne sont pas vraiment recteurs quelle que soit leur position, d'où la possibilité de l'absence de subordonnant.

Pour ces constructions, la question peut ainsi se poser, en l'absence de subordonnant notamment, de leur portée et position syntaxique. En effet, dans le système de l'ancien français (et finalement même du français parlé, comme on peut le voir dans les corpus oraux, au moment où nous devons résoudre la question du rattachement lors de l'étiquetage syntaxique des corpus oraux, comme les vocaux), il n'est pas toujours facile d'établir si le verbe se situe en incise, ou en position initiale, avec une portée sur ce qui suit. En position recteur, il reste possible de parler de subordination, un certain nombre d'indices montrant que le comportement comme régissant une subordonnée n'est pas le même qu'en incise, comme je l'ai montré : portée de la négation, de la modalité, diversité des personnes et des temps et mode, que je vais développer ci-dessous. En effet, les études précédemment citées sur les recteurs faibles et parenthétiques ont montré que leurs emplois comme incidente épistémique (et comme recteur faible suivi d'une complétive introduite ou non), étaient restreints à la première personne, pour *croire* par exemple, au présent de l'indicatif.

Après ma thèse, j'ai mené plusieurs travaux sur les verbes parenthétiques, y compris en impulsant la recherche collective à ce sujet, à travers l'organisation d'une journée d'étude organisée à Nanterre avec Mathieu Avanzi en 2010 « Entre rection et incidence : des

constructions verbales atypiques ? », qui a donné lieu au numéro thématique de la revue *Linx* 61 (Glikman et Avanzi (dir.) 2009 [2012]⁴¹ [ii]), et au colloque international *Les « Verbes Parenthétiques » : Hypotaxe, Parataxe ou Parenthèse ?* organisé à Nanterre en 2012 avec Mathieu Avanzi, Sophie de Pontonx et Stefan Schneider, après lequel nous avons publié en 2015 le volume *Parentheticals Verbs* (Schneider, Glikman et Avanzi (éds) 2015 [iii]). Je reviens ici sur la question des indices permettant d'établir le statut de ces constructions, et donc de les définir comme variantes pouvant entrer en concurrence.

Dans mon article publié dans le volume 60 de la revue *Travaux de Linguistique*, « La rection verbale entre micro- et macro-syntaxe » (Glikman 2010b [7]), j'ai réinvesti les rapports « prosodie » (tels qu'on peut les observer sur les états anciens de langue à travers la versification) - syntaxe pour distinguer les différentes réalisations de ces verbes.

L'étude de la position dans le vers m'a ainsi permis de faire des hypothèses sur le rattachement (rattachement à gauche dans l'exemple 13, à droite dans l'exemple 14) de ces constructions verbales, ainsi que sur leur statut syntaxique :

(13) E si n avrez, ço quid, # de plus gentilz. (*Roland* 150)

[et vous en aurez, je crois, # de plus nobles]

(14) que ceanz a assez de vitaille / ne cuit devant un an me faille (*Renart* 1785-86)

[car ici il y a assez de vivres, je ne crois pas [qu'] il m'en manque avant un an]

Notre étude a ainsi permis d'observer une correspondance régulière entre analyse prosodique et analyse syntaxique, les cas de rattachement à gauche correspondant à une analyse comme construction incise, et ceux de rattachement à droite à une analyse comme construction rectrice. La notion de rattachement prosodique, et par là, le critère de position, semble donc pouvoir servir d'indice supplémentaire pour fonder l'analyse syntaxique de ces verbes, entre rection et incidence. (Glikman 2010b [7] : 40)

Cette hypothèse est appuyée dans l'article par l'analyse d'indices complémentaires : présence du subjonctif *faille* dans la subordonnée (14), portée de la négation (14), (15) et de l'interrogation (7 repris sous 16), et d'énoncés comparables introduits (17), d'une part dans les cas de rection (avec rattachement à droite), alors que dans les cas d'incidentes, le verbe se trouve en position insérée entre des éléments dépendants, et l'on trouve des énoncés équivalents construits explicitement comme incident, avec « si com » (19) (tous les exemples sont tirés de Glikman 2010b [7]) :

⁴¹ Le numéro est effectivement paru en 2012 (la journée elle-même ayant été organisée en 2010), mais la revue ayant connu des retards de publication, et les numéros étant associés à une année, le numéro 61 est listé comme un numéro de l'année 2009.

- (15) ne quidez mie je vos hace. (*Renart* 803)
[ne croyez pas que je vous haïsse]
- (16) Quïas le guant me caïst en la place, | Cum fist a tei le bastun devant Carle ? (*Roland* 764-65)
[tu as cru que le gant me tomberait des mains, comme fit avec toi le bâton devant Charles?]
- (17) ne cuit qu'autre merci en aies (*Renart* 741)
[je ne crois pas que tu en aies d'autres grâces]
- (18) E si n avrez, ço quid, # de plus gentilz. (*Roland* 150)
[et vous en aurez, je crois, # de plus nobles]
- (19) que vos onques n'eüstes cure | de Renart ne de son deduit | ne de s'amor, si con je cuit. (*Renart* 194-96)
[que jamais vous n'avez eu cure / de Renard ni de son plaisir / ni de son amour, ainsi que je crois]

De plus, dans ce que j'ai appelé l'« observation du comportement », on retrouvait une diversité de personne, temps et mode dans la position initiale, tandis que les positions de « rattachement à gauche » se trouvaient effectivement restreintes à la première personne au présent de l'indicatif. Cela conduit à l'hypothèse qu'il n'existerait pas d'incise antéposée, la position initiale conservant, malgré la notion de « recteur faible », des propriétés de verbe recteur.

Ces différents indices devaient donc théoriquement permettre de distinguer les énoncés suivants, si l'on applique strictement la notion de rattachement prosodique comme codant la syntaxe, bien qu'ils puissent également être considérés comme ambigus et être analysés des deux manières (Glikman 2010b : 38) :

- (20) Eüstes vos tel livroison, | ce croi, tiex .c. cous i eüstes, (*Renart* v. 1716-17)
[vous avez eu un tel traitement, je crois, cents coups de cette nature y avez eu]
- (21) Dist Oliver: « sire cumpainz, ce crei, | De Sarrazins purum bataille avoir » (*Roland* v. 1006-07)
[Olivier dit: « Sire compagnon, je crois, nous pourrons avoir bataille avec les Sarrazins]

Les frontières prosodiques peuvent donc constituer un indice pour établir les variantes, mais ne constituent pas pour autant une condition nécessaire ni suffisante. Prises en considération avec d'autres critères, elles peuvent cependant permettre de faire des hypothèses qu'on peut étendre à d'autres énoncés.

La cooccurrence avec ce

Reste alors la question de la présence du pronom *ce*, qu'on retrouve dans l'une comme l'autre construction, et qui pose la question de la saturation de la valence du verbe et, par conséquent, du rôle que sa présence peut jouer dans l'établissement de contextes comparables ou non. Cette

question se pose d'autant plus que dans les états anciens de la langue, les complétives peuvent être introduites par *ce ... que*, pouvant être conjoint ou disjoint.

Le rôle de *ce* et son évolution soulèvent de nombreuses questions, que ce soit dans les constructions en « *ce ... qu-* » complétif ou relatif, en fonction objet, attribut ou sujet, ou en emploi autonome. Entrent en effet en compte plusieurs problématiques liées à *ce* mais pas uniquement. C'est le cas des constructions impersonnelles, dans lesquelles *ce* peut apparaître, mais qui méritent un traitement à part, de même que la question des structures clivées ou pseudo-clivées, ainsi que les questions d'accentuation et de prosodie (en particulier comme explication du remplacement de *ce* par des formes « pleines » comme *ceci* ou *cela*), qui sont liées à l'évolution de l'accent en français. Plusieurs auteurs traitent spécifiquement de la question de *ce que* introducteur de complétive, comme Magali Rouquier (Rouquier 1990, et plus récemment sur les structures clivées 2007, 2014), mais aussi Claude Muller (1997), qui pose à partir du passage *ce ... que* à *que* l'hypothèse d'une origine corrélatrice de la complémentation propositionnelle. D'après Bernard Combettes (1996) et Michel Pierrard (1996), l'alternance *que/ce que* se met progressivement en place, *ce que* étant peu à peu restreint aux contextes après préposition, notamment dans la complémentation indirecte.

Dans le cas de la position comme recteur de complétive, j'évoque dans l'article (2010b [7]) la possibilité de trouver une complétive introduite avec la présence du pronom (phénomène bien connu de *ce que* introducteur de complétive) en contexte comparable, ce qui me permet de dire que la présence de *ce* ne suffit pas à remettre en question l'analyse, mais qu'on peut au contraire postuler qu'il existe une forme réalisée de manière asyndétique de *ce que* introducteur de complétive (Glikman 2010b [7] : 39-40) :

(22) Ço sent Rollant que la mort li est pres (Roland v. 2259)

['ce' sent Roland que la mort est près de lui]

(23) Ço sent Rollant de sun tens n'i ad plus. (Roland v. 2366)

['ce' sent Roland [que] de son temps est épuisé]

L'analyse repose évidemment sur la notion de contexte comparable, qui est le principe de base de l'étude de la variation et de l'établissement des variantes, mais qui n'est pas évidente à établir en syntaxe, et plus encore dans l'étude des états de langue anciens (→ 2.1.). Dans le cas présenté ici, il s'agit bien du même verbe, et les deux énoncés sont construits de manière relativement symétrique pour qu'on puisse les juger équivalents, mais le contexte n'est toutefois pas strictement identique, si par contexte on entend vraiment la réalisation lexicale, ou même encore les groupes syntagmatiques en présence (sujet nominal réalisé dans un cas mais pas dans l'autre, etc.). De même, nous avons vu que *ce* peut apparaître dans différentes constructions, y

compris de manière autonome, et, inversement, les verbes peuvent avoir différentes réalisations. Pour les différentes réalisations des verbes, comme je le développerai ci-dessous, une analyse du « comportement » en « système » permet de faire des hypothèses relativement précises de ses constructions fréquentes. Enfin, on pourrait aussi postuler qu'il s'agit là non d'une complétive, mais bien de deux énoncés juxtaposés, avec un *ce* cataphorique, comme le propose Claude Muller. Nous l'avons évoqué, les éléments de coréférence peuvent aussi servir d'indice, comme Bernard Combettes l'a montré dans ses travaux. Il utilise notamment la notion de cataphore pour montrer l'intégration progressive des constituants périphériques. En effet, il postule que la cataphore (en particulier dans un complément périphérique antéposé) est un phénomène très tardif dans l'histoire du français. Dans un énoncé du type « quand il vient, Paul mange », Combettes s'appuie sur le fait que le pronom *il* ne renvoie pas à *Paul* avant le 19^e siècle, pour montrer que ces constituants ne sont pas strictement intégrés à l'origine :

La relation cataphorique ne devient possible qu'au moment où, dans l'énoncé complexe, le degré de dépendance entre les propositions est assez élevé pour permettre le jeu de facteurs syntaxiques. (Combettes 2017 : 127)

Le cas de *ce...que* introducteur de complétive, avec *ce* disjoint, est un peu différent des contextes analysés par B. Combettes, mais il reste possible de s'en inspirer et d'en tirer une analyse de la relation, sous forme de syllogisme peut-être, mais qui peut tout de même constituer une piste à explorer : le fait que la cataphore n'est possible qu'en présence d'un degré de dépendance élevée amène à confirmer d'une part le lien de dépendance élevée entre *ce* et *que* disjoint (ce que la tradition ne remet guère en question), mais aussi entre *ce* et tout autre élément propositionnel en lien cataphorique avec *ce*.

Je n'ai jusqu'ici malheureusement pas suffisamment exploité les travaux de B. Combettes dans cette optique, mais j'ai tout de même effleuré du doigt une distinction d'analyse selon la saturation référentielle de *ce*. Dans une communication (non publiée) présentée au colloque *ComplémentationS* en 2010⁴², j'ai tenté d'analyser de manière plus systématique la présence de ce pronom dans cette position et les conséquences sur l'analyse de ces structures. La question était de savoir s'il s'agissait d'une « double complémentation » ou d'un cas de dislocation, dans des occurrences telles que citées en (22, 23). Les conclusions de cette étude étaient que l'analyse syntaxique peut différer selon la saturation référentielle de *ce*. Dans le cas d'un *ce* cataphorique,

⁴² Glikman J. (2010) « La double complémentation : double marquage ou dislocation ? », Colloque *ComplémentationS*, Santiago de Compostela, Espagne, 20-23 oct. 2010.

la saturation référentielle est à chercher dans la suite de l'énoncé, elle se fait de manière intraphrastique, on a affaire à une construction de type « ce ... que » disjointe, y compris en l'absence de « que » (cf. ex. 24). Dans le cas d'un *ce* anaphorique, il reprend des éléments antérieurs, c'est un pronom plein, on peut considérer qu'il sature une place valencielle (cf. ex. 25).

- (24) **Ço** ad juret li Sarrazins espans, / Se en reregarde troevet le cors Rollant / Cumbatrat sei a trestute sa gent (Roland v. 611-13)

[Le Sarrazin d'Espagne a juré (cela) [que] s'il trouve Roland à l'arrière-garde, il le combattra avec toute son armée]

- (25) Quant li baron oïrent **chou**, si disent a l'empereur qu'il s'en retourneroient (Clari, p. 57)

[Quand les barons entendirent cela, ils dirent à l'empereur qu'ils s'en retourneraient]

Avec Tom Rainsford⁴³, nous avons abordé à nouveau la question dans une communication (non publiée) présentée au colloque *Diachro 6* en 2012. Dans cette intervention, nous interrogeons d'une façon plus générale la question du *ce* objet et de sa quasi disparition dans la plupart des contextes (en dehors de quelques contextes figés, comme « ce faisant »...), à travers l'étude des occurrences de *ce* dans le corpus SRCMF. Nous avons alors proposé une typologie des emplois de *ce* et des pistes pouvant expliquer leurs disparitions selon le type. Nous avons alors proposé de distinguer cinq types : annonce de discours direct (26), avec une complétive (27), en incise (28), en emploi seul anaphorique (29), avec une relative (30), observant pour chaque type les constructions en concurrence.

- (26) **Ço** dist Rollant : « cornerai l'olifant » (Roland v. 1702)

[(ce) dit Roland : « je sonnerai l'olifant »]

- (27) **Ço** dist li reis que sa guere out finee (Roland v. 705)

[(ce) dit le roi que sa guerre a pris fin]

- (28) Ja, **ce** cuit, l'ore ne savra / Qu'esperance traï l'avra (Yvain v. 2661)

[Jamais, (ce) [je] crois, on ne saura que l'espoir l'aura trahi]

- (29) « Mes ostiex est apareilliez a vos toz, entrez en la joie qui ja ne faudra. » Et quant il avoit **ce** fet, si venoit a l'ainzné des .ii. chevaliers et disoit : [...] (Gaal, p. 131, l. 13)

[... Et quand il eut fait (cela), il vint à l'ainé des deux chevaliers et dit...]

- (30) Si ferai **ce** que mes cuers vialt (Yvain v. 5169)

[je ferai ce que mon cœur veut]

⁴³ Glikman J. et T. M. Rainsford (2012) « *Ce* objet ou attribut : étude diachronique », *colloque Diachro 6*, Louvain, 17-20 octobre 2012.

Les données montraient qu'un déclin des emplois de *ce* dans la plupart de ces contextes s'observe dès l'ancien français. Dans le cas du déclin de « *ce que* » comme introducteur de complétive, le fait avait été signalé déjà par Magali Rouquier (1990). Les résultats n'étaient cependant alors pas assez aboutis pour justifier une publication dans l'un comme dans l'autre cas, et j'ai mis de côté à ce moment-là l'étude de *ce* objet, laissant la question ouverte. Pour l'analyse du lien syntaxique, vu les différentes configurations possibles et l'existence de *ce que* comme introducteur de complétive, la présence de *ce* n'est en tout cas pas un critère suffisant pour rejeter l'hypothèse de l'intégration syntaxique de la proposition non introduite.

Dans ma présentation des modes d'introduction des complétives dans la *Grande Grammaire Historique du Français* (Glikman 2020b [30]), j'ai opté pour l'analyse développée par Combettes (1996) et Pierrard (1996), celle d'une mise en place progressive de la répartition entre *que* et *ce que* introducteur de complétive. Le problème s'est en effet reposé lors de mon travail préparatoire sur les complétives pour la GGHF. En particulier, cela pose la question de l'analyse des complétives en fonction sujet, type « *que tu viennes me fait plaisir* », et de leur version 'détachée' « *ça me fait plaisir que tu viennes* », qui n'est alors plus analysée comme sujet dans la tradition grammaticale, mais comme « suite de formes impersonnelles » par exemple chez Riegel, Pellat et Rioul (1994 : 492-93). Nous avons dû aborder la question de l'analyse de ces constructions lors de l'annotation syntaxique du corpus SRCMF. En ancien français, la question peut se poser d'autant plus que le sujet pronominal peut très bien ne pas être réalisé. Avec Nicolas Mazziotta, Beatrice-Barbara Bischof et Tom Rainsford, qui étions alors tous annotateurs dans le projet, nous avons étudié toutes les réalisations de *ce*, en particulier sujet des constructions impersonnelles mais pas uniquement, dans le *Tristan* de Bérout (Mazziotta et al. 2012 [9]). Nous avons repris dans cette analyse la distinction entre *ce* sans complétive, anaphorique, et *ce* cataphorique avec complétive :

1. *Ce* peut être associé, sémantiquement et syntaxiquement, à une complétive, continue ou non.
2. Sans complétive, en particulier dans une incidente, *ce* est généralement anaphorique. La résolution de l'anaphore est externe à la proposition où *ce* se trouve et l'intégration syntaxique de la cible de l'anaphore est absente.
3. Avec complétive (cas peu fréquents dans l'extrait étudié), *ce* est cataphorique et la résolution de la cataphore est interne à la proposition où *ce* se trouve, avec un degré d'intégration syntaxique supérieur. En grammaire dépendancielle, la complétive dépendra de *ce*. En grammaire de constituants, *ce* et *que* + P seront les constituants immédiats d'une structure éventuellement discontinue.
4. Enfin, *ce* n'est pas obligatoire, mais fait partie du paradigme *ce/ce que* + P/*que* + P, qui contient un membre explicitement référenciel (*que* + P). (Mazziotta et al. 2012 [9] : 52)

Cette étude nous a permis d'avancer que *ce* sujet n'est jamais impersonnel, à l'exception de ses occurrences avec un type de verbe spécifique, « venir » temporel, qui est le choix d'annotation que nous avons également retenu dans le corpus SRCMF :

Ce est toujours analysé comme [SjPer](#) [sujet personnel], y compris quand il entre dans des tournures dites impersonnelles ou présentatives. Cependant, dans la construction très particulière en *venir* suivi d'une indication temporelle (*au matin...*), le pronom *ce* est assimilé au *SjImp* [sujet impersonnel]:

- (10) quant ce vendra au departir [...]. (TristBérM4 p. 83 v. 2686) [\[Afficher l'analyse\]](#)

(Bischof et al. 2012 [10], <http://srcmf.org/fiches/SjImp.html>, §5 *Ce*)

Ceci nous a conduits à distinguer les structures en verbe + complétive des structures à construction attributive + complétive, qui ont reçu deux analyses différentes dans le corpus SRCMF. Dans le premier cas, l'analyse de la complétive reste la même, que *ce* soit exprimé ou non, tandis que, dans le second cas, l'analyse de la complétive diffère selon la réalisation ou non d'un autre élément, *ce* ou *il* :

§2. Tournure impersonnelle du type: verbe + "complétive"

Que *ce* soit avec *il*, *ce* ([SjPer](#) [sujet personnel], voir [§5](#)), ou *un sujet non exprimé*, la "complétive" est toujours analysée comme un [Cmpl](#) [complément] du [V](#) [verbe] :

- (3) "S'il li plect que je muire | ce sera sauvement a m'ame." (SGraalIVQuesteM p. 254 v. 32-33) [\[Afficher l'analyse\]](#)

- (4) que s'il avient | que ele arive en estrange païs l'en savra bien | qui ele est. (SGraalIVQuesteM p. 242 v. 28-30) [\[Afficher l'analyse\]](#)

- (5) Or covient que je te die qui sont li dui | chevalier qui erent en lor compaignie, li ainznez de cels | qui les sivoit (SGraalIVQuesteM p. 136 v. 17-19) [\[Afficher l'analyse\]](#)

§4. Tournure impersonnelle du type: construction attributive + "complétive"

Dans le cas des tournures impersonnelles à attribut du sujet du type *avis*, *ainsi* construites avec une "complétive", si le [Sj](#) [sujet] est *il* ou *ce* ([SjPer](#), voir [§5](#)), la "complétive" est assimilée à un [ModA](#) [modifieur attaché] du sujet (voir [ModA^{§10}](#)) En revanche, si le [SjImp](#) est *non exprimé*, la complétive est considérée comme le [SjPer](#) de la construction *estre avis*, *estre ainsi*.

Dans ces tours, les mots *avis* ou *ainsi* est toujours [AtSj](#) [attribut du sujet] (voir aussi [AtSj^{§5}](#)).

- (8) "Puis qu'il est einzi, fet li rois, que nos avons | si bone remembrance de vos.' or me dites | s'il vos plect ou je lairé cest escu [...]." (SGraalIVQuesteM p. 34-35 v. 20-22) [\[Afficher l'analyse\]](#)

- (9) et li ert | avis que ele disoit (SGraalIVQuesteM p. 254 v. 19-20) [\[Afficher l'analyse\]](#)

(Bischof et al. 2012 [10], <http://srcmf.org/fiches/SjImp.html>)

Suite à mes travaux préparatoires pour la GGHF sur les complétives, j'ai repris et poursuivi l'étude des complétives en fonction sujet, présentée dans une communication au colloque *Diachro IX* en 2019⁴⁴, encore non publiée. Je ne suis pas encore satisfaite de l'aboutissement de la recherche, la question étant connexe à plusieurs autres points, comme l'évoque la figure 1 ci-dessous :



Figure 1: les complétives en fonction sujet et questions connexes

En effet, l'étude des complétives⁴⁵ en fonction sujet suppose de mener une réflexion sur ce que signifie la fonction sujet. Elle nécessite aussi d'analyser la concurrence avec les structures de type *le fait que* :

- (31) Que tu viennes me fait plaisir
- (32) Ça me fait plaisir que tu viennes
- (33) Le fait que tu viennes me fait plaisir

Les structures en *le fait que* permettent ainsi de suppléer les complétives dans les contextes où une complétive n'est pas la construction « préférée » (→ 4), comme je l'ai proposé

⁴⁴ Glikman J. (2019) « Les propositions complétives sujet dans l'histoire du français » *Diachro IX* · le français en diachronie, Université de Salamanca (Espagne), les 28, 29 et 30 mars 2019, Facultad de Filología, Palacio de Anaya, Plaza de Anaya, Salamanca.

⁴⁵ Le terme *conjonctive* serait d'ailleurs peut-être plus approprié dans ce contexte, comme me l'a fait remarquer Nicolas Mazziotta.

également avec la concurrence *malgré que* / *malgré le fait que* (Glikman 2019 [26]). *Le fait que* agit ainsi comme un double translateur (si l'on reprend la terminologie de Tesnière 1976 [1959]), par une sorte de double nominalisation à travers la présence d'une conjonction et d'un nom, « fait », dont on peut considérer que dépend la subordonnée. Ces structures en « le fait que » constituent ainsi des structures d'évitement, les propriétés nominales de « fait » lui permettant d'apparaître dans des contextes où l'on attend (ou préfère) un argument nominal.

De même, j'ai eu à aborder de nouveau l'opposition *que* / *ce que* dans mes travaux sur la concurrence *de façon (sorte, manière) que* / *de façon (sorte, manière) à ce que*, présentés notamment au colloque *Diachro X* à Paris en 2022 (communication encore non publiée)⁴⁶. Dans ce cas, toutefois, la présence de *ce* s'explique de façon tout à fait attendue par la présence de la préposition *à*. Ce qui est moins attendu est le développement d'une construction redondante, potentiellement par analogie avec la construction infinitive *de façon à*. Mes différents travaux menés jusqu'ici m'incitent cependant à privilégier l'hypothèse d'un développement d'un subordonnant « redondant » en *à ce que*, dont on observe également l'expansion dans la complétude verbale (cf. « je demande (à ce) qu'il vienne »), et dont le développement va dans le même sens que l'extension des constructions en *le fait que* (→ 4).

Dans l'un comme dans l'autre cas, on a l'impression d'un marquage redondant de la subordination. Ce marquage redondant en français moderne peut symboliser une étape supplémentaire dans l'évolution générale de la subordination et des relations phrastiques, vers de plus en plus de marquage explicite, comme j'ai pu le voir dans mes travaux sur la parataxe et sur les incidentes. Dans l'un comme l'autre cas, la question du genre discursif entre cependant en compte (→ 3 & 4). Cela pose aussi la question de la grammaticalisation potentielle de *le fait que* et *à ce que* en subordonnant. L'évolution du système actuel du français tendrait donc d'une part vers des subordonnants de « moins en moins » subordonnants, des subordonnées pouvant être de plus en plus autonomes, et d'autre part, des marqueurs de subordination de plus en plus redondants et explicites. Cette analyse rejoint le mouvement général de l'évolution de la phrase complexe que nous avons proposé avec Bernard Combettes dans la GGHF : un mouvement vers une intégration plus forte des subordonnées, le développement des marqueurs explicites et non

⁴⁶ Glikman J. (2022) « Variation et mécanismes d'analogie dans l'évolution des locutions conjonctives formées sur *façon, manière* et *sorte* », Colloque international *DIACHRO X Le français en diachronie*, 31 mai au 2 juin 2022, Sorbonne Université.

ambigus, mais aussi un phénomène que Combettes qualifie de « décondensation de la phrase complexe » :

En ce qui concerne l'écrit, on observe une évolution vers ce qu'on désigne habituellement comme « phrase complexe », caractérisée par des relations de dépendance plus ou moins fortes et une clôture syntaxique et sémantique. [...]

Il faut signaler aussi l'évolution du système des subordonnants. Le développement de la lecture autonome, peut-être, et la diversification des textes, sûrement, amènent un nouveau besoin : apporter plus de précision là où l'emploi de *que* ou des autres conjonctions simples entraînait parfois des ambiguïtés. Le système se diversifie et se développe en MF, période pendant laquelle apparaissent de nombreuses créations [...]. De la même manière, les phénomènes de subordination sans marqueur tendent à disparaître de l'écrit normé, tendant ainsi à plus de marquage des propositions de manière explicite et non ambiguë.

Une tendance moderne pourrait sembler aller à l'encontre de l'évolution générale : les phénomènes d'ajout après le point, mouvement inverse de l'intégration et du resserrement de la phrase et qui ne concerne que la langue écrite dès lors qu'il y a ponctuation, même si ces ajouts d'après coup existent également à l'oral, où ils sont signalés par une prosodie particulière. [...] Ce type de construction, qui correspond à un processus de décondensation de la phrase complexe, se développe en FContemporain, du moins dans certains genres textuels. On retrouve ainsi, toutes proportions gardées, une structuration symétrique à celle que présentaient certains énoncés de l'AF de type subordonnée temporelle + « principale », avec une proposition initiale en « ajout ». L'autonomisation de la « subordonnée » en ajout après le point semble l'évolution naturelle de ce tour, par exemple pour des propositions introduites par *sauf que* ou par *d'autant que*, qui n'apparaissent quasiment plus qu'isolées par une ponctuation forte. Bien qu'il s'agisse de l'écrit, ce type de structuration serait à replacer dans le cadre très général de la tendance à la « décondensation » de l'énoncé, qui est une caractéristique de l'évolution du français parlé (➔ 43.2.5). (Combettes et Glikman 2020 [28] : 1340-1342)

J'ai également traité la question de *ce* dans les incises dans mes travaux menés avec Stefan Schneider sur l'évolution des constructions parenthétiques en français, dans lesquels nous avons étudié la mise en place progressive du patron moderne « je VERBE » (Schneider et Glikman 2015 [15], Glikman et Schneider 2018 [23]). Dans Schneider et Glikman 2015, nous avons ainsi interrogé les trois pistes d'origine possible des parenthétiques modernes : issus d'une proposition rectrice (théorie proposée par Thompson et Mulac 1991 a et b notamment, avec le passage de « I think that P » à « P, I think » via « I think P »), d'une proposition subordonnée (type « comme je pense », hypothèse retenue pour l'anglais par Brinton), et enfin une origine juxtaposée, hypothèse soutenue par la présence du démonstratif *ce*. Nous avons tenté de répondre à la question pour le français à travers l'analyse des différentes constructions de l'ancien français au français de la Renaissance, qui peuvent se résumer selon nous à trois types dans l'histoire du français : les constructions avec démonstratif (« *ce*-parentheticals »), celles

avec une conjonction explicite (« *comme*-parentheticals ») et celles ayant le format moderne (« modern parentheticals »).

We can, however, evaluate the probability of the three hypotheses in the light of our current knowledge: subordinate clauses provide the less probable source for modern parentheticals, superordinate clauses are more probable as source, while juxtaposed sentences seem to offer the most plausible explanation. (Schneider et Glikman 2015 [15] : 185)

Concernant le rôle de *ce*, les données montraient à la fois son importance dans les premiers textes, et sa disparition progressive au profit du développement de la forme « je + V » :

But if *ce*-parentheticals originated modern parentheticals we have to explain the transition from *ce croi* to *je croi*. Both types of parentheticals co-occur from the early period. In the corpora examined there are no clues indicating an evolution through intermediate states as the following: *ce croi* → *ce croi je* → *je ce croi* → *je crois*. The construction *je ce croi* is not attested. Hence, the direct replacement of *ce* with *je* seems more plausible. In general, the transition from *ce croi* to *je croi* fits in well with what we know about the preferred argument structure of Old and Modern French. (Schneider et Glikman 2015 [15] : 184-185)

De même, les données montraient aussi le développement, en termes de proportion, des constructions introduites par une marque explicite, type « comme ». Le développement des structures en « comme » rejoint ce que nous avons vu précédemment pour la subordination et le développement du marquage explicite et non ambigu, et duquel on pourrait rapprocher d'autres phénomènes bien décrits par ailleurs, comme la disparition de la construction directe du complément du nom type « la mort le roi ». Ce développement semble en particulier lié au développement de la prose et à certains genres textuels, comme nous l'avons montré dans Glikman et Schneider 2018 [23] sur les constructions parenthétiques :

La forte association des CP [constructions parenthétiques] dans les textes anciens au vers et leur présence dans le DD et dans les interventions du narrateur pour la période médiévale confirment qu'elles peuvent être assimilées à « des marques représentatives de la proximité communicative et de l'oral représenté » (Guillot 2009). Les CP de type moderne, sans marque explicite, existent dès les premiers textes, et sont même pour plus de la moitié des occurrences relevées présentes dans la période médiévale, dans les textes à proximité communicative. Les CP avec marques explicites, représentées à toutes les époques, semblent connaître moins de contraintes à leur réalisation, et deviendront en effet majoritaires après la Renaissance (Schneider 2012). On voit ainsi à travers l'étude des CP ce qui a déjà été observé par ailleurs, un changement qui se produit à partir du moyen français et de la Renaissance dans le rapport aux textes et à l'écrit, et dans le rapport de l'écrit à l'oralité, et le renforcement des marques explicites dans la prose, dès les premiers textes en prose. Ces marques vont se développer dans le même mouvement que la prose se développe elle aussi. La présence de marques explicites dans les CP à l'écrit a pu faire penser que la CP moderne était de création récente. Nos recherches montrent qu'elle existe depuis l'ancien français, mais qu'elle reste cependant

fortement associée à l'oralité. Ce serait donc la proximité des textes anciens à l'oralité qui aurait permis d'observer la CP moderne, le développement de l'écrit en prose privilégiant les constructions marquées. [...] Ici, il semblerait que l'évolution des CP est davantage conditionnée par les contraintes stylistiques liées au développement de la prose. (Glikman et Schneider 2018 [23] : 332-333)

Nous reviendrons sur ces questions dans la partie suivante (→ 3), où nous aborderons plus généralement la question de la variation.

L'analyse de *ce* est ainsi loin d'être simple, quand on prend en considération le nombre de structures différentes dans lesquelles *ce* peut apparaître, qui posent toutes des questions spécifiques. Elle fait encore l'objet de recherche de la part de nombreux chercheurs, comme on peut le voir par exemple dans le colloque qui y a été consacré à Strasbourg en 2018⁴⁷, et aux deux numéros de revues coordonnées par Annie Kuyumcuyan et Anne Theissen sur la question (Kuyumcuyan & Theissen éd. 2020a et b). Elle est également revenue à différentes reprises dans mes travaux, sous différents angles d'approche, et mériterait encore d'être approfondie.

Pour l'établissement des variantes, la présence de *ce* ne peut ainsi pas, à elle seule, servir de critère définitif, plusieurs cas pouvant se présenter. Là encore, la présence de *ce* est à considérer en fonction de ses contextes de réalisation.

Autres critères

En ce qui concerne les autres critères, dans ma contribution au volume *La Parataxe* (Béguelin, Avanzi & Corminboeuf éd. 2010) (Glikman 2010a [6]), j'ai proposé une liste de marqueurs possibles de la subordination, dans l'objectif annoncé d'« établir des critères formels de reconnaissance de la parataxe » (titre de l'article), « parataxe » étant pris ici au sens de subordination asyndétique, tout en mettant en avant le fait que ces marqueurs sont les mêmes en ancien français, mais aussi en latin, moyen français et français moderne : marquages de type lexical, à travers la présence d'un certain type de verbe recteur ou d'un élément corrélatif, le mode verbal, mais aussi la présence de la négation, l'insertion d'une proposition en *se* ('si') entre les deux structures phrastiques et un changement dans l'ordre des mots, en particulier le verbe en première position. L'idée était de montrer que les énoncés « il vient je pars » et « il viendrait je partirais » n'ont pas tout à fait le même statut syntaxique, les modes et temps verbaux du second énoncé contraignant davantage une analyse en termes de dépendance que dans le premier.

⁴⁷ Voir programme : https://lilpa.unistra.fr/fileadmin/upload/DUN/lilpa/Actualites/Programme_Ce_disant.pdf.

Le principe de base de cet article était cependant problématique dès le départ, posant un problème méthodologique. En effet, j'annonçais dès le début de l'article que mes relevés portaient « uniquement sur des cas de parataxe avérés » (Glikman 2010a [6] : 357), dont les caractéristiques devaient alors servir à étendre l'analyse à d'autres cas de parataxe. On comprend alors la circularité de l'analyse : les caractéristiques qui nous permettent de faire l'analyse sont des éléments qui contraignent l'analyse comme subordonnée, mais n'en sont pas pour autant nécessairement l'origine. Cette circularité était évoquée dans l'article même :

Il existe certes ici un risque d'effet circulaire lié au fait que ces contextes contraignants constituent justement, en partie du moins, les indices qui influent sur l'analyse en tant que construction paratactique. Il n'en reste pas moins qu'il est possible de s'interroger sur la nature de ces particularités et de les considérer comme des éléments marqueurs. (Glikman 2010a [6] : 357-58)

Ainsi, j'ai évoqué la possibilité d'un marquage lexical, par la présence d'un type de verbe particulier, ou d'un élément corrélatif. Cependant, ce n'est pas parce que ce verbe peut s'observer dans ce type de construction qu'il en est le moteur ou que la construction n'est pas possible avec d'autres verbes. On retrouve pourtant ce genre d'analyse pour le français moderne, avec des structures comme « avoir beau » ou « à peine » qui appellent une deuxième proposition (cf. Conti et Béguelin 2010, Gachet 2010) :

- (34) Notre amour mûrissait **à peine que**, déjà, / Tu brûlais mes chansons, crachais sur mes violettes, / Et faisais des misères à mes chats... [Georges Brassens, *Le mauvais sujet repent*, 1952 <Frantext>] (exemple tiré de Gachet 2010)

La question du type de verbe est en premier lieu problématique, puisque la valence verbale n'est pas définie a priori et ne suffit donc pas pour invoquer la question de la saturation de la position objet. Cependant, cela est utile quand il s'agit d'étudier toutes les réalisations d'un verbe dans un texte, pour émettre à partir de cela des hypothèses sur ses constructions possibles, ce qui correspond à ce que j'appelle l'analyse en système, qui consiste à analyser, dans un espace donné, le fonctionnement d'un item. Bien sûr, pouvoir en déduire un lien de dépendance syntaxique est valable dans ce cas uniquement pour les compléments régis (dans mon cas les complétives), et non des compléments non régis (comme les subordonnées circonstancielles étudiées notamment par Combettes comme nous l'avons évoqué).

La présence d'un élément corrélatif fait davantage consensus, et on la retrouve dans les descriptions grammaticales des différents états de langue. Dans mon article sur les relatives (Glikman 2016 [18]), cet indice s'est également révélé important :

- (35) Dans cevaliers, se Diex m'aït, se je quidaisse que vous fuissies si pres de moi comme vous estes, je n'eüsse pas dit de teles paroles ai-je dites. (*Tristan* I, 101, 28) (in Glikman 2016 [18])

[Seigneur chevalier, si Dieu m'aide, si je pensais que vous étiez si près de moi comme vous l'êtes, je n'aurais pas dit de telles paroles [que celles que] j'ai dites]

J'ai inclus dans la catégorie « corrélateur supérieur » dans cet article également les déterminants définis, dont la définitude est saturée par la présence de la relative :

- (36) car quant il regarde l'espee mesure Tristrans tient [...] (*Tristan* III, 250, 14) (in Glikman 2016 [18])

[car quand il regarde l'épée [que] tient Tristan...]

Les temps et modes utilisés dans chacune des propositions constituent également un argument récurrent d'indice d'intégration, en particulier la présence du subjonctif, que j'ai déjà évoquée précédemment, et qui est traditionnellement défini comme un mode fortement associé à la subordination, bien que cela ne soit pas ses seuls emplois possibles, dans les différents états de langue. J'y ai également inclus la question de l'impératif du verbe potentiellement recteur, dans des énoncés tels que :

- (37) Et sachiés bien, se je pooie, | Service et honor vous feroie, (*Yvain*, 999)

[Et sachez bien [que], si je pouvais, je vous rendrais service et honneur]

Bien entendu, la présence de la négation permet notamment de reconnaître la subordination, par l'examen de sa portée, comme le rappelle Creissels (2006b : 190-191). Plutôt qu'un marqueur, il faudrait donc plutôt la traiter comme un indice. C'est d'ailleurs à ce titre qu'elle est convoquée dans l'article sur les liens entre rattachement prosodique et lien syntaxique que l'on vient d'évoquer (Glikman 2010b [7]). La relation de cause à effet peut être biaisée : ce n'est pas parce que les structures sont subordonnées que la négation est présente, mais bien parce qu'elle est présente que nous *reconnaissons* ces structures comme subordonnées. C'est d'ailleurs souvent le cas : la plupart de ces marques ne sont pas des marques de parataxe mais des éléments qui contraignent l'analyse syntaxique, comme dans le cas des relatives, où la présence d'éléments sémantiquement « vides », comme les pronoms dans les patrons de type « n'i ad celui ... » pousse à analyser la proposition suivante comme une relative venant spécifier le pronom (voir Glikman 2016 [18]) :

- (38) Il avoit deus ymages jetés de coivre en forme de femme, si bien faites et si natureument et si beles que trop ; si n'i avoit **chelui** n'ait bien vint piés de haut. (Clari XCI 1-4) (cité dans Glikman 2016 [18])

[Il avait deux représentations de cuivre en forme de femme, si bien faites, si naturelles, si belles, que [c'était] trop ; il n'y en avait aucune [qui] ne fasse bien vingt pieds de haut]

Ainsi, la présence de la négation, outre le fait que l'étude de sa portée sert d'indice d'intégration, n'est pas une marque en soi, comme j'ai pu le dire dans Glikman (2010a [6]). L'observation de sa présence systématique, ou non, peut cependant permettre de vérifier les contextes de distribution et de voir s'il s'agit d'une contrainte dans la distribution des variantes. La présence de la négation n'est ainsi pas une contrainte pour la réalisation de la parataxe, puisque dans les contextes où elle est systématique, on la retrouve aussi dans les réalisations introduites (cf. Glikman et Verjans 2013 [13]). Cette présence systématique de la négation concerne aussi bien la principale, par exemple avec les verbes *pooir muer* : *Carles li magnes ne poet muer n'en plurt* « Charlemagne ne peut s'empêcher de pleurer » (cité par Buridant 2019 : 1128), que la subordonnée, par exemple avec *garder* : « *Tybert, » ce dist Morans, // « garde # sor li ne fier, (Berthe 604) [Tibert, dit Moran, // prend garde à ne pas la frapper]* (cité dans Glikman 2009 [5]). Je reviendrai dans le chapitre suivant sur ces questions de distribution et de variation (→ 3.).

La présence d'une proposition en *se* (« si ») est intéressante, mais là encore, elle montre peut-être davantage l'intégration de la proposition en *se* elle-même que le lien de subordination avec le verbe de la première construction :

- (39) *car il set bien, se la reïne le set, ele l'en savra maugré si grant a son escient qu'il ne trouvera jamés pes envers li. (La Mort le roi Artu, p. 10, § 14, lg 20-21, cité dans Glikman 2010a [6])*

[car il sait bien [que], si la reine l'apprend, elle lui en voudra tellement à son avis qu'il ne trouvera jamais de paix auprès d'elle]

- (40) *Et sachiés bien, se je pooie, / Serviche et honor vous feroie, (Yvain, 999, cité dans Glikman 2010a [6])*

[et sachez bien, si je le pouvais, je vous rendrais service et honneur]

L'hypothèse avancée ici est que la présence d'un marqueur de subordination explicite constitue déjà un marqueur suffisant de frontière de proposition. Dans Glikman (2008b [3]), j'ai ainsi retenu cet indice, la position de la *se*-P, pour distinguer entre une analyse comme verbe recteur ou comme incise du verbe *savoir* dans des énoncés comme (41) et (42) (les ex. (41) à (44) et leur traduction littérale sont tirés de Glikman (2008b [3]) sur *Le Roman de Renart*) :

- (41) *Sachiez, se vos le me donez, / bien vos sera guerredonnez (1509-10)*

[Sachez [que], si vous me le donnez, [cela] vous sera bien rendu]

- (42) *S'il est panduz, sachiez de fi, / tuit si parent ierent honi, (1443-44)*

[S'il est pendu, sachez assurément, tous ses parents seront honnis]

Cette analyse est appuyée dans l'article par la comparaison avec d'autres « énoncés comparables » dans le même texte : un énoncé avec pronom anaphorique, dans la position incise (ex. 43) et la présence d'un équivalent avec expression du *que* dans la position rectrice (ex. 44) :

(43) Par toz les sainz de Belleant, / se mes en oi ne tant ne quant, / sachiez le bien, sanz demorance / panduz sera a une branche. (1465-68)

[Par tous les saints de Bethléem, si dorénavant [j']en entends parler en quoi que ce soit, sachez le bien, sans délai, [il] sera pendu à une branche]

(44) bien sachiez que, se Renart vit, / tel le conperra q'ainz nel vit. (1769-70)

[sachez bien que, si Renart vit, celui-là le payera cher qui ne l'a jamais vu]

La présence et la position d'une proposition en *se* peut ainsi constituer un indice pour l'établissement du statut des propositions, en association avec d'autres critères. L'apport de cet indice pour l'analyse de l'intégration des *se*-P reste cependant à faire.

Analyse en système

Comme nous l'avons vu jusqu'ici, j'ai souvent associé l'examen des critères pour établir le statut syntaxique à l'analyse d'autres éléments contextuels permettant de confirmer ou infirmer une hypothèse. C'est ce type d'analyse globale, par le recours à l'analyse du contexte et à l'étude « du comportement en système », que j'appelle l'analyse en système, que j'avais commencé à développer au chap. 5 de ma thèse.

Par *système*, j'entends la vérification en corpus des contextes comparables. Ces contextes nous permettent, en particulier pour les états anciens de langue, de se substituer aux tests linguistiques en donnant à voir les éventuelles modifications possibles. En effet, il n'est pas possible de faire de tests sur les états anciens de langue, mais il est possible de vérifier si des énoncés comparables peuvent attester des transformations possibles ou non d'une construction donnée. Cela pose bien sûr le problème de savoir définir ce qu'est un énoncé comparable (→ 2.1.). Cela me semble cependant possible dans un contexte donné, par exemple autour d'une forme spécifique, comme une forme verbale, pour laquelle on peut relever toutes les occurrences et en comparer les réalisations. C'est une méthode que j'ai plusieurs fois appliquée par la suite.

Dans mon article sur le cas des relatives (Glikman 2016 [18]), l'analyse de toutes les réalisations de la séquence « avoir (a) non » dans Clari, qui apparaît dans des constructions ambiguës entre relativisation et parenthèse, m'a ainsi permis de faire l'hypothèse de deux schémas distincts, le premier ((nom donné) + avoir + a non) présentant davantage d'indices d'intégration syntaxique que le second ((objet/personne) + avoir + (a) non + (nom donné)). Là encore, bien sûr, se pose la question de ce qui est comparable (→ 2.1.). Dans mon étude sur

malgré que et *à cause que* par exemple (Glikman 2019), je n'ai finalement pas comparé les réalisations de ces items avec leurs « équivalents », ce qui m'avait, à juste titre, été reproché par un relecteur anonyme (→ 2.3.2).

Bien sûr, les outils modernes permettent beaucoup plus facilement de regarder l'ensemble des constructions d'un mot, d'une catégorie grammaticale dans un corpus large, voire d'une fonction, comme dans le corpus annoté en syntaxe SRCMF ou dans celui que nous sommes en train de constituer dans le projet sur les sms vocaux. Dans le cas de corpus déjà annoté, bien évidemment, il faut avoir conscience que les choix d'annotation sont déjà une analyse en soi (→ 3.2.). Pour le chapitre sur les complétives de la GGHF, j'ai ainsi utilisé le corpus SRCMF pour extraire tous les verbes régissant des complétives. Cependant, le fait d'avoir commencé, durant ma thèse, par des dépouillages manuels, a été très formateur et m'a certainement permis de mieux prendre conscience de certains phénomènes propres à un texte, de prendre en compte le contexte, ce que des extractions peuvent parfois masquer. En particulier pour des états de langue anciens, se familiariser avec le « système » global du texte⁴⁸ reste important. Cela permet de penser le texte comme un système clos : vu les variations et la variabilité de l'ancien français, les conclusions sont toujours à prendre avec précaution sur la question de comment, ou non, les généraliser. Étudier le fonctionnement à l'intérieur d'une unité texte (dont on connaît en outre la possible mouvance (Zumthor 1981⁴⁹) ou variance (Cerquiglini 1989)), permet au moins de dire qu'on a un fonctionnement dans un système qui se donne comme tel. En termes méthodologiques, cela amène évidemment à penser la description au sein d'un ensemble comparable, à l'échelle d'une seule œuvre parfois, dont les conclusions restent à étendre peu à peu, en prenant garde aux généralisations hâtives qu'on peut en faire. Mais n'est-ce pas là le propre de toute démarche sur corpus, comme le rappelle Christiane Marchello-Nizia :

Enfin, comme tout travail sur corpus, l'analyse procède nécessairement et de façon systématique à des allers et retours entre la formulation de règles de fonctionnement de plus en plus précises et fines, et leur test sur de nouveaux énoncés ; entre la description

⁴⁸ Par *texte*, j'entends ici la production étudiée. Cette production peut correspondre au texte édité, ce qui a été mon cas lors de ma thèse. La majeure partie des bases de données actuellement existantes sont également composées de textes édités, dont les variantes (de manuscrits) peuvent constituer des variations (l'absence de l'apparat critique est d'ailleurs l'une des remarques souvent adressées à l'encontre des bases de données constituées de numérisation d'édition). Elle peut également correspondre au texte du manuscrit, qu'on peut vouloir étudier à lui seul (le manuscrit comme témoin), ou encore aux productions d'un même locuteur (ou scripteur) dans un corpus oral (ou écrit tel que le corpus 14). L'idée principale est ici de délimiter une unité, restreinte mais qui se donne comme un tout cohérent. En cela, l'unité « texte » (pris dans ce sens élargi) me paraît révélatrice d'un état de langue.

⁴⁹ Voir Lagorgette 2019 pour une synthèse de la notion et de son évolution dans l'œuvre de Zumthor.

empirique dont on induit des règles de fonctionnement, et une transformation de ces régularités en hypothèses grammaticales généralisables que l'on projette sur un autre corpus. Le but visé est en effet de mettre au jour une grammaire qui prédise tout ce que l'on trouvera non seulement dans les énoncés du corpus, et c'est la moindre des choses, mais aussi dans les énoncés de corpus de plus en plus vastes de la même langue. (Marchello-Nizia 1995 [2010] : 19)

Avant de tester les hypothèses sur des corpus de plus en plus larges et d'en généraliser les résultats, la prise en compte de l'unité texte, comme un système qui se donne comme tel, permet ainsi de prendre la mesure du fonctionnement des faits linguistiques en contexte et peut servir de critère pour l'analyse.

Jusqu'ici, j'ai traité de la question d'établir le lien syntaxique, afin de confirmer le statut effectivement semblable de deux formes de réalisations syntaxiques, permettant de les traiter comme deux variantes d'une même structure. Dans d'autres cas, la question de la mise en évidence des variantes passe par d'autres critères et se pose différemment, point sur lequel je reviendrai (→ 2.3.). Au préalable, je souhaite faire dans la section suivante un bilan intermédiaire sur les questions et perspectives liées d'une manière plus générale à cette problématique de l'établissement du lien syntaxique.

2.2.3. Questions, applications et perspectives

Établir l'unité maximale

Au commencement, donc, il y eut le verbe. Le verbe, défini traditionnellement, pour le français du moins, comme pouvant porter les informations de personne, de nombre, de temps, d'aspect et de mode. Le verbe comme pouvant constituer le noyau réactionnel d'une « phrase ». Ceci amène évidemment à la question de la définition même de la phrase, en tant qu'unité maximale, qui est l'une des grandes questions de l'analyse syntaxique. Question des plus anciennes, à la fois des plus simples et des moins évidentes, comme le rappelle Gilles Siouffi en ouverture de son volume sur l'« histoire de la phrase française » :

À cette question simple, tellement simple qu'elle en est désarçonnante, on aurait envie de répondre comme le faisait Audistin d'Hippone (IV^e siècle) lorsqu'on lui demandait ce qu'était le temps : « Si on ne me pose pas la question, je le sais. Si on me la pose, je ne le sais plus... » (Siouffi (éd.) 2020 : 7)

Cette question occupe une part importante dans le domaine, comme on peut le voir dans les différentes publications sur le sujet (Siouffi (éd.) 2020 phrase, Beguelin et al. (éds) 2019, Gautier 2006, entre autres) et la part que cela prend dans les grammaires. Il ne s'agira pas ici de faire un état des lieux de cette question, chose qui dépasserait largement le cadre de cette

synthèse. Travailler en syntaxe, sur le lien syntaxique, sur les rapports entre propositions, tout comme constituer des corpus annotés en syntaxe, suppose cependant nécessairement d'avoir une position sur cet objet. Je me contenterai dans cette section de faire le point sur les questions qui sont apparues au cours de mes travaux, les positions que j'ai pu prendre et comment ces positions ont pu évoluer, ainsi que sur les perspectives que cela pourrait ouvrir.

Définir la *phrase*, au sens syntaxique, pose en effet des problèmes au niveau théorique, dans son rapport au modèle et son lien avec la définition même du domaine :

On définit traditionnellement la syntaxe comme « l'étude de l'organisation des mots dans la phrase ». Une telle définition est problématique, puisqu'elle suppose que l'on peut définir les notions de mot et de phrase avant de définir ce qu'est la syntaxe. Dans cet ouvrage, la notion de mot ne sera définie qu'au début de la quatrième partie et l'unité maximale de la syntaxe ne sera discutée que dans la sixième et dernière partie de l'ouvrage. (Kahane et Gerdes 2022 : 5)

Cela pose aussi des problèmes d'un point de vue pratique, notamment autour des choix à poser, avec leurs implications, quand il s'agit de constituer un corpus annoté en relation syntaxique (→ 3.2.).

Comme me le disait Alain Berrendonner lors d'une discussion privée, lorsque je lui demandais de m'expliquer la différence entre l'école de macro-syntaxe d'Aix (autour des travaux de l'équipe du GARS et de Claire Blanche-Benveniste, Blanche-Benveniste et al. 1990) et celle de Fribourg (autour de ses propres travaux, qui ont donné lieu à la *Grammaire de la période*, Groupe de Fribourg 2013), il existe deux façons de prendre la route de Neuchâtel à Fribourg : en partant de Neuchâtel ou en partant de Fribourg. C'est ainsi qu'il m'expliqua la distinction qu'il voyait entre ces deux écoles, l'une partant des unités minimales pour en délimiter les limites maximales, l'autre, la sienne, partant des limites maximales. On l'aura compris, à partir de ce qui a été dit jusqu'ici et du titre même de cette section, j'ai adopté dans mes travaux un positionnement « par le bas », partant du verbe et des phénomènes de réaction verbale pour établir les limites maximales. Au départ, je m'étais même limitée à définir le verbe comme noyau de la proposition, appelée *construction verbale CV* (sans lien ici avec les grammaires de constructions), m'intéressant uniquement à l'étude des relations entre ces CV, aux types de liens qui sont tissés à partir du verbe, sans chercher à définir l'unité maximale. C'est une position qui opte pour une définition minimale sans préjuger de l'usage maximal en discours. Ce choix de définition minimale est également celui qui a été retenu, aussi pour des raisons pratiques, dans le projet d'annotation syntaxique du corpus SRCMF, où chaque verbe constitue le noyau d'une phrase, et où le choix avait été de ne pas faire de coordination au niveau phrastique : « On considère qu'il n'y a jamais [Coo](#) [=coordination] de Snt [=phrase]: chaque Snt

contient un seul V [=verbe] principal » (Bischoff et al. 2012 [10], sv *phrase* (Snt), <http://srcmf.org/fiches/Snt.html>) (voir aussi notamment Mazziotta 2012). C'est également le choix que nous faisons actuellement pour l'annotation syntaxique du corpus des Vocaux, le parti pris d'un découpage maximal. Ainsi, la constitution d'un corpus annoté en syntaxe nécessite de faire des choix et d'établir, autant que possible, des critères permettant la reproductibilité et la cohérence du corpus (idéal inatteignable dans une annotation manuelle, comme chacun sait), et limiter l'unité maximale à un verbe principal peut être un moyen de tendre vers ce but. Même ainsi la question n'est pas simple, en particulier sur l'oral (cf. Benzitoun et al. 2010), mais le problème est transversal aux différents états de langue (cf. Lavrentiev 2010). C'est justement ces problèmes de limites de la rection qui ont conduit au développement des écoles de macro-syntaxe citées précédemment. L'analyse macro-syntaxique vise à traiter les éléments liés par une connexion qui ne se laisse pas définir par les critères traditionnels de la grammaire. Dans cette approche, la micro-syntaxe est restreinte aux phénomènes de rection. Ainsi, le projet de constitution d'un corpus de français parlé annoté en syntaxe, *Rhapsodie*, distingue unité de rection⁵⁰ et unité illocutoire⁵¹, qui correspond à un « acte de langage », définissant par-là deux unités maximales, les frontières de l'une comme l'autre ne se recoupant pas nécessairement. Si la notion d'unité illocutoire ne m'a pas été utile dans mes travaux sur l'ancien français, de même que je n'ai pas appliqué les concepts de macro-syntaxe, elle peut reprendre toute sa pertinence pour la délimitation des phrases dans le corpus des Vocaux. Une définition par le bas donc de la « phrase », éventuellement restreinte autant que possible autour d'un noyau de rection, qui en définit les bornes, mais tout en conservant la visée illocutoire, voire l'intentionnalité. Mais est-ce toujours de la syntaxe ?

La définition de l'unité maximale au final dépend du point de vue qu'on veut poser sur l'objet, et des objectifs, éventuellement pratiques, qu'on vise. Mais surtout, la définition de l'unité maximale en syntaxe pose la question de la définition de la syntaxe elle-même. Définir la phrase délimite ainsi le domaine et le modèle.

⁵⁰ « Une UR est une unité construite autour d'une tête, qui n'est syntaxiquement dépendante d'aucun élément de rang supérieur dans un texte ou discours donné » (Benzitoun et al. 2010 : 2076).

⁵¹ « On appelle *unité illocutoire* une portion de discours comportant un unique acte illocutoire, soit une assertion, soit une interrogation, soit une injonction » (*ibid.* 2078).

Les limites de la syntaxe

Se poser la question de l'intégration des propositions, et ainsi, de la limite entre une ou deux phrases, suppose pourtant d'avoir une limite. J'avais alors une définition floue de la phrase, comme le lieu où opèrent les règles de la syntaxe : il y a syntaxe, et donc phrase, tant qu'il y a lien syntaxique. Bien que circulaire, cette définition est assez répandue dans le domaine, comme l'ont rappelé Kahane et Gerdes (2022 : 5, cité *supra*). Partant de là, la question de la définition de la phrase est liée à la question de la définition de la syntaxe : la « phrase » est le domaine, la limite, d'application des règles de la syntaxe, dans ma perception des débuts. D'autres types de relation entrent en jeu au-delà. La question de la parataxe constitue la recherche de ce moment de bascule si difficile à saisir qui fait passer de la syntaxe au texte et aux règles de structuration textuelle (cf. Benzitoun et Sabio 2010 pour la même question sur le français parlé). Ce problème a été très souvent abordé dans la littérature, dans l'étude de la corrélation notamment, autour de la problématique de pouvoir établir la limite entre une corrélation construite syntaxiquement ou se situant au niveau du discours (Deulofeu 2001, Roig 2015). La question de l'unité maximale amène directement à la question de la définition de la syntaxe, question actuelle, comme on peut le voir dans l'appel à communication récent pour le colloque « Français parlé : données, représentations, questionnements théoriques » qui s'est déroulé à Turin les 16-17 juin 2022 : « Les notions de construction verbale (syntaxique), d'énoncé (pragmatique) et de structure thème-rhème (informationnelle), souvent utilisées dans les descriptions de l'oral, sont partiellement incommensurables par rapport aux notions de phrase et de discours et remettent en question le bien-fondé de la séparation entre domaine syntaxique et domaine sémantico-pragmatique. » pour conclure l'appel sur « Où en est-on des recherches sur la macrosyntaxe et, plus largement, sur les frontières de la syntaxe ? » (appel à communication⁵²).

Pendant longtemps, donc, j'ai pensé que les limites de la phrase étaient les limites de la syntaxe, la phrase étant le domaine d'application des règles de la syntaxe. Martinet (1985) débute ainsi sa *Syntaxe Générale* par la définition de la syntaxe (bien que « définir l'objet de la syntaxe n'est pas chose facile » *ibid.* 13) comme « l'examen de la façon dont les unités linguistiques douées de sens se combinent, dans la chaîne parlée, pour former des énoncés »,

⁵² Colloque organisé par Ruggero Druetta (Université de Turin), Maria Margherita Mattioda (Université de Turin), Patricia Kottelat (Université de Turin), Carole Conti (Université de Turin), Claire Martinot (Sorbonne Université), Sonia Gerolimich (Université de Udine), Isabelle Stabarin (Université de Trieste), Giulia D'Andrea (Université du Salento). <https://www.dipartimentolingue.unito.it/do/avvisi.pl/Show?id=2sil>. Appel complet consultable ici : <https://www.fabula.org/actualites/106063/francais-parle--donnees-representations-questionnements-theoriques.html>.

« point sur lequel peuvent tomber d'accord les linguistes contemporains » (Martinet 1985 : 13). J'étais même allée dans ma thèse jusqu'à déclarer : « S'il n'est pas possible d'établir les limites de l'unité maximale en syntaxe, est-il toujours possible de parler de syntaxe ? » (Glikman, thèse, 2009 : 52), tout en ouvrant déjà une brèche : « Tout dépendra donc de ce qu'on voudra que soit, ou non, pris en charge par la syntaxe. » (Glikman, thèse, 2009 : 78). En effet, se poser la question de l'unité maximale, c'est se poser la question de la délimitation du champ de la syntaxe. Comme nous venons de le voir dans l'appel du colloque de Turin, la délimitation des unités tend à fragiliser la séparation entre les domaines d'analyse. Peut-être que le problème vient justement du lien que je faisais moi-même au départ, entre unité maximale et règles de la syntaxe : peut-être que les règles de la syntaxe peuvent dépasser les limites de son unité maximale ? ou que l'unité maximale peut englober d'autres types de liens ? On rejoindrait ainsi l'approche de Benzitoun et al. (2010), qui ont montré que les frontières entre unités de réaction et unités illocutoires ne se recoupent pas toujours. Mais pour avoir à la fois suivi les discussions autour de Rhapsodie et SRCMF au départ, et avoir, de nouveau, les mêmes types de discussion avec Christophe Benzitoun et Nicolas Mazziotta sur le projet Les Vocaux (qui a dit que l'histoire ne se répétait pas ?), il ressort que face aux données, il reste des cas où nous ne savons pas. Au point que Nicolas Mazziotta a un jour proposé, lors d'une discussion, de développer un modèle de « syntaxe floue ».

On en revient à la question principale : qu'est-ce qui relève de la syntaxe ? Si l'on reprend la définition de Ducrot et Todorov citée *supra* (→ 2.2.1.), la syntaxe traite « d'ordre des mots » et des phénomènes de « réaction (accord ou régime) », ainsi que des « fonctions ». La syntaxe est l'étude des « liens » et non des « mots », si l'on repart de Tesnière (1976 [1959]), mais l'entrée dans son étude passe par les mots et leur distribution, leurs caractéristiques morphosyntaxiques, etc. (ce qui nous ramène à notre questionnement de départ : comment accéder au lien syntaxique ?). Par certains aspects, la syntaxe se trouve ainsi réduite à l'analyse de l'ordre des mots. Celui-ci, qu'il soit plus ou moins grammaticalisé comme en français (Marchello-Nizia 1995), est en effet une composante traditionnelle de la syntaxe, comme en témoigne notamment la place qui lui est accordée dans la *GGHF* (Marchello-Nizia et al. (éds) 2020). La syntaxe doit-elle se réduire à l'étude de l'ordre des mots ? En un sens, au moins il s'agit là d'un observable concret. Kim Gerdes et Sylvain Kahane (Gerdes et Kahane 2006 et Kahane et Gerdes 2022) distinguent ainsi l'ordre des mots de la topologie, véritable composante de la syntaxe :

On renvoie traditionnellement à la branche de la syntaxe qui s'intéresse aux questions d'ordre en parlant d'ORDRE DES MOTS. Nous préférons parler de TOPOLOGIE et cela pour au moins trois raisons.

Premièrement, nous considérons que les unités minimales de la syntaxe sont les syntaxèmes et non les mots et donc que les questions d'ordre commencent avec l'ordre des syntaxèmes à l'intérieur des mots.

Deuxièmement, nous ne nous intéressons pas seulement à l'ordre relatif des syntaxèmes (ou des mots ou des unités syntaxiques en général), mais plus généralement à la POSITION qu'occupent ces éléments dans une structure qui est plus riche qu'un simple ordre linéaire. C'est ce que désigne la topologie, qui est l'étude des lieux ou places (du grec *topos* 'lieu').

Troisièmement, le terme *topologie* (du grec *logos* 'étude'), à la différence de *ordre des mots*, renvoie clairement à un domaine d'étude de la langue ; il permet de plus de dériver l'adjectif *topologique* et de parler par exemple de *structure topologique*, de *constituant topologique* ou de *niveau topologique*.

Le terme *topologie* provient de la tradition grammaticale allemande, où le terme est d'usage depuis le milieu du 20^e siècle (Kahane et Gerdes 2022 : 464)

Dans un autre ordre d'idées, on peut aussi se demander si les phénomènes de rection relèvent bien toujours de la syntaxe. En effet, est-ce que cela ne relève pas plutôt du lexique, chaque terme pouvant être considéré comme porteur de ses propres propriétés combinatoires, on pensera notamment aux approches de type lexique - grammaire (Gross 1984)⁵³. Quelle place accorder à la sémantique et à la pragmatique, voire à la « programmation », dans la délimitation de l'unité maximale, et comment l'articuler à la syntaxe ?

Au CMLF 2022, lors de l'exposé de Nicolas Gregov sur l'apposition et son traitement dans les grammaires (Gregov 2022), Franck Neveu, le remerciant de son travail, a évoqué alors le fait que traiter du phénomène de l'apposition revient à interroger les limites de la syntaxe. Forcément, dans un premier temps, j'ai soupiré dans mon coin, puisque, selon moi, c'est l'étude de la parataxe, de la limite entre une ou deux phrases, qui interroge véritablement les limites de la syntaxe. De fait, le lien entre parataxe et apposition existe, ne serait-ce que par la question de l'absence de lien explicite. Mais plus encore, après réflexion et diverses discussions, en particulier avec Nicolas Mazziotta, il apparaît qu'en effet, l'un comme l'autre interrogent le

⁵³ Le lexique-grammaire est consultable à l'adresse suivante (dernière consultation le 29.11.22) : <https://ladl.univ-mlv.fr/DonneesLinguistiques/Lexiques-Grammaires/Presentation.html>.

modèle, et, surtout, l'un comme l'autre posent le problème des constituants de même nature au même *niveau*, configuration qui pose problème à la plupart des modèles. Peut-être que la syntaxe ne se définit qu'à travers sa modélisation, et qu'il n'existe pas de syntaxe hors d'un cadre théorique (comme toute composante de la linguistique ?), le point de vue créant l'objet, comme disait Saussure.

Aujourd'hui, dans mes travaux, la question de la délimitation de la syntaxe s'est déplacée à l'appréhension du contexte en premier lieu : tout est contexte. En effet, une des pistes que je souhaite développer est le fait que des éléments de discours peuvent prendre une fonction de par leur position davantage que par leur forme, c'est-à-dire par l'évolution de leur emploi en contexte davantage que par l'évolution de leur forme. On pensera par exemple aux marqueurs de discours ou aux incidentes (cf. Laferrière 2018 sur la diversité lexicale des incises), ou aux termes d'adresse étudiés par Dominique Lagorgette et définis par un faisceau de critères, y compris contextuels :

Nous opposerons les GN désignatifs (sujet ou objet) dans le récit aux termes d'adresse, syntagmes nominaux détachés dans le discours¹ [en note : Pour une définition précise de ces deux notions, on se reportera à D. Lagorgette (1998 : ch.1). *Nous abrègerons dorénavant désignatifs par GNDés et termes d'adresse par T.A.*]. Nous proposons de définir les désignatifs comme les syntagmes nominaux de dénomination et de désignation des personnages. Les termes d'adresse sont des groupes nominaux renvoyant exclusivement à la deuxième personne (l'allocutaire) dans le discours direct. (Lagorgette 2002 : 1)

Dans la mesure où de nombreux exemples montrent (toutes époques confondues) que n'importe quel nom peut être employé comme TA ou insulte, les approches strictement syntaxiques ou lexicales ne peuvent suffire à caractériser ce type de SN, même si, bien entendu, les critères syntaxiques sont essentiels pour en délimiter les frontières – et il s'agit d'autant plus de critères syntaxiques que la ponctuation ne peut être fiable, ou même pertinente, selon les époques envisagées.

Outre son caractère détaché, qu'est-ce qui caractérise donc, finalement, un TA? D'une part, (7) sa grande liberté de positionnement dans les énoncés 19; d'autre part, (8) la présence de ces SN dans le discours direct (et nous quittons maintenant le domaine de la grammaire de phrase) ; enfin, (9) la capacité à modifier la valeur performative d'un énoncé. [...]

Ce seul exemple nous permet de dégager quelques critères de plus, cette fois d'ordre pragmatique : (10) intonation, (11) pertinence par rapport au contexte, (12) savoirs communs partagés ajustés entre L1 et L2, (13) intention de L1 et de L2 (14) dans la co-construction de l'interaction sont clairement à prendre en compte dans les opérations de calcul du sens. Bien souvent, l'identification des SN d'adresse passe en effet par des relectures a posteriori de l'entier d'un énoncé ou d'une séquence et/ou par des explicitations de L1 (ex. 1 et 2). Ceci n'est en rien étonnant dans la mesure où ces SN détachés sont par définition autonomes et averbaux : ils entrent dans la construction des chaînes de référence à un niveau moindre que les syntagmes intégrés (désignatifs) qui

ont des fonctions dans les phrases (sujet, objet), et nécessitent de la sorte, pour une interprétation correcte, la mise en relation de plus d'éléments (marques de 2e personne, analyse globale du contenu sémantique de l'énoncé). En dernier ressort, il nous semble que c'est donc la convergence des divers critères, tout juste évoqués, qui permet le mieux de caractériser le phénomène de l'adresse nominale : plus grand est le nombre de critères validés, plus on sera sûr d'avoir à faire à un SN d'adresse. (Lagorgette 2006 : 43-44)

Cette importance donnée au contexte rejoint, d'une certaine manière, la notion de paradigme, en lien avec les constructions, telle que développée par Lene Schøsler et ses collègues :

Nous définissons les constructions (et plus spécifiquement les *schematic constructions*) comme des structures grammaticalisées organisées en paradigmes d'oppositions, à l'instar des oppositions morphologiques (Kragh & Schøsler 2015). Ces paradigmes sont créés, réorganisés ou abandonnés comme les paradigmes morphologiques. En d'autres termes, les structures complexes peuvent se «grammaticaliser» ou se «dégrammaticaliser», c.-à-d. perdre leur statut grammatical, comme c'est le cas de la quasi-disparition du progressif français formé à l'aide d'un verbe auxiliaire et d'un participe présent, du type *il va chantant* (Schøsler 2004, 2005 ; Kragh & Schøsler 2015). Ces structures peuvent en outre se «regrammaticaliser» (à savoir s'intégrer dans un nouveau paradigme) : l'illustration étant la «regrammation» [sic] de *voici* et *voilà* qui s'intègrent dans le paradigme des marqueurs discursifs (Kragh 2021 ce volume). Le processus de grammaticalisation, de *dégrammaticalisation* ou de *regrammaticalisation* se poursuit selon les principes de l'opposition participative ou de *markedness* comme cela a été proposé par A. Timberlake (1977) et H. Andersen (2001).

Nos définitions comportent les traits caractéristiques suivants : pour qu'une construction fasse partie de la grammaire de la langue, elle doit avoir une *expression* (une forme)⁴ et un *contenu*, et ce contenu n'équivaut pas à la somme de ses éléments constitutifs⁵. L'organisation paradigmatique d'une construction prévoit que les membres de ce paradigme s'opposent entre eux grâce à des différences systématiques et prévisibles d'*expression* et/ou de *contenu*. Parmi ces membres, l'un est non marqué, c.-à-d. qu'il possède des traits plus généraux par rapport aux membres marqués, qui possèdent des traits plus spécifiques qui permettent de les distinguer (cf. § 4.3). (Kragh et Schøsler 2021 : 85)

J'ai également élargi la notion de contexte à la prise en compte du jugement des locuteurs, mes travaux récents ayant mis au jour l'importance de la notion de « préférence » dans la langue, davantage que celle de grammaticalité (→ 4.2.). Ce qui nous conduit à une approche cette fois-ci inverse de la syntaxe. En effet, la syntaxe est souvent vue comme la zone de la langue où s'opèrent les contraintes (combinatoire, etc.). Cette notion de contraintes, et plus encore dans une approche de la langue qui s'occupe de variantes et de variation, est centrale, et nous conduit à interroger les limites de la syntaxe non plus par « en bas », mais par « en haut », autour de la question : qu'est-ce qui est dans la grammaire, et qu'est-ce qui est dans le « style », dans le « choix », dans la « préférence », et qui relève des paramètres de variation et de l'intentionnalité du locuteur (→ 3).

Bilan intermédiaire

Dans cette section, j'ai présenté ce qui a été une préoccupation au centre d'une grande partie de mes travaux : établir le statut syntaxique des constructions, première étape nécessaire pour établir leur statut de variante. Cette étape est le préambule à l'analyse de leurs caractéristiques en tant que variantes : spécificités ou absence de spécificités, critères de distribution et / ou paramètres de variation, que nous aborderons dans le chapitre suivant (→ 3.). Dans ce que j'ai présenté jusqu'ici, il s'agissait surtout d'établir le lien syntaxique entre deux constructions verbales, lien de subordination ou non (« subordination » pris ici au sens traditionnel de subordination phrastique), notamment en l'absence de marque explicite de subordination. Du point de vue de l'évolution de la langue, ce qui est ressorti de mes travaux (en tout cas dans le cas de l'étude de la parataxe), c'est que l'existence d'un lien est ce qui ne change pas. J'avais d'ailleurs cherché à montrer dans ma thèse non pas ce qui changeait, mais ce qui ne changeait pas (ce qui peut sembler plutôt paradoxal pour une thèse en diachronie). J'ai continué à mettre en avant dans mes travaux ultérieurs, évoqués précédemment, que ce n'était pas le lien syntaxique qui avait changé au cours de l'histoire du français, mais éventuellement la manière dont il est exprimé et, dans le cas de la parataxe, la répartition des variantes (→ 4.).

La rédaction du mémoire de synthèse est aussi ce moment où l'on doit se demander ce que nos propres travaux ont apporté sur la question, c'est-à-dire ici, pour conclure cette section, sur l'établissement du lien syntaxique afin d'établir les variantes. Outre la description des phénomènes eux-mêmes, mes travaux ont contribué à mettre en avant l'importance de considérer le marquage du lien syntaxique comme pouvant provenir d'un faisceau d'indices, ainsi que l'importance de la prise en compte du contexte. En cela, mes travaux s'inscrivent dans une large tradition de l'analyse syntaxique et rejoignent plusieurs études sur la question comme j'ai essayé de le montrer jusqu'ici. En particulier, et c'est peut-être ce qui constitue davantage leur originalité, dans plusieurs contributions j'ai établi des parallèles avec les analyses sur le français moderne (pour interroger l'évolution, en bonne diachronie), et surtout avec les études sur le français moderne « oral » (→ 3.1 & 4.1). Ainsi, ce qui est ressorti de mes premiers travaux, c'est surtout la répartition de ces variantes et la réalisation de la variation. Une ouverture à la fin de ma thèse a été qu'au lieu d'avoir « disparue », la parataxe connaît une distribution de type « oral/écrit » relativement stable au fil du temps, ce qui expliquerait sa « disparition » dans le développement de la prose à la fin du moyen-âge, tout en confirmant que le développement de cette prose s'est fait avec de plus en plus de distance avec « l'oral » ou

l'oralité au sens conceptuel (→ 4.1.). Après ma thèse, j'ai à la fois poursuivi mes recherches sur la subordination et ses marqueurs, sur les constructions verbales atypiques (comme les incidentes), mais la question de la répartition des variantes m'a amené à m'intéresser à d'autres zones de variables dans la langue, pour en étudier la variation (→ 2.3.). C'est ainsi que je me suis de plus en plus intéressée au *statut* des variantes, ce qui m'a conduite à travailler sur la perception de la variation par le locuteur (→ 4.2.). Je me suis également intéressée aux paramètres de variation qui émergent de comparaisons plus globales, comme l'étude des spécificités de l'oral représenté par rapport à la narration et ses marqueurs, et les spécificités associées à un genre textuel, ainsi qu'aux méthodes pour analyser la variation (→ 3.).

Aujourd'hui, certaines questions sont toujours présentes. Par exemple, dans l'analyse que nous menons actuellement sur le corpus des Vocaux (voir Glikman et Fauth 2022 [34]), l'intégration de certaines constructions phrastiques pose toujours le même type de questions, et, comme le dit Christophe Benzitoun, certains cas restent indécidables. Nous en revenons à la question du *choix* lors de l'annotation, pour laquelle Nicolas Mazziotta a bien montré que l'annotation constitue déjà l'analyse (Mazziotta 2016), et pour lequel il faudrait pouvoir opter pour une annotation qui « laisse le choix » ou du moins montre l'ambiguïté, ce que les formats modernes type xml devraient pouvoir rendre possible. De la même manière, quand j'avais co-encadré une étudiante en mémoire de master de Traitement Automatique des Langues sur l'annotation des verbes en personne, mode et temps du texte de la *Queste del saint Graal* (sur l'édition numérique de Christiane Marchello-Nizia et Alexei Lavrentiev⁵⁴) (Bringer 2019), j'avais attiré son attention sur l'importance d'afficher le caractère ambigu de certaines occurrences, dont la forme pouvait correspondre à deux ou plusieurs analyses (ex. présent / passé simple *dit*). Bien entendu, cela n'empêche pas l'expert de réussir à déterminer l'analyse la plus probable en contexte. Conserver l'ambiguïté permet cependant de la laisser apparente et de ne pas forcer l'analyse.

Après ces considérations d'ordre plus général liées à la problématique de la détermination du lien syntaxique et des enjeux que cela soulève, je vais à présent revenir à l'établissement des variantes en revenant sur les autres zones de variables que j'ai étudiées.

⁵⁴ *Queste del saint Graal*, édité par Christiane Marchello-Nizia et Alexei Lavrentiev, Lyon, ENS de Lyon, 2019. Publié en ligne par la Base de français médiéval, http://catalog.bfm-corpus.org/qgraal_cm Dernière révision le 2018-11-30.

2.3. Autour d'autres zones de variables : établir les variantes en concurrence

*Whatever you want,
the choice is yours
So choose*

Gossip, *Heavy Cross*, 2009.

Comme je l'ai évoqué précédemment, les objets que j'ai traités ont naturellement évolué. Une nouvelle recherche a parfois simplement découlé de la précédente, à la manière d'un fil qu'on tire. C'est le cas de mes travaux sur la subordination en général⁵⁵, sur les verbes parenthétiques et les constructions incidentes⁵⁶, sur *ce*⁵⁷, et, d'une certaine manière, sur les locutions conjonctives⁵⁸ et sur les spécificités de l'oral représenté et des genres textuels. D'autres sont nées simplement d'un intérêt porté à un phénomène observé dans les productions, comme l'alternance *avoir été / être allé*, l'évolution de l'accord de *tout* ou de *espèce de*, les emplois modernes de *moult*⁵⁹. Ce dernier article est venu d'une réflexion faite par une de mes étudiantes en cours de diachronie à Strasbourg, lorsque je présentais le traditionnel remplacement de *moult* par *beaucoup* et *très* : « mais Madame, *moult* n'a pas disparu ! ». Enfin, d'autres ont résulté de nouvelles opportunités de collaboration. C'est le cas des travaux sur les chaînes de références⁶⁰, peut-être plus éloignés de mes objets d'étude habituels, auxquels j'ai collaboré grâce à ma participation aux projets MC4 puis DEMOCRAT évoqués en introduction. De même, mes travaux sur la variation en français contemporain sont le résultat d'une collaboration avec Mathieu Avanzi en particulier, dans les projets Français de Nos Régions et Donnez Votre Français à la Science (→ 3.2.).

Dans tous les cas, il s'agit toujours d'observer la réalisation d'un phénomène en fonction de différents paramètres et de mettre en avant la question du *choix*. Dans chacun de mes travaux, je m'attache à observer une zone de *variable* de la langue : décrire des zones d'alternance

⁵⁵ Glikman 2009 [5], 2010a [6],b [7], 2013 [11], 2016 [18], 2020a [29],b [30], Glikman et Verjans 2013 [13], Fagard, Pietrandrea et Glikman 2016 [17], Combettes et Glikman 2020 [28].

⁵⁶ Glikman 2012 [8], Glikman et Avanzi (dir) 2009 [2012]Mul [ii], Schneider et Glikman 2015, Schneider [15], Glikman et Avanzi (dir) 2015 [iii], Glikman et Schneider 2018 [23].

⁵⁷ Mazziotta, Bischof, Glikman et Rainsford 2012 [9].

⁵⁸ Glikman 2019 [26], Glikman et Bouard 2021 [31].

⁵⁹ Glikman et Patard 2022 [35], Bouard et Glikman 2018 [21], Verjans et Glikman 2022 [36], Glikman et Verjans 2021 [32].

⁶⁰ Glikman, Guillot-Barbance, Obry 2014 [14], Obry, Glikman, Guillot-Barbance, Pincemin 2017 [20], Schnedecker, Glikman, Landragin (dir.) 2017 [iv], Oberlé et al. 2018 [25], Capin, Glikman, Schnedecker, Todirascu 2021 [33].

possible et les conditions de réalisations de cette alternance, distribution libre ou contrainte, type de variation en jeu... pour lesquels il faut commencer par établir les observables de la variation.

Ces zones de variable constituent en effet un point d'entrée privilégié à l'étude du changement linguistique, comme nous l'avons évoqué en introduction. En cela, même dans l'étude de phénomènes relevant davantage du niveau lexical ou textuel, mon approche reste une approche de type syntaxique, à travers l'étude de la distribution des phénomènes et de leur *grammaticalité* et du passage d'un choix libre à un choix contraint. Le locuteur est-il réellement libre de *choisir*, ou doit-il suivre des règles, et si oui, lesquelles, et comment celles-ci évoluent-elles au cours du temps ? J'essaie ainsi de saisir ce moment de bascule et comment ces choix et règles se déplacent au cours de l'histoire du français. Dans l'analyse du changement linguistique, l'idée que je cherche à mettre en avant et qui ressort de mes travaux jusqu'ici, et que je souhaite poursuivre, c'est montrer que le changement se situe quand la variante change de statut (→ 4.).

Pour l'étude de ces zones de variable, la question d'établir les variantes se pose également. Comme je l'ai évoqué (→ 2.1.), quel phénomène peut être considéré comme comparable dans l'étude de locutions conjonctives, de *moult*, ou encore des chaînes de référence ? Comment établir un *contexte comparable* ? Selon les cas, je dois avouer que la définition de la variante n'a pas toujours été faite de manière explicite, la description du phénomène ayant été menée dans une comparaison implicite, par exemple dans l'article sur *moult*. Dans d'autres cas, plusieurs éléments pouvaient être considérés comme comparables, mais tous n'ont pas été étudiés, comme dans l'article sur *malgré que* et *à cause que*. D'autres études ont été menées explicitement sur des distributions de variantes, comme l'article sur *avoir été* employé pour *aller* au passé. Enfin, dans certains travaux, c'est le texte lui-même qui a été pris comme zone de variable, la réalisation des variantes étant considérée en comparaison sur une échelle plus large, comme la réalisation des chaînes de référence selon les textes (les variantes étant alors les différentes réalisations possibles de chaque maillon), ou les spécificités de l'oral représenté par rapport à la narration (les variantes étudiées ici étant les réalisations des différentes fonctions syntaxiques). Je n'aborderai pas ici tous les objets étudiés, mais seulement des exemples pour illustrer le rapport à la définition de la variante.

2.3.1. Une zone de variable à variantes implicites

Dans notre contribution sur les emplois modernes de *moult* (Glikman & Verjans 2021 [32]), nous nous sommes intéressés à l'évolution des réalisations de *moult* en corpus ainsi qu'à l'évolution des discours métalinguistiques le concernant (dictionnaires, grammaires, mais aussi commentaires linguistiques, notamment sur Internet, de non experts, cf. Siouffi et al 2012). De fait, dans notre article, l'élément de comparaison dans la distribution des usages (et du discours métalinguistique) est implicite. L'observation de la réalisation ou non réalisation de *moult* en corpus ne signifie pas que *moult* alterne dans ces contextes avec une réalisation « vide », mais bien évidemment avec un autre adverbe ou déterminant, selon les cas. Pour notre article, nous ne nous intéressions pas au remplacement de *moult* par *beaucoup* et *très*, largement étudié par ailleurs, ce qui explique que ces formes n'ont pas été prises en compte. Nous nous intéressions uniquement aux réalisations à partir du 16^e s. de *moult*, en particulier pour décrire le fonctionnement de ces emplois, et comprendre leur origine et leur évolution. L'observation des contextes d'apparition a permis de déterminer les fonctions de *moult* dans ces emplois modernes, ce qui nous a permis de mettre au jour le fait que « les emplois modernes de *moult* ne sont pas ceux qui étaient en usage au moment de son déclin au XVI^e s. » (Glikman & Verjans 2021 [32] : 388). Nous avons ainsi cherché dans quels types de textes ces emplois se développaient, et quel était le niveau de contrainte qui pesait sur ces emplois. L'un des corpus utilisés, le corpus de Leipzig⁶¹ permettait cependant de noter les mots avec un contexte similaire ainsi que les voisins de gauche et droite significatifs, que nous avons évoqués mais peut-être pas assez exploités. Les mots avec un contexte similaire ont mis en avant un rapprochement avec *maint*, que nous avons brièvement évoqué mais dont l'étude approfondie resterait encore à faire. L'évolution différente de *maint* par rapport à *moult* pourrait être rapprochée du fait que *maint* ne connaît pas d'emploi adverbial⁶², ses emplois n'ont pas été mis à mal par la tendance à la distinction catégorielle. De même, dans ce corpus, le voisin de gauche le plus significatif de *moult* est *Après*⁶³ et le voisin de droite *péripéties*. Cela est probablement lié au style journalistique fortement présent dans le corpus, donnant ainsi à voir des contextes « préférés », qui ont pu être les premiers contextes ayant favorisé le développement des emplois modernes

⁶¹ French mixed corpus based on material from 2012.

⁶² Cf. emplois recensés dans le DMF et le FEW, consultés le 24.10.22.

⁶³ Avec majuscule, la plateforme distinguant la casse. Il s'agit donc bien, en outre, d'un groupe placé à l'initial de la phrase graphique.

de *moult*, qui se sont ensuite étendus à d'autres contextes. Nous n'avons cependant pas cherché à vérifier les contextes comparables pour analyser les concurrents possibles. Cela ne nous avait pas semblé strictement nécessaire dans cette étude, dont la visée première était d'établir les types de contextes syntaxiques dans lesquels *moult* pouvait apparaître. Nous avons en effet montré que les emplois modernes de *moult* sont principalement des emplois comme déterminants et ayant une construction différente des emplois de *moult* au moment de sa disparition, montrant en cela qu'il s'agit bien d'une (re)création moderne :

Malgré sa disparition et son remplacement par *beaucoup* et *très* à la fin du Moyen Âge, *moult* connaît toujours des emplois modernes. Plutôt qu'une survivance, il semble plus probable que ces emplois modernes soient le résultat d'un emprunt à l'ancienne langue : *moult* ne se serait donc pas maintenu dans l'usage commun, mais aurait été réintégré à la langue et, en tant qu'emprunt, il est aujourd'hui complètement entré dans l'usage. [...] La prononciation et les emplois modernes ne sont également pas ceux qui existaient au moment de sa disparition, *moult* moderne étant majoritairement employé comme déterminant construit directement, une construction qui n'existait plus à la fin du Moyen Âge. Les rares emplois comme adverbe sont clairement marqués comme archaïsants ou ironiques, à la différence de l'emploi comme déterminant, où l'on peut penser que c'est la quantité forte qui est davantage marquée, comme renouvellement expressif de l'intensivité.

La formation du *moult* moderne aboutit ainsi à une sorte de forme mixte, à la fois réservée à l'emploi comme déterminant, mais qui hérite de la forme adverbiale son caractère invariable, créant ainsi une forme hybride entre le fonctionnement de *maint* et celui de *beaucoup de*. On a bien en français moderne une nouvelle création de forme, qui ne correspond ni à ses emplois au XVI^e s. (en *moult de*), ni à ses emplois médiévaux comme déterminant (où l'adjectif était accordé), uniquement déterminant, sans *de*, et sans accord. [...] Il reste à voir s'il est toujours stylistiquement marqué, mais les contextes des différentes occurrences tendent à montrer qu'il s'est émancipé de ses constructions préférées, et qu'il est de moins en moins restreint à un type de contenu presse. Cependant, c'est une forme encore rare à l'oral. On aurait alors affaire à un changement qui arrive par l'écrit. (Glikman & Verjans 2021 [32] : 401-402)

Une rapide recherche menée sur le même corpus sur la forme *Après*⁶⁴ donne cependant comme voisins de droite significatifs (nous conservons uniquement les formes potentiellement comparables à *moult*) *quelques*, *plusieurs*, des déterminants numéraux, et enfin *maintes* qui arrive finalement assez loin derrière *moult*, lui-même situé après *plusieurs*. En outre, la forme *beaucoup* n'apparaît pas dans cette liste. La question des constructions en concurrence mériterait ainsi peut-être d'être poussée plus avant.

⁶⁴ Requête menée le 24.10.22.

De la même manière, je conseille à mes étudiants de se concentrer sur les contextes de réalisation de leur objet pour en comprendre le fonctionnement et en mesurer le degré de diffusion. C'est le cas par exemple de Jonathan Fontaine qui prépare un mémoire de master sur les emplois de *full* en français⁶⁵. L'observation des emplois de *full* en français montre que ce mot peut avoir différentes portées possibles, et donc des emplois de types différents (tous les exemples proviennent du corpus SMS4Science, analysé par Jonathan Fontaine) :

- (45) En plus, on n'a full pas les mêmes amis.
- (46) Ça m'a full fait du bien de te voir aujourd'hui.
- (47) Je suis un peu full en ce moment mais pense très fort à toi.
- (48) Alors le bus est full mais si il y a des désistements, C'est pour nous.
- (49) Moi j'ai encore full trucs à faire!

Chercher des contextes comparables *a priori* dans cette étude n'aurait pas de sens, puisque c'est l'exploration des occurrences réalisées qui ont permis de mettre au jour les contextes d'emploi. De fait, *full* commute dans ces emplois avec d'autres adjectifs, déterminants et adverbes potentiels. Dans le cadre d'un mémoire de master, la première étape reste selon moi la description des contextes d'occurrences et la mesure de la diffusion à travers la diversité des sources présentant l'objet. Dans un travail plus poussé, on pourrait envisager une comparaison explicite des variantes en contexte comparable. Mais là encore, les variantes en concurrence, de type lexical, peuvent être nombreuses (cf. « le bus est *plein/rempli/bondé/complet...* »), l'efficacité d'une telle approche pose donc question. Il me semble véritablement plus efficace de se concentrer sur les réalisations du terme et de mesurer s'il existe des contraintes à ces réalisations (par ex. présence systématique d'une négation, contexte positif ou négatif, personne, temps...) et évidemment les paramètres de variation (que nous verrons dans le chapitre suivant). La mise en évidence des contextes d'emploi de *full* en français permettra ainsi bien sûr de les comparer avec les contextes d'emploi du mot en anglais et d'en mesurer également l'intégration.

2.3.2. Une zone de variable à variantes volontairement restreintes

Dans mes travaux sur *malgré que* et *à cause que* (Glikman 2019 [26]), la question se posait de savoir quels éléments prendre en considération dans l'étude. Un relecteur avait notamment

⁶⁵ Master 2 Sciences du langage à l'Université de Strasbourg, 2021-2022 et 2022-2023.

regretté l'absence de prise en compte de la totalité des modes d'expression de la concession pour mesurer le développement de *malgré que*. De fait, ces locutions connaissent plusieurs variantes possibles :

Dans ce paradigme sémantique de l'expression de la cause et de la concession, elles constituent donc à la fois des choix lexicaux, par rapport au choix d'autres locutions conjonctives, et des choix syntaxiques, par rapport au choix d'expression à l'aide d'une structure non verbale, comme un syntagme prépositionnel, Prép + SN. (Glikman 2019 [26] : 76)

Quels seraient les meilleurs équivalents à observer dans ce cas : *malgré le fait que*, *malgré* + GN, *bien que* ? *à cause de*, *à cause du fait que*... ? L'ensemble des modes d'expression de la concession et de la cause ? Prendre en compte l'ensemble des modes d'expression de la cause et de la conséquence est évidemment intéressant et constitue également une mise en système, mais une telle entreprise dépassait le cadre de cette étude, et aurait demandé une toute autre approche. L'objet de l'article était de comparer les réalisations de ces locutions en diachronie, et de mettre en perspective l'étude sur corpus avec l'évolution du discours grammatical et le jugement des locuteurs observé en enquête (Avanzi et al. 2016 [16], → 3.2.). J'ai pu m'appuyer sur les monographies existantes sur l'expression de la cause (Bertin 1997, pour l'ancien français) et de la concession (Soutet 1990) pour le cadre général. J'ai ainsi principalement exploré les réalisations de *malgré que* en corpus et dans le discours métalinguistique, pour en étudier le développement. L'idée principale était de voir si ces emplois se répandaient à période récente, dans différents types de sources. L'analyse des discours métalinguistiques permettait également à la fois de mesurer l'éventuelle diffusion du phénomène (en tant que « événement méta » comme l'ont proposé Siouffi, Steuckardt et Wionet 2012), y compris comme forme condamnée, et à la fois de montrer l'évolution du discours sur la forme, *à cause que* connaissant une stigmatisation relativement récente, tandis que *malgré que* commence à être de plus en plus toléré. Prendre en compte le discours grammatical comme indication de la diffusion d'une forme me semble nécessaire, mais dans certains cas, l'importance donnée au phénomène en question interroge. C'est en effet ce que j'avais constaté dans mes travaux sur *de façon à ce que*⁶⁶, forme pour laquelle le discours métalinguistique est très important alors que les attestations de la forme restent rares en corpus.

⁶⁶ Communication non publiée, Glikman J. (2022) « Variation et mécanismes d'analogie dans l'évolution des locutions conjonctives formées sur *façon*, *manière* et *sorte* », Colloque international *DIACHRO X Le français en diachronie*, 31 mai au 2 juin 2022, Sorbonne Université.

Dans l'étude sur corpus, j'ai cependant uniquement analysé les occurrences de *malgré que* et à *cause que*. La seule autre réalisation incluse a été *malgré que + en avoir*, que les grammaires donnent comme à l'origine du développement de la locution conjonctive *malgré que* (explication infondée selon mes analyses). De la même manière, lors de l'enquête menée auprès des locuteurs (Avanzi et al. 2016 [16]), les questions étaient volontairement orientées. Pour à *cause que*, les participants devaient simplement déclarer leur fréquence d'utilisation à partir d'un exemple. Mais pour *malgré que*, les variantes proposées n'incluaient pas d'autres variantes possibles, en particulier *bien que*, pour restreindre le choix des participants :

Pour la locution *malgré que*, le contexte donné était « Marie avait prévu d'aller faire un tour de vélo, mais il pleut. Elle est têtue. Cela veut dire », suivi de trois choix possibles (les réponses multiples étant autorisées) : « Malgré la pluie, elle va aller faire un tour de vélo / Malgré qu'il pleut, elle va aller faire un tour de vélo / Malgré qu'il pleuve, elle va aller faire un tour de vélo ». Pour la locution à *cause que*, le contexte donné était « Ce printemps, votre facture d'électricité a été moins élevée que l'année précédente, à cause qu'il n'a pas fait très froid cet hiver », et les participants devaient déclarer la fréquence d'emploi à laquelle ils pensaient utiliser l'expression, sur une échelle allant de 0 (jamais) à 10 (souvent). (Glikman 2019 [26] : 89-90)

J'ai ainsi choisi volontairement de restreindre les variantes en considération dans cette étude. La question du choix des variantes incluses dans une recherche sur corpus et dans un questionnaire linguistique pose de vraies questions. En corpus peut se poser le problème du coût de traitement, si les variantes sont trop nombreuses, trop larges, ou encore, évidemment, si elles risquent de générer trop de bruit (i.e. données qui ressortent de la requête sans cependant correspondre à la cible). Nous avons par exemple rencontré le problème avec Marie Didienne lors de son mémoire de master (Didienne 2019). Marie Didienne travaillait sur la diffusion des néologismes de la commission terminologique de la Délégation Générale à la Langue Française et aux Langues de France *FranceTerme*⁶⁷ et sur leur implantation face aux emprunts qu'ils ont pour objectif de remplacer. Autant une recherche sur *autocaravane / camping-car* peut donner des résultats exploitables, autant une requête sur *chat / dialogue* génère trop de bruit et de contextes non similaires pour être facilement traitée. Dans les enquêtes linguistiques se pose le problème d'élicitation de la forme voulue. En effet, si l'on avait inclus le choix *bien que*, nous prenions le risque que les participants privilégient cette réponse. L'enquête constitue en effet un cadre non pas de production spontanée, mais de déclaration d'usage. En cela, elle fait appel à la conscience linguistique du locuteur, en particulier à sa représentation de la norme. On

⁶⁷ <http://www.culture.fr/franceterme>.

s'attendait donc à avoir des réponses en partie « surveillées ». En donnant au contraire le choix entre *malgré que* + indicatif et *malgré que* + subjonctif, nous comptons jouer sur le poids normatif de l'opposition indicatif / subjonctif dans la subordination pour tromper cette vigilance, en nous appuyant sur les discussions sur les emplois du mode avec *après que*⁶⁸, que tous les locuteurs ayant été scolarisés ont de grandes chances d'avoir entendues à un moment ou un autre. En cela, les réponses à l'enquête ont donné des résultats relativement exploitables, même si la part de rejet de la locution conjonctive reste élevée.

Il me semble donc important de considérer pleinement les variantes que l'on veut tester et de justifier ces choix. Dans le cas de *malgré que*, l'étude mériterait d'être complétée, à la fois dans la recherche sur corpus et dans les enquêtes auprès du public. Dans la recherche sur corpus, on pourrait observer la répartition des différentes constructions de *malgré*. La réalisation comme préposition est peut-être cependant plus éloignée des contextes de *malgré que*. En effet, *malgré moi* n'a pas forcément d'équivalent propositionnel, à la différence de *malgré la pluie* qu'on peut plus facilement imaginer équivalent de *malgré qu'il pleut/pleuve*. La variante privilégiée à prendre en compte pour l'étude de *malgré que* resterait peut-être *malgré le fait que*, dont le développement serait intéressant à étudier. En effet, une rapide recherche menée sur *Frantext* sur la séquence « malgré le fait que/qu' » donne les attestations les plus anciennes en 1928⁶⁹, donc contemporain au recul de *malgré que* dans *Frantext* (cf. Glikman 2019 [26]). À cette recherche pourrait s'ajouter la variante *bien que*, première variante qui s'impose à l'esprit quand on évoque *malgré que*, ainsi que la variante de référence citée dans la plupart des discours métalinguistiques (parfois à côté de *quoique*). Une enquête linguistique qui prendrait en compte ces trois variantes donnerait certainement également des résultats intéressants à étudier. On pourrait ainsi peut-être avoir un aperçu des préférences des locuteurs, mais aussi d'une possible (re)motivation sémantique de la variation, telle que Gide, cité dans le *Grand Robert*, s'en fait le porte-parole :

« J'ai écrit avec Proust et Barrès, et ne rougirai pas d'écrire encore : malgré que, estimant que, si l'expression était fautive hier, elle a cessé de l'être. Elle ne se confond pas avec bien que, qui n'indique qu'une résistance passive ; elle indique une opposition ». (Gide, *Incidences, Lettre à P. Souday*, 13 oct. 1923, cité par *Le Grand Robert de la langue française 2016*, en ligne, cité ici d'après Glikman 2019 [26] : 86)

⁶⁸ Nous avons également interrogé les participants sur *après que* dans nos enquêtes, l'exploitation des résultats reste à faire.

⁶⁹ Requête menée le 26.10.22.

De la même manière, dans notre article sur *car* en contexte injonctif avec Daniéla Capin (Capin et Glikman 2017 [19]), nous avons inclus une brève comparaison avec *mais*, *or* et *donc* en contexte équivalent, ce qui nous avait permis de mettre au jour la plus grande flexibilité d'emploi de *donc*, pouvant expliquer l'extension de ses emplois au détriment de *car* dans ces contextes.

2.3.3. Une zone de variable à variantes explicites

Dans notre étude sur *être* mis pour *aller*, menée avec Adeline Patard (Glikman et Patard 2022 [35]), les variantes en concurrence sont au contraire explicitées, ainsi que les contextes de variation clairement définis :

Dans cette étude, nous nous sommes concentrées sur deux types de contextes locatifs où *être* et *aller* peuvent alterner et éventuellement constituer deux variantes alternatives dans un même paradigme, les contextes où *être* / *aller* est suivi d'une préposition spatiale (*j'ai été / je suis allée à la piscine*) et les contextes où *être* / *aller* est suivi d'un infinitif (*j'ai été / je suis allée voir ma mère*). (Glikman et Patard 2022 [35] : 2)

Nous avons malheureusement eu connaissance seulement après la publication de l'article d'Agnès Steuckardt et Béatrice Dal-Bo (2018) qui porte sur le même sujet. Cet article se concentre également sur les contextes suivis de l'infinitif et sur la période 1914-1916, permettant d'établir une comparaison fine entre *Frantext* et le *Corpus 14*. Dans notre contribution, nous avons également interrogé le *Corpus 14*. Nous sommes toutefois parvenues à un résultat légèrement différent du comptage des occurrences de *être* mis pour *aller* suivi d'un infinitif. Cette différence peut peut-être s'expliquer par le fait que nous avons inclus dans la recherche les formes passées de *être* construites avec l'auxiliaire *être* (relativement fréquentes dans le corpus), ce qui nous donne un nombre légèrement plus élevé de *être* mis pour *aller* que dans les résultats de Steuckardt et Dal-Bo. Notre étude est moins poussée que la leur pour ce qui concerne l'analyse sémantique des variantes, mais elle rejoint globalement leurs résultats quant à l'absence claire de distribution sémantique des deux formes. L'article de Steuckardt et Dal-Bo prend également en compte le discours grammatical. Notre contribution rejoint leurs résultats pour ce qui concerne l'évolution du discours grammatical, d'une opposition sémantique à une opposition de registre.

Dans l'une comme dans l'autre étude, il manque l'observation des autres contextes dans lesquelles les variantes *avoir été* / *être allé* peuvent alterner, en particulier les contextes de type *cette robe te va bien* et *comment ça va*, pour lesquels, au contraire, la forme *avoir été* est

privilegiée, y compris dans le discours métalinguistique, où l'emploi de *aller* apparaît comme une hypercorrection (cités d'après Glikman et Patard 2022 [35] : 3) :

En outre, dans les emplois figurés, *avoir été* l'emporte largement : Cette horloge va bien, elle n'a jamais si bien ÉTÉ. [...] Les phrases suivantes, à cet égard, paraissent étranges : Et tu as quitté librement Ferbannes ? / - Oh ! ce n'est pas allé tout seul ! (Billy, Madame, p. 250.) (Grevisse et Goosse 152011 : 1109, §833)

Si *aller* est employé au sens figuré en parlant de la santé, d'un vêtement, d'une horloge, de qqch. qui "va bien", on doit, aux temps composés, employer *être* (*avoir été*) : *Il a été plus mal hier, il va mieux aujourd'hui. Ce commerce va bien, a bien été. Ce chapeau lui va bien. Cette robe ne lui va pas bien et d'ailleurs ne lui a jamais bien été.* (Hanse 1989 : 65)

Cette poursuite de l'étude est à faire. Elle pourra en outre interroger l'éventuelle répartition sémantique qui se serait mise en place, la forme passée de *être* étant préférée dans les contextes abstraits, tandis que *aller* est la forme préconisée dans les contextes concrets. L'étude de *être* pour *aller* constitue également un excellent point d'entrée pour observer les effets de la norme sur les usages. C'est en outre un exemple intéressant de variation sur un temps long, les deux variantes étant en concurrence sur plusieurs siècles. L'évolution des discours grammaticaux sur la question nous donne ainsi accès au changement non seulement dans les usages (observables en corpus), mais aussi dans le statut des variantes.

Les études sur les rapports entre normes (prises ici en tant que discours normatif préconisé) et usages sont souvent, de fait, menées à partir de variantes explicites, la forme rejetée par la norme et celle préconisée constituant les deux variantes, le contexte étant donné par le discours normatif lui-même. C'est dans ce type d'approche que j'ai mené l'étude sur l'accord de *tout* devant adjectif avec Bérengère Bouard (Bouard et Glikman 2018 [21]), celle sur le développement de l'accord de *un/une espèce de* avec le nom qui suit avec Thomas Verjans (Verjans et Glikman 2022 [36]). Enfin, les variantes régionales correspondent également à des variantes explicites, que nous avons exploitées dans nos travaux dans le cadre des projets *Français de nos Régions* et *Donnez votre Français à la Science* (cf. Avanzi et al. 2016 [16], Glikman et al. 2018 [22], Goldman et al. 2018 [24]). Autant que possible, j'essaie de faire travailler les étudiants sur ce type de variantes explicites. Il y a quelques années, j'avais ainsi un cours de linguistique diachronique en master Sciences du langage qui se prêtait bien à des exposés que les étudiants préparaient en groupe, sur de telles concurrences, facilement délimitables pour une première approche des requêtes en corpus et de la variation.

Mes travaux récents sur *de façon (sorte/manière) que / de façon (sorte/manière) à ce que* présentés à *Diachro X* portent également sur des variantes explicites. Dans ce dernier cas, j'ai

cependant introduit en outre une comparaison avec les constructions infinitives équivalentes, comme origine possible des variantes émergentes par analogie. Lors de l'exposé, on m'a notamment interrogée sur une éventuelle opposition sémantique entre *de façon que* et *de façon à ce que*. La locution *de façon que* peut elle-même exprimer la manière, la conséquence ou consécution, ou encore le but. Bernard Combettes, en particulier, m'a fait remarquer qu'il était possible que se soit mise en place une répartition sémantique entre *de façon à ce que* et *de façon que*, la variante émergente *de façon à ce que* apparaissant avec un effet de sens spécifique et plus restreint. B. Combettes suggère ainsi la mise en place d'une distinction *de façon à ce que* pour exprimer le but (entraîné par le sémantisme de *à* + infinitif) et *de façon que* pour la manière. Cette suggestion reste à vérifier en corpus, la distinction entre les deux effets de sens n'étant en outre pas toujours évidente à établir par des éléments contextuels vérifiables. Cependant, dans l'étude du changement, le fait qu'une variante commence à apparaître dans un contexte spécifique est tout à fait possible, voire attendu. La question serait alors d'étudier le moment où cette nouvelle variante soit étend ses contextes, soit supplée complètement l'ancienne variante dans ce contexte de départ, mettant alors en place une distribution sémantique des variantes. En outre, il me semble qu'il faut absolument prendre en compte le mécanisme de remotivation sémantique lors de l'apparition de variantes, les locuteurs ayant tendance à vouloir trouver une distinction entre deux formes équivalentes⁷⁰. Cette distinction peut alors être de registre (avec la stigmatisation d'une forme), sémantique, ou encore distributionnelle, comme avec ce commentaire que nous avons recueilli dans l'enquête sur *être allé / avoir été* : « Au pluriel je dis « nous sommes allés » », France, 47 ans, réponse « avoir été » (cité dans Glikman et Patard 2022 [35] : 13). Elle n'apparaît cependant pas dès le départ, si l'on se réfère à l'approche proposée par Mario Barra-Jover, dans laquelle la variante est d'abord *invisible*, suivant un schéma de type « Opposition paradigmatique > variante invisible > variante visible partielle > variante visible générale > disparition de l'une des possibilités » ou « Variante invisible > variante visible partielle ou générale > re-motivation de la variante » (Barra-Jover 2009 : 111). Ces questions soulèvent selon moi une piste intéressante à exploiter pour comprendre le changement linguistique. Le moment où la communauté change de point de vue sur la variante pourrait constituer une des phases du changement linguistique. Analyser ce moment de bascule contribuerait ainsi à notre compréhension du changement linguistique (→ 4).

⁷⁰ Cette tendance est probablement liée au besoin d'économie du langage, qui tendra à réduire les formes redondantes en leur attribuant une éventuelle nouvelle fonction, selon le principe une forme = une fonction.

2.3.4. *Le texte comme variable*

Dans une autre partie de mes travaux, d'une certaine manière, c'est le texte même qui constitue la zone de variable, dans lequel nous avons observé la répartition de différents types de variantes. Bien évidemment, dans les autres cas l'observation des variantes se fait également dans les textes et l'analyse de leur représentation dans les différents types de textes est prise en compte. Il me semble important cependant de présenter ici certains travaux dont l'approche diffère de ceux que j'ai présentés jusqu'à maintenant. Je classe dans cette catégorie les études que j'ai réalisées avec Nicolas Mazziotta sur les spécificités syntaxiques de l'oral représenté et des genres textuels ainsi que les travaux auxquels j'ai collaboré sur les chaînes de références.

Avec Nicolas Mazziotta, nous avons en effet mené des recherches sur les spécificités de l'oral représenté en ancien français avec une approche de type statistique (Glikman et Mazziotta 2013 [12]). Dans cette étude, nous sommes partis de l'observation de la réalisation des fonctions syntaxiques dans deux catégories posées a priori : les passages identifiés comme de l'oral représenté et ceux identifiés comme de la narration, dans les quelques mille premières phrases de *La Queste del Graal* (édition numérique de Christiane Marchello-Nizia et Alexei Lavrentiev, disponible sur la BFM). Cette étude s'appuie sur l'annotation syntaxique que nous avons menée dans le projet SRCMF évoqué en introduction. En cela, elle était complémentaire de l'étude de Céline Guillot-Barbance et de ses collègues (Guillot et al. 2013) publiée dans le même volume, qui se basait sur la représentation des catégories morphosyntaxique (parties du discours) dans des textes de la BFM. Ainsi, dans cette étude, les variantes sont réellement toutes les différentes fonctions syntaxiques possibles (telles que définies dans le projet SRCMF) : sujet, objet, circonstant... mesurées à l'échelle du texte, et mises en relation avec une opposition discours direct / narration. L'application d'un test χ^2 et d'un test de chance a ensuite été utilisé pour mettre au jour les associations significatives, ce qui nous a permis de montrer la pertinence de l'opposition :

Nous avons posé *a priori* une opposition entre les deux types discursifs que sont le discours direct et la narration, et l'étude que nous avons menée a permis de prouver la pertinence de cette distinction par une analyse fondée exclusivement sur les structures syntaxiques. Les deux types discursifs, soumis à des mécanismes d'articulation énonciative similaires en raison du médium employé et obéissant à la même grammaire, présentent des différences nettes dont l'intérêt pour l'histoire du français et la typologie linguistique est prépondérant.

Qui plus est, ces différences confirment une opposition qui s'aligne sur la polarité *oralité* vs *écrit*, car les structures préférées les plus prégnantes rapprochent l'oral représenté de l'oral observé, alors que la narration emploie préférentiellement des articulateurs propres

à l'écrit. Tous ces faits nous inclinent à penser que l'oral représenté a des traits caractéristiques véritables de l'oralité médiévale. (Glikman et Mazziotta 2013 [12] : 57)

Dans Mazziotta et Glikman (2019 [27]), nous avons poursuivi ce travail en incluant deux autres textes de type distinct, *La Chanson de Roland* et *La Conquête de Constantinople* de Robert de Clari, annotés dans le cadre du projet SRCMF. Dans cette étude, il s'agissait toujours d'observer les réalisations des fonctions syntaxiques comme autant de variantes dont on étudie la répartition. Cette fois-ci, nous avons donc, en plus du paramètre discours direct / narration, un paramètre « texte » permettant de mesurer la *distance* éventuelle entre chaque texte. Enfin, nous avons également inclus le paramètre de la distinction entre « phrase » (= propositions principales et indépendantes) et « subordonnée ». Étant donné le nombre de variables prises en compte dans cette étude, nous avons utilisé la méthode de l'analyse factorielle des correspondances multiples : « Il s'agit d'une technique statistique heuristique qui met en évidence, notamment par des représentations graphiques du type de fig. 2⁷¹, les tendances d'attraction et de répulsion de modalités catégorielles » (Mazziotta et Glikman 2019 [27] : 5), qui nous a permis de mesurer quel paramètre avait le plus de poids dans la répartition des fonctions syntaxiques :

Les analyses ont montré que la principale partition est celle qui oppose les propositions indépendantes (→2.2) aux subordonnées et *roland* aux autres textes. [...] Toutefois, l'OR [Oral Représenté] de *roland* partage des caractéristiques avec celui de *qgraal* et de *clari* [...] De manière plus générale, la présence d'une opposition OR/narration ne fait aucun doute pour les textes étudiés, mais il ne s'agit pas du seul paramètre influant sur la syntaxe. Malgré une indéniable stabilité des caractéristiques de l'OR et de la narration dans des textes aussi différents que ceux dont nous avons parlé, notre étude montre également qu'il est nécessaire de distinguer l'analyse des subordonnées de celle des principales. Le type de texte doit aussi être pris en considération. L'opposition entre *roland* et les autres textes est plus forte que l'opposition entre oral représenté et narration. (Mazziotta et Glikman 2019 [27] : 13-14)

Toujours dans une approche du texte comme variable et autour de la question des spécificités liées aux genres discursifs, j'ai participé à plusieurs travaux sur l'étude des chaînes de références, avec mes collègues Céline Guillot-Barbance, Vanessa Obry et Bénédicte Pincemin (Glikman et al. 2014 [14], Obry et al. 2017 [20]), d'une part, et avec mes collègues de Strasbourg, d'autre part, dans le cadre du projet DEMOCRAT (Oberlé et al. 2018 [25], Capin et al. 2021 [33]). Dans ces cas, nous nous sommes intéressés aux réalisations des éléments

⁷¹ Voir p. 578 du volume des travaux.

constitutifs de la chaîne de référence, en décrivant à la fois les paramètres des chaînes elles-mêmes (nombre, longueur...) et des différents *maillons* (= expression référentielle contenue dans une chaîne) dont elles sont constituées, et en particulier leur catégorie morphosyntaxique. Là aussi, l'objectif était de montrer les spécificités des textes à partir de ces caractéristiques, et leur évolution, ce que nous avons pu faire dans une première étude sur des textes narratifs :

Nous avons fait l'hypothèse et avons essayé de montrer que la composition des chaînes de référence varie en fonction des types de textes (récits brefs/longs, récits en vers/en prose, nombre de référents principaux). Comme l'ont démontré de nombreuses études sur la période moderne (voir Schnedecker & Landragin 2014, introduction de ce volume), le genre discursif semble conditionner la forme des chaînes de référence pendant la période médiévale. (Glikman et al. 2014 [14] : 58)

Nous avons ensuite poursuivi cette étude à travers l'analyse des chaînes de référence dans les récits brefs sur un empan temporel plus large :

Notre étude a permis de mettre en lumière plusieurs spécificités des CR des récits brefs anciens, parmi lesquelles certaines se trouvent aussi dans les faits divers modernes (densité référentielle des animés humains, nombre faible et taille réduite des CR, fréquence des noms relationnels) et d'autres sont plus typiques des états anciens du français et peuvent être relevées autant dans les genres brefs que dans les récits longs (fréquence du sujet zéro, stabilité et monotonie fortes ; voir Glikman, Guillot-Barbance & Obry 2014). Notre analyse montre, par ailleurs, une évolution diachronique dans la composition des CR : le statut d'indicateur de saillance élevée d'abord attaché à l'article indéfini semble s'atténuer au XVI^e siècle et le nom propre paraît le remplacer dans ce rôle. De nouvelles expressions, encore rares, se développent à partir du XV^e siècle (*ledit*, *lequel*). Cette évolution serait à confirmer avec d'autres textes des mêmes périodes.

Ce qui ressort en outre de notre étude, c'est l'impact du type de référent, central ou non, sur les CR. Ce paramètre, même s'il n'est pas nécessairement mis en avant de manière explicite dans les travaux sur les CR, est toujours primordial, et ce sont les chaînes les plus longues, celles qui correspondent aux référents centraux, qui sont les plus étudiées et qui permettent les contrastes d'un genre à l'autre. Dans le cas présent, les changements qui affectent ces CR caractérisent plus largement l'évolution des récits brefs et l'individuation croissante des personnages importants (marquée notamment par la progression de la fréquence du nom propre en début de CR). (Obry et al. 2017 [20] : 108)

Dans Oberlé et al. (2018) [25], nous nous sommes centrés sur l'étude de textes encyclopédiques. Pour la période médiévale, nous avons pris en compte les bestiaires, qui ont notamment en commun le fait qu'ils comportent des rubriques centrées sur des animaux (catégorie retenue pour comparaison dans les textes postérieurs) et une visée didactique. Cette étude nous a permis de mettre en avant l'évolution diachronique du genre à travers l'analyse des chaînes de référence.

Ces approches nécessitent la prise en compte d'un grand nombre de données, en particulier de données enrichies d'informations complémentaires, comme nous venons de le voir, ce que le développement des corpus annotés rend possible. Cependant, les approches quantitatives ne remplacent pas les approches qualitatives, dont elles sont souvent complémentaires, comme nous le montrons dans Capin et al. (2021) [33] :

Premièrement, il y a une nette différence entre la coréférence révélée par l'annotation opérée sur un grand ensemble de textes, qui ne semble guère pouvoir être envisagée en dehors d'une conception paradigmatique des phénomènes (Schnedecker 2019) et ce qu'elle est quand elle est filtrée par notre connaissance empirique des textes et de leurs modes de structuration. Nous avons ainsi montré que la notion de chaîne de référence n'est opératoire que dans certaines des configurations évoquées supra, à savoir les cas où la coréférence opère intra-paragraphiquement (et/ou inter-paragraphiquement proche). Dans les cas d'une trop grande distance entre les mentions, la notion de chaîne n'est pas forcément très intéressante, ce qui a pu expliquer que certains travaux antérieurs (Schnedecker, Glikman & Landragin 2017) ont choisi d'opter pour un redémarrage des chaînes à chaque nouveau paragraphe. (Capin et al. 2021 [33] : 103-104)

Les approches statistiques et les mesures sur un grand nombre de données enrichies peuvent ainsi fournir des résultats intéressants. Les variantes, dans ce cas, sont définies a priori, généralement dans le modèle d'annotation, comme nous avons pu le voir dans l'exemple de l'étude des spécificités syntaxiques du discours direct (Glikman et Mazziotta 2013 [12] et Mazziotta et Glikman 2019 [27]), qui s'appuie sur l'annotation syntaxique menée dans le projet SRCMF. Le travail sur corpus ne remplace cependant pas la connaissance fine des textes et des contextes, peut-être encore plus en diachronie, comme j'essaie toujours de le transmettre à mes étudiants, seul moyen de comprendre les réalisations des variantes « en système ». Enfin, de telles approches nécessitent également de prendre clairement la mesure de ce que constitue vraiment la variante à l'étude. Dans les cas présentés, j'avais moi-même participé à l'annotation des corpus et aux discussions à l'origine de l'établissement des choix d'annotation, dont j'avais donc une pleine connaissance, mais dont il importe de documenter explicitement les choix lors de la diffusion (→ 3.2.).

Ces études, en interrogeant les rapports entre discours direct et narration ainsi qu'entre les différents genres textuels, contribuent ainsi à la question des paramètres à prendre en compte dans la variation que j'aborde dans le chapitre suivant (→ 3.).

3. Analyser la variation. Autour des données et des méthodes : une approche en linguistique variationnelle

3. Analyser la variation. Autour des données et des méthodes : une approche en linguistique variationnelle

*Un petit poisson, un petit oiseau
S'aimaient d'amour tendre
Mais comment s'y prendre
Quand on est dans l'eau*

Juliette Gréco, *Un petit poisson, un
petit oiseau*, 1966.

La conception du changement linguistique à laquelle j'adhère repose sur une approche qui place la variation au centre du processus. Après avoir présenté (→ 2.) la problématique de l'établissement des variantes, comme observables de la variation, nous allons maintenant aborder la question des paramètres, ou axes, de la variation. Sans prétendre à l'exhaustivité sur cette question largement débattue, je mettrai l'accent sur les points que je retiens dans mes travaux, montrant ainsi mon ancrage dans la linguistique variationnelle⁷² (→ 3.1.).

Autour de ces thématiques se retrouve la question d'évaluer les choix. Une fois établies les variantes, nous abordons dans ce chapitre la manière d'établir leurs caractéristiques, leurs critères de distribution et conditions de réalisation, en d'autres termes, les paramètres selon lesquels elles varient. Ce genre d'analyse contribue bien sûr à l'étude des phénomènes eux-mêmes, mais cela nous apprend aussi quelque chose sur la variation elle-même, sur ses paramètres et sur leur validité pour la description. Enfin, cela contribue à notre compréhension du changement. Évaluer les choix permet d'étudier le rapport entre grammaire et usages. Étudier la variation permet ainsi de comprendre le passage entre contrainte et choix. Mon objectif est de saisir le moment de bascule : pour moi, le changement linguistique a lieu quand la variante change de statut. Ainsi, en diachronie, étudier la variation ne doit pas se limiter à étudier la répartition des variantes, mais bien à étudier le changement dans la répartition, dans le statut des variantes (cf. Glikman 2019 [26], Glikman & Patard 2022 [35]). Cela correspond à l'analyse des changements dans la phase de l'évaluation, pour reprendre la terminologie du cadre de Weinreich, Labov et Herzog (1968). L'objectif, à travers l'établissement du statut des variantes, est d'établir les préférences d'une langue à partir des choix et des représentations.

⁷² Je retiens ici le terme de « linguistique variationnelle », à la suite de Völker (2009), Glessgen, Kabatek et Völker (éds) (2018), Dufter (2018), qui distinguent « linguistique variationnelle » des approches sociolinguistiques, auxquelles est plutôt associé le terme de « linguistique variationniste » (cf. par ex. chez Poplack et Dion 2021), voir *infra*.

Ces notions d'évaluation et de préférence sont pour moi au centre du changement et seront abordées au chapitre suivant (→ 4).

À la suite de ces questions d'ordre théorique, j'aborderai la question plus pratique de l'accès aux données qui permettent d'analyser la variation. Comme nous l'avons vu avec la question des variantes, établir des contextes strictement comparables n'est pas toujours chose facile. Cela est d'autant plus vrai si l'on veut mesurer les différents paramètres de variation. Ce phénomène est bien connu en diachronie et, d'une manière générale, en linguistique de corpus, autour des questions de leur représentativité, y compris pour les périodes contemporaines. Comment comparer ce qui n'est pas toujours comparable, *comment s'y prendre* ? Quelles données nous permettent d'avoir accès aux différents paramètres ? Ce sera l'occasion de parler des méthodes ainsi que de la constitution et du partage des données avec la communauté (→ 3.2.).

3.1. Quels paramètres pour la variation ?

*Je manque d'application comme les GSM
T'as réparé le mal que j'avais laissé
T'es mille fois trop bonne pour que j'aie le seum
Mais là, tu me contrôles et ça tu le sais
Tu continues à danser sur des hits sales
T'as mis un pull XL, on voit quand même tes lolos
J'pourrais être ton beau gosse, j'pourrais t'faire du vice car
T'allumes toutes les flammes et j'suis docile comme un boloss*

Thérapie Taxi ft. Roméo Elvis, *Hit Sale*, 2017.

3.1.1. Quelques éléments de définition

La notion de variation est désormais largement répandue dans différents travaux linguistiques, la plupart des courants considérant comme acquis le fait que les langues varient, quelle que soit la place qu'ils accordent à l'étude de la variation en elle-même. Introduite et diffusée notamment dans le cadre de la sociolinguistique labovienne, elle s'est étendue à d'autres approches, mais reste divisée entre les approches centrées sur le système et celles centrées sur le locuteur :

Ce sont les travaux du sociolinguiste américain William Labov (entre autres, l'ouvrage fondateur de 1972) qui, à travers une école très productive, le variationnisme, ont implanté le terme en sociolinguistique, souvent non sans un certain flou au-delà du constat que toutes les langues vivantes varient. Deux directions des travaux sur la variation peuvent être observées, centrées d'un côté sur le fonctionnement du système, de l'autre sur le rôle des producteurs de discours. La notion de variation se situe ainsi au centre de débats majeurs des sciences du langage, engageant rien moins que la

conception de la langue et ses effets sociétaux. Et c'est sans doute pourquoi on rencontre aujourd'hui le terme dans des travaux fort divers, dont beaucoup n'ont pas d'orientation sociolinguistique. (Gadet 2021 : 331)

Cette distinction a pu servir de démarcation entre les courants qui relèveraient d'une sociolinguistique et ceux qui se réclament d'une linguistique variationnelle :

Il est vrai que, tout comme la pragmatique linguistique et la sociolinguistique, la linguistique variationnelle est basée sur l'ouverture vers l'extralinguistique. Mais alors que la plupart des travaux en sociolinguistique et pragmatique visent en premier lieu à l'analyse du contexte social et du lien avec les objectifs communicationnels poursuivis par le locuteur, la linguistique variationnelle cherche à interpréter la variation comme faisant partie d'un (sous-) système linguistique, c'est-à-dire à identifier les régularités normatives [en note : Le terme se réfère ici aux effets de solidarité qui forment des sous-systèmes, non à la norme scolaire/prescriptive.] auxquelles elle est soumise. (Völker 2009 : 33)

Cependant, comme le conclut Françoise Gadet, « La notion de variation, très présente dans la littérature linguistique actuelle, conserve une part de flou, tout en s'avérant difficilement contournable pour les (socio)linguistes. » (Gadet 2021 : 335). La question de la variation, en particulier de sa théorisation, est ainsi au centre des préoccupations de nombreux chercheurs, comme en atteste la création de la revue *Language Variation and Change* aux Presses universitaires de Cambridge dans les années 80⁷³, les cycles de conférences DIA plus récemment⁷⁴ et le nombre de publications liées. Elle est encore matière à débat, si l'on en croit la récente publication de 2018 des actes du colloque DIA IV dont le titre est en lui-même évocateur : *Repenser la variation linguistique* (Glessgen, Kabatek et Völker (éds) 2018). La problématique des paramètres à prendre en compte fait plus encore l'objet de larges controverses, que ce soit autour du statut à accorder à chaque paramètre, de leur définition (voire de leur légitimité), des délimitations de leur champ ou encore des rapports qu'ils entretiennent les uns avec les autres. Mon objet n'est pas ici de faire un état de la question sur le sujet, tant de la variation que de ses axes, mais de montrer le cadre dans lequel s'inscrivent mes travaux. Je ferai ainsi le point sur les paramètres que j'ai explorés jusqu'ici, pour montrer à la fois la continuité et l'évolution de mes recherches. Cela me permettra de mettre en avant la manière dont mes travaux et les travaux des étudiants que j'encadre s'intègrent dans le champ de la linguistique variationnelle.

⁷³ <https://www.cambridge.org/core/journals/language-variation-and-change>. La revue, dont l'un des principaux éditeurs est W. Labov, est d'ancrage sociolinguistique.

⁷⁴ Voir Glessgen, Kabatek et Völker (2018) pour l'historique de la création de ces conférences.

Jusqu'ici, j'ai souvent mentionné le problème de la délimitation entre la contrainte et le choix. Dans cette optique, j'évoquais, en premier lieu, la limite entre la contrainte grammaticale et le choix (de type stylistique), qui peut alors relever d'un paramètre de variation. Comme le rappellent Dufter et Stark (2003), la définition même de la variation suppose l'existence d'un choix pour le locuteur :

Nous proposons comme point de départ théorique la notion de variation pour désigner le fait que les locuteurs, dans certains cas, semblent avoir le choix entre plusieurs formes et constructions ayant la même signification ou bien entre plusieurs réalisations d'un même élément linguistique. (Dufter et Stark 2003 : 81)

La notion de choix serait ainsi ce qui distinguerait les éléments liés à la variation, en particulier conçue comme extérieure au système :

la variable est vue comme un choix en un point unique du système, choix d'où découlent les variantes, et sur lequel pèsent des facteurs contextuels (Godard 1992 : 52)

Du point de vue qui nous occupe, le fait que les variables soient ou non associées à une dimension sociale ne met pas en jeu la vision du système syntaxique : le système offre des possibilités, qui se trouvent être ou ne pas être exploitées par des systèmes autres que le système linguistique. (Godard *ibid.* : 60)

On peut comprendre à partir de cela qu'il faut que le choix soit libre, du point de vue de la grammaire, pour que les facteurs de variation extralinguistiques puissent être considérés. C'est sur ce point qu'on peut établir la distinction entre la notion de distribution (au sens grammatical) et de variation. Cet aspect a vu les premiers lieux de développement de l'étude de la variation : « Faute de mieux, on se bornait à constater des variantes libres là où on ne trouvait pas de règle de distribution » (Dufter et Stark 2003 : 81).

La question du choix n'est cependant pas si évidente, y compris dans les paramètres de la variation, comme le rappellent Glessgen et Schøsler : « L'on peut simplement retenir que la liberté de choix est moindre pour la diatopie que pour la diastratie et qu'elle est plus développée pour la diaphasie. » (Glessgen et Schøsler 2018 : 22). La possibilité pour le locuteur de choisir n'est pas toujours vraie et peut varier, à la fois selon l'axe (diatopique, diastratique ou diaphasique⁷⁵) et selon l'époque. Tout le travail du linguiste pourrait alors être de chercher à établir dans quelle mesure un usage relève d'un choix ou d'une contrainte et comment cette distinction évolue (cf. Glessgen et Schøsler 2018 : 31). En outre, comme nous l'avons évoqué

⁷⁵ Définis traditionnellement respectivement comme variation selon l'espace, selon la stratification sociale ou selon la situation (cf. « les occasions, circonstances et finalités de l'emploi de la langue dans le discours » (Coseriu 1998 : 9)).

(→ 2.1.) à propos de l'établissement des variantes, analyser la variation en syntaxe et l'intégrer à la grammaire ne va pas de soi (cf. Dufter, Fleischer et Seiler (éds) 2009).

Dans tous les cas, il importe de commencer par vérifier les facteurs de distribution de type syntaxique ou morphosyntaxique, comme la personne, le temps, la présence de la négation, avant d'élargir aux autres facteurs. C'est ainsi que j'ai procédé dans certains travaux (cf. Glikman et Verjans 2013 [13]) et c'est ce que je demande toujours à mes étudiants de faire, afin d'établir les contraintes éventuelles sur la réalisation d'un phénomène. Cependant, dans certains cas, les contraintes distributionnelles ne constituent pas une distribution complémentaire uniquement syntaxique, mais font bien partie d'un processus de variation à l'origine d'un changement. Cela a été montré pour le développement du pronom sujet (Prévost 2011a), pour la disparition de la déclinaison (Schøsler 1984), ou encore avec les verbes parenthétiques (Blanche-Benveniste et Willems 2007) et le développement des marqueurs discursifs à base verbale (Dostie 2004). Avec Stefan Schneider, nous avons notamment mis en évidence comment ces verbes parenthétiques peuvent en outre être considérés comme de véritables marqueurs d'oralité (au sens conceptuel) (Glikman et Schneider 2018 [23]).

3.1.2. Quels paramètres retenir ?

À la suite de Glessgen et Schøsler (2018), je pense que tout énoncé est porteur en lui-même de toutes les dimensions de la variation et que l'analyse gagnerait à prendre en compte la mesure de ces dia-variations en termes de degré de marquage :

tout énoncé comporte en même temps une dimension diatopique, diastratique et diaphasique, s'inscrit dans le continuum entre immédiat et distance, représente une modalité médiale et appartient à un genre textuel donné. Autrement dit : tout énoncé fait appel à chacune des dimensions diasystématiques qui sont ainsi à tout moment coprésentes. (Glessgen et Schøsler 2018 : 17)

C'est également la position adoptée par Siouffi (2020) :

Les sujets s'expriment d'une manière ou d'une autre parce que des paramètres extérieurs influent sur les situations de communication – paramètres classiquement divisés en diatopiques, diastratiques et diaphasiques, éventuellement diamésiques (→ chap. 6), et aussi parce qu'ils répondent à des stratégies pragmatiques particulières. (Siouffi 2020 : 68)

Mais même cela posé, il n'en reste pas moins que la distinction entre les différentes dimensions n'est pas toujours simple (cf. entre autres Glessgen et Schøsler 2018, Dufter et Stark 2003, Dufter 2018). Cela est vrai à la fois du point de vue théorique et lorsqu'il s'agit de

Analyser la variation. Autour des données et des méthodes : une approche en linguistique variationnelle

caractériser les faits langagiers. Pour Coseriu (1998), il s'agit là d'un faux problème, puisque tout énoncé s'inscrit, de fait, dans toutes les dimensions :

Tout fait de langue est *en même temps* fait dialectal, fait de niveau et fait de style (du moins, en tant que fait appartenant à la langue commune et au niveau et au style « neutres ») et devrait être « indexé » comme tel dans le lexique et dans une grammaire de la variété de la langue historique. (Coseriu 1998 : 14)

Cela ne l'empêche pas de conserver les notions de dialectes, niveaux et styles, en tant qu'unités « par définition *unitaires* (homogènes) d'*un seul point de vue* » (*ibid.* 14), tout en précisant plus loin :

Un seul et même système linguistique peut être en même temps dialecte et niveau, dialecte et style de langue, niveau et style de langue, selon la perspective dans laquelle il est considéré et selon la fonction qui lui revient dans la communauté (*ibid.* 15)

La plupart des études ultérieures ont cependant souvent cherché à catégoriser les faits, première étape vers une identification des idiolectes, quand bien même on l'aborderait sous la forme de continuum. Comme l'ont rappelé Glessgen et Schøsler (2018), la classification des phénomènes dans des axes peut en outre n'être que temporaire, le principe même de la diffusion étant basé sur l'extension progressive des « contextes » d'emploi et donc, à mon sens, de la répartition dans les « axes », point sur lequel je les rejoins pour dire qu'il s'agit là d'un domaine d'étude fructueux :

De notre point de vue, la grande variabilité dans la composition des axes à un moment historique donné, qui est réelle et incontournable, représente tout au contraire un noyau d'interrogations riches et porteuses. (Glessgen et Schøsler 2018 : 45)

Il me semble important de distinguer la question de la catégorisation des faits, le « faux problème » évoqué par Coseriu, de celle de l'établissement de variétés. La notion de variété est largement répandue et utilisée, tant en sociolinguistique qu'en linguistique variationnelle, y compris par les locuteurs eux-mêmes, qui peuvent avoir conscience de la diversité (cf. Siouffi 2020). Le cas des dialectes est l'exemple le plus connu, mais ce n'est pas le seul, les locuteurs ayant également conscience des distinctions des autres axes, même diachronique. Les phénomènes d'hypercorrection, d'accommodation et d'insécurité linguistique viennent, de fait, de l'existence (et de la conscience) de variétés hautes, de prestige et de variétés basses. L'antiquité avait déjà conscience de ce phénomène, qualifié de « style » (cf. Siouffi 2020 : 121). L'une des difficultés théoriques et terminologiques tient aussi, comme me l'a fait remarquer Dominique Lagorgette, dans le fait que les concepts liés à la variation sont utilisés conjointement par la sociolinguistique et par la linguistique variationnelle, mais d'une manière

différente, voire à des fins différentes. Il me semble cependant qu'à notre époque, la notion de groupe se définit davantage par une identité linguistique que par une identité sociale. Le groupe social reste une réalité, tout comme la notion de prestige ou d'insécurité, mais la manière dont on le constitue demande peut-être à être réévaluée. Julie Peuvergne, qui a beaucoup travaillé à la constitution de corpus avec Françoise Gadet ainsi qu'avec Bernard Laks dans le cadre de PFC notamment, me rappelait à ce sujet l'importance de « laisser émerger » les catégories « par le bas », plutôt que de les plaquer à priori sur les données.

La question de l'analyse des paramètres de la variation se pose plus encore lorsqu'il s'agit d'accéder aux données nous permettant d'étudier les différentes dimensions (→ 3.2.). Pour des états anciens de langue, l'accès aux différents paramètres de la variation s'avère d'autant plus compliqué que les productions qui nous parviennent sont fixées sous forme de textes, dont les « genres » constituent le principal observable pour certaines de ces variations, en particulier la diaphasie :

Le cas est particulièrement flagrant en linguistique diachronique où la carence et l'hétérogénéité des données permettent très difficilement une évaluation indépendante de chacun des facteurs diasystématiques. (Glessgen et Schøsler 2018 : 38)

Si déjà la diatopie pose de solides problèmes pour l'observation ciblée, cela est encore plus vrai pour la diastratie et la diaphasie, qui sont pratiquement inséparables dans l'observation à l'époque contemporaine. Pour les époques historiques, les moyens restreints de nos sources mènent souvent à une neutralisation entre diastratie et diaphasie, plus encore entre diaphasie et genres textuels : ces derniers sont le seul lieu d'observation concret pour la diaphasie (cf. Glessgen 2005).

L'analyse diasystématique doit se résoudre par conséquent à l'idée que ses paramètres de base ne sont pas atteignables de manière directe et immédiate et qu'il est essentiel d'admettre une part importante de réflexion épistémologique dans tous les travaux de type variationniste. (Glessgen et Schøsler 2018 : 44)

Sans aller plus loin sur ces questions théoriques, je vais consacrer les paragraphes qui suivent à montrer comment j'intègre les différents paramètres de la variation à ma recherche.

En tant que diachronicienne, j'ai toujours abordé la variation, et la question des paramètres de variation lui est intimement liée dès les premières approches. En effet, dès ma thèse, j'avais rencontré ce type de questionnement : comment analyser les différences que l'on peut observer entre *La Chanson de Roland* et *La Conquête de Constantinople* de Robert de Clari, quand on sait que ces deux textes diffèrent tant du point de vue de la forme (vers / prose), du genre (chanson de geste / chronique), de leur date de composition (10^e / 13^e siècle), que de leur marquage diatopique (anglo-normand / picard) ? C'est ce qui m'a souvent conduite à centrer mes études sur un fonctionnement au sein d'un micro-système (le texte en tant que témoin). Ces

analyses peuvent ensuite être comparées, mais les conclusions à tirer de ces comparaisons sont toujours à prendre avec précautions, sur leur portée et ce que ça nous apprendrait sur l'établissement d'une variété. De fait, je ne parle jamais de variété dans mes travaux, et j'attire toujours l'attention des étudiants sur la difficulté de définir des variétés et l'importance de la notion de continuum. Les frontières sont toujours posées à posteriori, y compris (et surtout) en diachronie. Ainsi, je commençais mon cours d'introduction au moyen français, en deuxième année de licence Sciences du Langage, par l'observation des introductions de différentes grammaires du moyen français (par ex. Marchello-Nizia 1997 et Martin et Wilmet 1980), en particulier leur délimitation temporelle de leur objet, entre les grammaires à empan large et celles à empan réduit. J'aborde ainsi avec eux la question de la périodisation, qui peut varier selon les points de vue, les éléments pris en compte évidemment et, surtout, élément important selon moi que je mets particulièrement en avant auprès des étudiants, selon les objectifs. Ainsi, le degré de granularité adopté peut dépendre de la perspective du chercheur, comme l'introduction de la notion d'un « très ancien français » ou d'un « français pré-classique ». J'insiste également auprès des étudiants sur le fait que personne n'a conscience du changement et sur le principe de l'intercompréhension de proche en proche, aussi bien en diachronie qu'en diatopie. Cela n'empêche évidemment pas d'avoir une certaine perception du changement, tant dans les innovations, qui peuvent être remarquées, que dans le sentiment de « vieillissement » langagier, préoccupations au centre de nombreux travaux sur le changement linguistique, comme en témoigne le récent numéro de *Linx* « Entre vieillissement et innovation : le changement linguistique » coordonné par Gaétane Dostie, Sacha Diwersy et Agnès Steuckardt (2021). J'ai ainsi proposé dans notre dernière offre de formation la création d'un cours « interface synchronie / diachronie : variation et changement en cours » en deuxième année de licence de Sciences du langage. L'objectif était de montrer aux étudiants le lien entre variation et changement et, si possible, de leur montrer que la diachronie n'est pas qu'une affaire de « vieux livres poussiéreux ». Pour essayer de leur montrer des exemples plus proches d'eux et qui peuvent tout à fait illustrer les notions d'innovation et de vieillissement, je leur faisais analyser une chanson d'Aya Nakamura, *Djadja*⁷⁶, connue pour son lexique particulier. De même avec la chanson *OK Boomer* de McFly et Carlito⁷⁷, qui reprend l'expression qui s'est popularisée récemment, illustrant ainsi le concept de diffusion, tout en associant aux

⁷⁶ Cf. <https://youtu.be/iPGgnzc34tY>.

⁷⁷ Cf. <https://youtu.be/1-ac8jxb66U>.

« boomers » des expressions ressenties comme vieilles. À côté de cela, je donnais aussi des passages d'une enquête de Maigret de Simenon pour illustrer à la fois la proximité de la langue et, déjà, le vieillissement, par exemple dans les expressions à renouvellement rapide (cf. « faire la bombe »). À l'inverse, il me faut toujours clarifier auprès des étudiants une habitude qui s'est installée dans la répartition des cours à l'Université de Strasbourg. En effet, par tradition, les cours à Strasbourg sont divisés en trois grandes catégories, « linguistique synchronique », « linguistique diachronique » (parfois « linguistique romane et diachronique ») et « phonétique ». Cependant, comme je le précise aux étudiants, certains cours de « linguistique diachronique » peuvent s'avérer être des cours de *linguistique synchronique d'états anciens de langue*, comme les cours d'ancien français, de moyen français ou de français pré-classique. Même si, de fait, ces états anciens sont le plus souvent mis en perspective et replacés dans une histoire de la langue, il n'en reste pas moins que reconnaître les verbes au présent du subjonctif dans un texte du 12^e siècle a tout d'une approche synchronique. Cette mise au point entre synchronie et diachronie se retrouve dans l'évolution des appels à communication de la section diachronique du *Congrès Mondial de Linguistique Française*, qui a ajouté explicitement aux premiers appels cette distinction, notamment sur l'impulsion de Sophie Prévost, coordinatrice de la section :

Diachronie et histoire de la langue

L'étude des changements du français, qui, pendant deux siècles, a surtout fait l'objet d'analyses empiriques, connaît depuis quelques décennies un renouveau théorique capital.

La section « Diachronie, histoire de la langue » aura pour objectif ce renouveau. La diversité des approches (grammaticalisation, grammaire universelle, réanalyse, typologie diachronique, réflexion sur l'analogie, sociolinguistique historique, linguistique textuelle, pragmatification...) a permis de réelles avancées dans la connaissance de l'évolution du français et des langues en général.

La section invite à soumettre des illustrants de toutes ces approches, dans tous les domaines de la langue, qu'il s'agisse d'écrit ou d'oral, et que les analyses soient plus proprement descriptives ou plus spécifiquement théoriques. (Appel à communication, CMLF 2008)

Histoire du français : perspectives diachronique et synchronique

Les études proprement diachroniques, portant sur l'évolution de phénomènes à travers les siècles ou sur des diachronies courtes (y compris de la langue des XX-XXI^{èmes} siècles) sont encouragées, quel que soit le domaine dont elles relèvent (phonétique, morphologie, syntaxe, sémantique, ou pragmatique), qu'il s'agisse d'écrit ou d'oral, et que les analyses soient descriptives ou plus spécifiquement théoriques. Seront également accueillis des travaux visant à approfondir ou discuter des théories sur le changement.

Enfin, des études synchroniques consacrées à une période ancienne précise, antérieure au XX^{ème} siècle, trouveront également leur place dans cette session. (Appel à communication, CMLF 2014 et suivants).

Dans mes travaux, les différentes dimensions sont inégalement abordées. De fait, la dimension diachronique (elle-même souvent considérée comme à part) y est largement présente, bien qu'à des échelles différentes. Certains travaux, en particulier sur l'ancien français, portent plutôt sur une approche synchronique des états anciens de langue. Cependant, la dimension diachronique n'en est jamais totalement exclue. D'une part, ces états anciens de langue peuvent être mis en perspective avec les états précédents (notamment le latin) et suivants (ne serait-ce qu'avec la prise en compte de l'aboutissement en français moderne). D'autre part, les données auxquelles nous avons accès pour l'ancien français sont par elles-mêmes porteuses de variation, y compris diachronique (cf. Verjans 2012), à travers la mouvance des textes tels qu'ils nous sont transmis (cf. Zumthor 1981, ou variance chez Cerquiglini 1989). C'est pourquoi, dans certains travaux, j'ai privilégié l'étude d'un fonctionnement au sein d'un texte, pour une mise en système à une échelle réduite (Mazziotta et al. 2012 [9], Glikman et Mazziotta 2013 [12]). Dans d'autres travaux, plusieurs textes sont pris en compte, toujours sur la période médiévale (Glikman 2012 [8], Glikman, Guillot-Barbance et Obry 2014 [14], Glikman 2016 [18], Mazziotta et Glikman 2019 [27]). Dans ce cas, les différences observables entre les textes sont toujours à considérer avec précaution, notamment sur les conclusions à en tirer, différents facteurs variant d'un texte à l'autre comme le rappellent Glessgen et Schøsler (2018, cf. ci-dessus) :

Reste encore à étudier si cette particularité de *roland* est liée au genre, à la forme (en vers tandis que les autres textes de l'étude sont en prose), ou encore à l'espace temporel (presque 2 siècles séparant ce texte des deux autres). (Mazziotta et Glikman 2019 [27] : 14)

Dans d'autres travaux, j'adopte une optique de diachronie « large », rendue possible par le développement des corpus et des bases de données. C'est le cas pour les travaux auxquels j'ai collaboré dans le cadre du projet Democrat sur les chaînes de référence, qui consistait en l'élaboration d'un corpus diachronique annoté en référence (Obry et al. 2017 [20], Oberlé et al. 2018 [25], Capin et al. 2021 [33]). C'est également le cas des travaux que j'ai menés pour la GGHF, pour laquelle un corpus diachronique et, autant que possible, équilibré avait été constitué (cf. Prévost 2020 ; Glikman 2020a [29] et b [30]). Les approches en diachronie large m'ont également permis de mener des travaux associant les analyses des discours grammaticaux, en particulier à partir du 16^e siècle, aux études des « usages » en corpus (Bouard

et Glikman 2018 [21], Glikman et Bouard 2021 [31]). Ma recherche s'est dans ce cas nourrie de mon enseignement. C'est en effet le fait d'avoir eu à assurer pendant plusieurs années à l'Université de Strasbourg le cours de français pré-classique et classique en 3^e année de Lettres et de Sciences du Langage qui m'a conduite à m'intéresser de près à la période post-médiévale et au développement de la grammatisation du français. À la suite de mes recherches dans le cadre de ce cours, j'ai proposé un cours d'histoire de l'orthographe et un cours sur l'histoire de la norme en français pour les étudiants de licence de Sciences du Langage et de Lettres. Dans tous mes travaux qui concernent des périodes plus récentes, je prends désormais en compte dès que possible l'évolution du discours métalinguistique sur les phénomènes étudiés, comme autant d'indicateur du statut de la variante et de son évolution (cf. Siouffi, Steuckardt et Wionet 2012 ; Glikman 2019 [26], Glikman et Verjans 2021 [32], Glikman et Bouard 2021 [31], Glikman et Patard 2022 [35]). Comme le montrent Glessgen et Schøsler (2018), les « marquages » peuvent en effet se déplacer d'un axe de variation à l'autre :

le marquage diatopique peut évoluer d'une variation selon l'utilisateur vers une variation selon l'usage. Ils montrent également que les effets de la chaîne variationnelle entre diatopie et diastratie peuvent être de nature variable et que l'action de la diatopie n'a rien de 'mécanique'. (Glessgen et Schøsler 2018 : 31)

Le discours grammatical, quand il existe, et son évolution, de même que le discours épilinguistique (Canut 1998) et les remarqueurs « modernes » (Siouffi, Steuckardt et Wionet 2012), sont de parfaits indicateurs de ces déplacements (→ 4.2.).

Dans mes projets sur les périodes contemporaines, mon optique reste diachronique, bien qu'incluant d'autres paramètres de variation, dont le lien avec la diachronie est justement ce qui m'intéresse. C'est le cas en particulier des enquêtes dans le cadre des projets *Français de nos Régions* (désormais FRR) et *Donnez votre Français à la Science* (désormais DFS). C'est également cette optique diachronique que j'ai mise en avant dans le projet qui m'occupe actuellement, dont l'une des composantes est la constitution d'un corpus de vocaux. J'ai d'abord déposé ce projet sous le titre « Le rapport oral / écrit et les nouveaux modes de communication : étude de diachronie contemporaine » pour ma demande de délégation CNRS au laboratoire ATILF (CNRS & Université de Lorraine) (2021-2023⁷⁸). Je l'ai ensuite adapté, suite aux premiers résultats, dans le projet *Oralité et diachronie : une voie d'accès au changement linguistique* (Idex Recherche exploratoire 2022-2024, Université de Strasbourg). Ce projet vise

⁷⁸ Délégation 2021-2022, incluant un semestre de CRCT au titre du CNU, renouvelée en 2022-2023.

à accéder à des productions de français parlé spontané non surveillé par la constitution d'un corpus de SMS vocaux, à l'image des corpus qui ont été fait pour les SMS écrits (cf. Panckrust et al. 2013, Stark 2016-2018) (cf. Glikman et Fauth 2022 [34] pour une présentation du projet et des premiers résultats).

Concernant les autres axes, de fait, les projets FRR et DFS incluait une bonne part de variation diatopique, que Mathieu Avanzi a exploitée dans différents articles et atlas linguistiques (Avanzi 2017, 2019a, b, 2020). Dans cet axe, on peut observer des phénomènes comme le terme *GSM* présent dans la citation en exergue de cette section (3.1.), clairement distribué diatopiquement et spécifique à la Belgique francophone⁷⁹. En ce qui me concerne, c'était aussi les aspects méthodologiques qui m'intéressaient (cf. Avanzi et al. 2016 [16], Glikman et al. 2018 [22]). Sur les aspects tant diachroniques que diatopiques, les enquêtes menées dans le cadre de ces projets nous donnaient accès à la diffusion des termes, comme nous avons pu le voir dans les quizz DFS, que nous avons pu comparer aux résultats des enquêtes FRR. Les quizz DFS étaient construits sous forme de questions à choix multiples sur la connaissance du lexique régional. Ces questions étaient établies sur la base des résultats des enquêtes FRR. Elles comportaient ainsi effectivement une réponse correcte. Chaque réponse était accompagnée d'un texte informatif et, le cas échéant, d'une carte géolinguistique tirée des résultats des enquêtes FRR. Par exemple, la question correspondante au passage cité infra était : « Qu'est-ce qu'une *chocolatine* ? » et les réponses possibles : « Une tranche de pain recouverte de pâte à tartiner », « Un chocolat fourré à la praline », « Une pâtisserie de pâte feuilletée fourrée d'une barre de chocolat, aussi appelée "pain au chocolat" », « Une boisson au chocolat chaud aromatisée à la vanille ». Les participants obtenaient à la fin un score de réussite et pouvaient répondre à un court questionnaire de données personnelles (localisation notamment). Ainsi, le quizz était constitué dans une idée de diffusion des résultats de la recherche auprès du grand public à travers une activité ludique, mais permettait aussi d'en tirer des informations pour la recherche, comme ici sur la diffusion du terme *chocolatine*, non pas nécessairement comme terme employé dans leur langue quotidienne (ce qui correspond plutôt aux enquêtes FRR), mais comme terme connu, et donc, d'une certaine manière, faisant tout de même partie de leur stock lexical :

⁷⁹ Thérapie Taxi est un groupe français, mais Roméo Elvis est effectivement belge. Voir le billet de Mathieu Avanzi sur la distribution des dénominations du téléphone portable en français d'Europe : <https://francaisdenosregions.com/2017/04/25/smartphone-natel-gsm/>

Pour les 1077 participants ayant renseigné leur origine, nous avons pu établir des cartes de diffusion des items selon les résultats aux questions¹⁵. On peut ainsi comparer l'aire d'utilisation de *chocolatine* (effectuée d'après l'origine des participants au quiz, sur données lissées, plus les réponses de la zone sont correctes, plus la couleur est foncée), par rapport à son aire d'emploi (effectuée d'après les résultats aux précédentes enquêtes, Avanzi et al. 2016) :

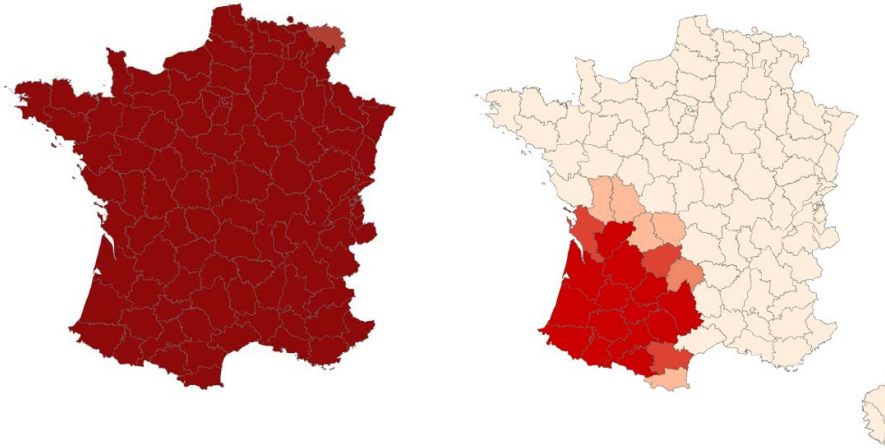


Fig. 1. Carte de réponses correctes au quiz vs carte d'emploi (enquêtes) de *chocolatine*.

(Glikman et al. 2018 [22] : 7-8)

Ils nous donnaient également accès aux représentations des locuteurs, y compris sur les représentations des variations diatopiques, comme dans l'exemple de la prononciation de 80. La question du quizz était « comment dit-on 80 en Belgique ? » et les réponses incluaient *quatre-vingts*, *huitante* et *octante*. La prononciation *quatre-vingts*, qui est effectivement la prononciation en usage en Belgique, a été dans la figure ci-dessous sélectionnée principalement par les Belges (qui connaissent leur usage), tandis que les Français ont eu tendance à privilégier les autres réponses. Cela montre que les Français ont connaissance de la différence de prononciation des nombres entre la France et la Belgique (ce qui est en effet le cas pour 70 et 90), mais n'en connaissent pas exactement le détail :

Enfin, dans le cas de la prononciation de 80, l'effet est inverse : en effet, les zones de réponses correctes sont principalement en Belgique francophone, alors que le terme effectivement utilisé en Belgique, en l'occurrence *quatre-vingt*, est utilisé dans l'ensemble de la France et de la Belgique :

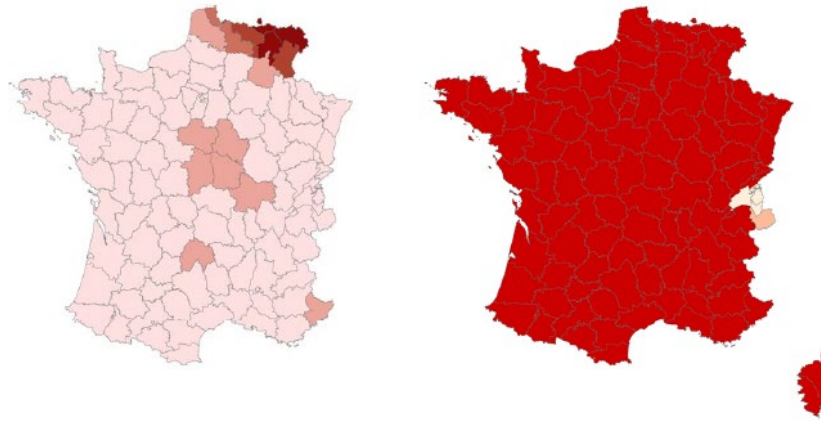


Fig. 6. Carte de réponses correctes au quiz vs carte d'emploi (enquêtes) de *quatre-vingt*.

Comme nous l'avons évoqué ci-dessus, ce résultat montre la représentation linguistique que les francophones de France ont sur le français de Belgique : ils sont au courant de la variation sur la dénomination des nombres, mais sans savoir l'appliquer dans son détail, et, par une sorte d'hypercorrection, ajoutent une spécificité régionale là où elle n'existe pas. (Glikman et al. 2018 [22] : 10-11)

Pour des phénomènes moins marqués diatopiquement, comme certaines variantes relevant plutôt de la grammaire que du lexique, j'ai pu observer un léger effet lié au pays, les participants français déclarant davantage utiliser les variantes acceptées par la norme (cf. Glikman 2019 [26] sur *malgré que* et *à cause que*, Glikman et Patard 2022 [35] sur *avoir été mis pour aller*). Les âges des participants ont également été analysés et ont donné quelques résultats, quoique de manière moins marquée pour les phénomènes que j'ai traités :

Lors des manipulations des données recueillies, mon collègue Mathieu Avanzi et moi avons en effet pu constater une corrélation : plus le participant est âgé et plus il a un diplôme élevé, moins il déclare utiliser la locution *malgré que*. (Glikman 2019 [26] : 90)

L'âge ne semble pas avoir d'effet particulier sur la répartition des variantes (Glikman et Patard 2022 [35] : 12)

Pour d'autres phénomènes, comme la diffusion de *full* étudiée par Jonathan Fontaine dans son mémoire de master (en cours), on peut s'attendre à ce que l'âge ait un impact plus significatif sur l'emploi du terme.

Nous avons également demandé les niveaux d'étude et la catégorie socio-professionnelle dans les enquêtes FRR, mais cela s'est révélé peu productif pour les phénomènes étudiés. Au regard de cette expérience, je n'ai pas demandé ces informations lors de la constitution du

corpus *Les Vocaux*⁸⁰. D'une part, je ne voulais pas décourager les participants par des formulaires fastidieux à remplir. D'autre part, je n'avais pas d'hypothèse a priori sur une éventuelle variation diastratique (et/ou diaphasique, selon la catégorie dans laquelle on considère que le niveau d'études et la catégorie socio-professionnelle renvoient). Comme je l'enseigne à nos étudiants dans le cours « Recueil et traitement des données » de 2^e année de Sciences du Langage, tout protocole de recherche doit être en lien avec les hypothèses explorées. Je laisse effectivement de côté les traditionnels axes diastratique et diaphasique, tant par probable habitude de diachronie, car ces axes sont plus difficilement observables et tendent, dans les états de langue anciens, à se donner à lire principalement à travers les genres et les représentations linguistiques, cf. Glessgen et Schøsler (2018 : 44) cités ci-dessus, que par doute sur leur efficacité, au moins à l'époque actuelle, comme le met en avant Andreas Dufter (2018) :

Mais les problèmes ne s'arrêtent pas là, car la diastratie et la diaphasie mériteraient aussi d'être passées en revue : il y a quarante ans déjà, André Martinet remarqua à propos de la diastratie dans le français hexagonal que « [b]eaucoup de traits langagiers socialement distinctifs se sont estompés ou ont disparu » (1979, 66). Wunderli (1992, 184) va dans le même sens lorsqu'il affirme au sujet de la diastratie qu'« il n'existe aucune variété définissable exclusivement par tel ou tel point donné sur cet axe [...] ». Sans doute, cette difficulté est à mettre en rapport avec celle de distinguer entre diastratie et diaphasie dans la description des données empiriques. Il est significatif dans ce contexte que dans Koch/Oesterreicher (2011), les phénomènes linguistiques relevant des deux dimensions sont traités en commun dans les chapitres consacrés au français, à l'italien et à l'espagnol modernes. Toutefois la distinction semble non seulement peu tranchée dans les travaux descriptifs, mais en plus, elle n'est pas non plus toujours établie de la même manière par les théoriciens de la variation (cf. Dufter/Stark 2003 pour une revue critique de différentes tentatives de délimitation). [...]

Reste la diaphasie, seule dimension disponible pour localiser les différents choix qu'un même locuteur peut faire dans différentes situations de communication [...]. Ici encore, il faudrait s'interroger dans quelle mesure l'ensemble des choix stylistiques se laisse ordonner sur une seule dimension. Un « axe diaphasique » reflète certainement les intuitions traditionnelles d'une échelle de degrés de formalité, mais il resterait à voir jusqu'à quel point une « sociolinguistique des locuteurs » (Gadet 2000) n'a pas besoin d'un équipement terminologique plus souple afin de permettre aux chercheurs une description plus nuancée de l'interface entre langue et contexte situationnel, en incluant aussi le rôle du destinataire. (Dufter 2018 : 70)

⁸⁰ La nouvelle loi sur la protection des données personnelles nous conduit en outre à limiter autant que possible les informations personnelles demandées et à en justifier la nécessité.

Outre les problèmes de définition théorique, le problème de rattachement des unités décrites se pose également, comme évoqué précédemment. Dans la citation extraite de la chanson *Hits Sales* donnée en exergue, on peut se demander à quel axe de variation rattacher des termes comme *seum*, *boloss*, *faire du vice*, quand on peut envisager à la fois un effet lié à l'âge, la classe ainsi que le groupe de pairs et quand, par ailleurs, on y trouve également des indicateurs clairement diatopiques, par exemple *GSM*. Ainsi, j'ai trouvé plus important de recueillir auprès des participants des informations sur leur usage et leur pratique de ce mode de communication encore peu connu, comme la fréquence d'emploi ou encore le type et le nombre de destinataires. Cela nous permet à la fois de mieux comprendre ce nouveau canal et la place qu'il occupe dans les pratiques des locuteurs. C'est aussi un moyen d'obtenir des indications sur le type de destinataire. Nous avons en effet demandé aux participants s'ils utilisaient plutôt les vocaux avec peu de personnes, intimes et toujours les mêmes, ou avec un cercle plus large, comme les collègues, voire tout le monde. Malheureusement, le type de recueil ne permet pas de connaître le destinataire de chaque message recueilli (→ 3.2.), mais nous avons au moins un aperçu des pratiques de chaque participant. J'ai également demandé l'âge (l'emploi des vocaux était jusqu'à récemment encore peu répandu et complètement inconnu de certains), la ville d'origine (nous avons vu avec Mathieu Avanzi dans les enquêtes FRR que ce paramètre était le plus influant sur les usages) et la ville de résidence actuelle.

Les difficultés que je viens d'évoquer à propos de la définition et de la délimitation des axes diastratique et diaphasique ne remettent évidemment pas en question l'intérêt pour la recherche des corpus autres que littéraires, en particulier les corpus de « peu lettrés ». Je pense notamment au Corpus 14, constitué de lettres échangées entre les « Poilus » et leur famille pendant la Première Guerre mondiale, développé à Montpellier, que je consulte de manière systématique dans mes études en diachronie large ou sur les variations contemporaines. Le développement de corpus de sources jusqu'ici peu exploitées, pour lesquels l'intérêt se fait croissant ces dernières années (voir les derniers colloques *Diachro* et *SIDF*), a certainement beaucoup à nous apprendre (voir entre autres Lagorgette et Larrivée (éds) 2013, Ayres-Bennett et al (éds) 2018 [v]).

Jusqu'ici, dans mes travaux, j'évoque ainsi assez peu les dimensions diastratique et diaphasique. Pourtant, ces dimensions, telles que les définissent Glessgen et Schøsler (2018) en termes de continuum, pourraient assez bien être intégrées à mon approche. En effet, partant de la problématique de caractériser les productions et non les locuteurs, ils proposent de transposer les axes diastratique et diaphasique en échelles, les énoncés s'inscrivant plutôt vers un pôle ou vers l'autre. Du point de vue des énoncés, ils considèrent que l'impact le plus important de la

dimension diastratique est l'action du prestige linguistique (Glessgen et Schøsler 2018 : 23) et le « degré de spécificité situationnelle ou contextuelle des énoncés » (Glessgen et Schøsler 2018 : 21) pour l'échelle relevant de la diaphasie :

Le continuum diaphasique se développe donc, dans notre idée, entre un pôle de haute spécificité voire un haut degré de contextualisation linguistiques et un pôle de faible spécificité linguistique. (Glessgen et Schøsler 2018 : 25)

La prise en compte du contexte et de la spécificité et, dans une moindre mesure, du prestige, évoquée ici au sujet de la diastratie et de la diaphasie, se retrouve en partie dans la définition des genres. La question des genres est évidemment le premier élément observable dans les corpus diachroniques. C'est un paramètre auquel je prête systématiquement attention, de fait, en travaillant sur corpus. Certains de mes travaux ont notamment contribué à décrire les spécificités de certains genres, participant ainsi à montrer la validité de ce paramètre (→ 4.).

J'emploie la notion de *genre* ici pour désigner toute production langagière, et non uniquement les traditionnels « genres textuels », dans la lignée de plusieurs travaux en linguistique variationnelle et intégrant la notion de traditions discursives⁸¹. Dans cette conception, une conversation entre amis est un *genre*, tout comme un entretien d'embauche, un éditorial de journal de presse, une dissertation de philosophie au baccalauréat, ou un article de loi. Dans cette optique, toute production langagière s'inscrit dans une tradition, à un moment historique donné, dans un contexte donné, à destination d'un public donné et en présente les caractéristiques liées.

À la question des genres, qui, ainsi, ne concernent pas que les *textes*, mais bien toute production langagière, se trouvent donc étroitement liées les questions de visée, de portée, de degré de préparation, de destinataire, d'intention et de programmation⁸². Ces paramètres rejoignent les paramètres descriptifs du continuum distance communicative / immédiat communicatif tel qu'il a été conçu par Koch et Oesterreicher (2001), qui est le modèle que j'utilise majoritairement dans mes travaux. En effet, un autre paramètre que j'ai exploré jusqu'ici, outre les genres, est également la question des rapports oral / écrit, en tant que médium, de l'oral représenté, en particulier en diachronie du français, ainsi que, d'une manière plus générale, la variation selon ce que Koch et Oesterreicher ont clairement distingué du

⁸¹ Cf. Winter-Froemel 2020 pour une synthèse sur la notion de traditions discursives.

⁸² L'importance de cette notion de *programmation*, que j'avais jusqu'alors peu rencontrée, a été à juste titre rappelée par Jeanne-Marie Debaisieux lors d'une intervention au colloque en hommage à Claire Blanche-Benveniste à l'ATILF (Nancy) en octobre 2021.

médium en proposant le continuum conceptuel distance / proximité communicative. Koch et Oesterreicher ne sont pas les premiers à s'interroger sur cette variation qu'on désigne parfois sous le nom de variation diamésique, bien qu'elle ne concerne parfois sous ce nom que l'opposition médiale⁸³ et non conceptuelle⁸⁴. Il me semble cependant que leur approche présente l'avantage d'associer leur théorisation à une liste de critères, non finie, permettant de classer les productions de manière concrète, alliant ainsi théorie et pratique. En outre, les critères qu'ils proposent, comme le destinataire, le degré de préparation, etc. me semblent véritablement des paramètres fondamentaux dans l'évaluation des productions. Dans le cours que nous avons créé en 1^{re} année de Sciences du Langage, « Langue, code, communication », dans lequel j'enseigne le module sur « oral / écrit », je sensibilise les étudiants à ces questions. Après leur avoir montré les préjugés sur l'oral (qu'on retrouve notamment dans les manuels scolaires), souvent confondu avec le registre familier, j'amène les étudiants à cette distinction entre le médium et le degré de distance / proximité communicative. Auprès de ces étudiants de 1^{re} année, la notion du continuum conceptuel est parfois difficile à saisir, et tend à être facilement confondu avec les registres. De fait, certains aspects de définition de registre se rapprochent des critères servant à caractériser l'opposition distance/proximité, comme l'adaptation au destinataire. Une fois cette distinction mise en place, un autre aspect essentiel mais pourtant difficile à leur transmettre réside dans les spécificités qui restent liées au médium lui-même. Comme le rappelle Christophe Benzitoun⁸⁵, les codes oral et écrit, par essence, n'ont pas le même fonctionnement et vont donc avoir des caractéristiques et des réalisations qui leur sont propres, comme le marquage morphologique du pluriel (<les murs> vs [lemyR]). Ainsi, malgré la distinction à faire entre le médium et le degré de proximité, l'un comme l'autre jouent un rôle dans la caractérisation et les spécificités des productions.

Là encore, mon projet actuel sur le corpus *Les Vocaux* est dans la continuité de mes travaux. Les SMS vocaux relèvent en effet de l'oral du point de vue du médium. En tant que SMS, ils relèvent plutôt de la proximité communicative, en tant que message privés, destiné à un destinataire connu. On peut également s'attendre à ce que le degré de préparation soit faible, à

⁸³ C'est-à-dire correspondant à la matérialité du support, au code, graphique ou phonique.

⁸⁴ Correspondant à la conceptualisation, « parlé » ou « oralité », au regard de « écrit » ou « scripturalité », redéfinie en continuum distance ou proximité communicative chez Koch et Oesterreicher.

⁸⁵ Avec qui j'ai beaucoup discuté lors de la préparation de ce cours, en tant que spécialiste de l'oral, formé à Aix auprès de Claire Blanche-Benveniste et José Deulofeu. Je le remercie encore de m'avoir confié certains de ses supports pour monter ce cours.

la différence de vidéos YouTube ou TikTok, souvent issues de montage et publiques. En outre, les premiers résultats ont montré la présence de différents types de discours dans les vocaux. On y trouve en effet des énoncés présentant des marques de l'interlocution : adresses directes à l'interlocuteur à la deuxième personne (50), questions directes (51), mais également des messages de type narratif, comprenant des passages de narration et d'oral représenté (52) (les ex. 50 à 52 sont tirés de Glikman et Fauth (2022 [34]) :

(50) je t'envoie ce petit message (01_03)

(51) et qu'est-ce qui t'a déclenché cette pensée ... explique-moi (06_03)

(52) et il dit bonjour alors on est là genre bonjour ... et euh ... il sort à #name euh ... ouais euh vous pensez à bien garder les distances de sécurité hein alors elle dit ben oui vous voyez bien qu'ils sont espacés ... oui oui non mais je vois je vois hein (05_08)

Ce corpus permettra ainsi de contribuer à la description de ces différents discours, dont l'oral représenté, dans ce nouveau mode de communication.

3.1.3. Bilan intermédiaire

Ainsi, j'exploite dans mes travaux les dimensions des variations diachronique et, dans une moindre mesure, diatopique, des genres, des rapports oral / écrit et de la représentation de l'oral ainsi que l'opposition distance et proximité communicative, les représentations et la norme, et leur évolution. En particulier, mes travaux récents me conduisent à un intérêt croissant pour l'étude du changement dans le statut des variantes, à la fois dans les productions et dans les représentations (→ 4). Même si, jusqu'ici, j'ai peu eu recours aux notions de diastratie et diaphasie, il me semble évident qu'il existe des variations inter- et intra- individuelles. L'adaptation au contexte (y compris au destinataire) et l'intentionnalité du message (incluant la programmation), constituent des points primordiaux à prendre en considération, quel que soit le nom sous lequel on les classe.

À travers ces questions, on comprendra que mon approche est plutôt centrée sur le locuteur, au sens de la linguistique variationnelle, et son utilisation des possibilités du système (Verjans 2014) (→ 4.), tout en incluant les notions de représentations, de groupes de prestige, d'hypercorrection, etc. Cela me permet de travailler à la fois en diachronie et sur les variations contemporaines, avec une cohérence dans les approches et dans les méthodes utilisées. En effet, quels que soient l'époque considérée et le type de production, j'utilise autant que possible les mêmes outils conceptuels et descriptifs ainsi que la même méthode d'analyse. J'étudie les réalisations, tout en essayant de ne pas plaquer de catégorie a priori. C'est ainsi que j'essaie

d'amener mes étudiants à la recherche. Pour l'année 2022-2023, j'ai accepté de suivre une étudiante qui souhaite faire son mémoire de Master Sciences du Langage sous ma direction sur « la langue des jeunes ». Pour moi, comme je le lui ai expliqué, la « langue des jeunes » n'existe pas (cf. aussi Dufter 2018 : 70), ce qui n'empêche pas d'étudier les variations selon l'âge, le contexte de production, l'intention communicative, etc. ainsi que les représentations et les jugements des locuteurs, auxquels j'accorde de plus en plus d'importance dans mes travaux récents.

Une question qui se pose alors est d'accéder aux moyens d'étudier la variation, point que je vais aborder dans la section suivante (→ 3.2.). Le développement des corpus facilite aujourd'hui grandement la recherche et le recours aux enquêtes peut efficacement compléter les données quand les corpus sont insuffisants. Les enquêtes FRR ont en outre permis d'accéder non seulement aux réalisations, même s'il ne s'agit pas tant de production que de la conscience linguistique du locuteur, mais aussi au jugement des participants sur les variantes, point sur lequel nous reviendrons au chapitre 4.

3.2. Quelles données pour analyser la variation ?

*Sean Paul is the only human in
history that speaks doctor's
handwriting* 🤖🤖🤖

Commentaire public de Mothebro
Bro, 2021⁸⁶

Après avoir évoqué dans la section précédente les orientations théoriques que j'adopte pour analyser la variation, je vais présenter maintenant les données concrètes que j'ai jusqu'ici utilisées. Cela sera ainsi l'occasion d'aborder de manière plus concrète les questions de méthode ainsi que celles de l'accès aux données, de leur enrichissement et de leur mise à disposition pour la communauté.

J'ai eu la chance à la fois de commencer mes recherches par une approche exhaustive du *texte* comme corpus et d'être également très tôt familiarisée avec les méthodes et outils sur grands corpus. J'ai en effet commencé mes premières recherches par le dépouillement manuel d'œuvres, rendu notamment nécessaire par mes études sur la parataxe, recherche de l'absent.

⁸⁶ Commentaire public de Mothebro Bro sur la video YouTube Sean Paul - *Give It Up To Me* (feat. Keyshia Cole) [Disney Version] (Official Video) daté « il y a 9 mois » au moment de la consultation (consulté le 04.11.21) <https://www.youtube.com/watch?v=bMoNV9oWs2E&list=RDMmJqEioMMJAU&index=27>.

Cette méthode permet d'approfondir une réelle connaissance des textes et des particularités linguistiques de chacun. C'est un phénomène bien connu que les œuvres du moyen âge varient, tant par leurs traits diatopiquement marqués que par leur forme (vers / prose), leur insertion dans un genre, leur datation... autant de paramètres qui peuvent influencer sur leur langue. Malgré les points communs qui permettent tout de même aux linguistes de parler d'« ancien français » ou de « moyen français », il me semble important de se familiariser avec les particularités de chaque texte et d'observer les occurrences en contexte. Les phénomènes étudiés doivent absolument être replacés en contexte et leur fonctionnement étudié dans le cadre du texte en tant que système. En effet, malgré les variantes inhérentes aux textes médiévaux, leurs réécritures et leur mouvance (Zumthor 1981, Cerquiglini 1989), voire dans certains cas la pluralité des voix qui les composent, le texte constitue à mon sens un premier pas vers une compréhension d'un fonctionnement en système, avant de pouvoir éventuellement en étendre les conclusions. Ainsi, il me semble important de regarder les textes eux-mêmes, tels qu'ils se donnent comme objet, et non uniquement à travers les bases de données. Les bases de données donnent accès à un grand nombre d'occurrences directement comparables en contexte, ce qui est primordial et constitue un apport indéniable. Cependant, pour certains phénomènes, la connaissance du texte même est absolument nécessaire, de même que la place du phénomène au sein de la langue de celui-ci (le texte⁸⁷ étant alors considéré en tant que témoin d'un état de langue, cf. Verjans 2012). Quand une seule occurrence d'un fait linguistique est observée, nous pouvons interroger sa place dans la langue du texte : est-ce parce que c'est la seule occurrence de la variable, ou parce que les autres occurrences du même texte présentent l'autre variante ? Pour donner un exemple concret, si l'on considère la variante *de façon à ce que* (*malgré que*, *avoir été* + infinitif...), est-ce la seule expression de cette variable dans le texte, ou est-ce la seule occurrence de cette variante, alors que le texte contient d'autres réalisations en contexte équivalentes sous la forme *de façon que* ? Les conclusions à en tirer sur la répartition des variantes ne seront ainsi pas les mêmes selon le cas. C'est ainsi que, dans plusieurs travaux (Glikman 2009b [5], 2010b [7], 2012 [8], Mazziotta et al. 2012 [9]), j'ai travaillé sur un nombre plus restreint de textes⁸⁸, ce qui m'a permis d'observer le « comportement en système » :

⁸⁷ Texte toujours pris ici en tant que production langagière qui se donne comme une unité (plus ou moins) cohérente (voir note 47).

⁸⁸ Dans le même ordre d'idées, je regarde si les différentes variantes se trouvent chez le même auteur (ou scripteur, dans le cas des correspondances du *Corpus 14*) ou non, sous quelle répartition, etc.

L'observation du comportement de nos constructions verbales, dans notre corpus, par rapport aux personnes et aux temps compatibles selon leur position, nous apporte un indice supplémentaire qui vient confirmer notre analyse d'incise vs rection par correspondance entre rattachement prosodique et analyse syntaxique.

En effet, dans les configurations de rattachement à gauche, la forme verbale apparaît à la première personne du présent, qui est l'unique personne et le seul temps auxquels elle peut apparaître lorsqu'elle est en position incise, le blocage sur la première personne du présent étant caractéristique des incisives avec ces verbes (Blanche-Benveniste, 1989).

Dans les configurations de rattachement à droite, au contraire, on trouve les formes verbales en question à différentes personnes et à différents temps, comme dans l'exemple [16] ci-dessous, de même qu'à différentes modalités, comme on avait en [12] et [13]. Cette absence de contrainte est caractéristique d'un emploi libre, tel qu'il apparaît quand le verbe est en position recteur (Glikman 2010b [7] : 37)

La question des éditions modernes et de leur mode de constitution doit évidemment être elle aussi prise en compte. En particulier, sur les périodes plus récentes (pratiquement dès la période post-médiévale), la modernisation de l'orthographe est quasi systématique⁸⁹. Dès que possible, avec mes étudiants de Licence en cours de français pré-classique et classique, j'ai recours aux œuvres originales, dont la consultation est maintenant largement facilitée grâce au site Gallica de la BNF⁹⁰. L'analyse des variantes (au sens « philologique » ici), auxquelles nous pouvons parfois avoir accès grâce à des éditions synoptiques quand elles existent, ainsi que les mises en prose, constituent aussi des sources prometteuses pour l'étude de la variation. Bien que j'en aie exploré quelques-unes dans ma recherche doctorale, sur les excellents conseils d'Annie Bertin, ma directrice de thèse, ce sont des sources que j'ai trop peu exploitées par la suite. De fait, je n'ai jusqu'ici pratiquement jamais exploité les manuscrits eux-mêmes, travaillant la plupart du temps sur des éditions modernes, à l'exception, comme je viens de le mentionner, des extraits d'œuvres de la Renaissance pour mes cours. Mais là encore, j'utilise peu ces œuvres originales pour les études de diachronie large, où je me sers plutôt des ressources électroniques existantes. Cependant, j'initie toujours mes étudiants en premier lieu à la réalité du texte médiéval et à la distance entre le texte édité et le texte du (ou des) manuscrit(s), tant dans le travail de l'éditeur

⁸⁹ Il suffit de regarder quelques introductions aux éditions de la collection XVII^e siècle chez Classiques Garnier pour en avoir une idée. Pour la Renaissance ce n'est peut-être encore pas le cas, par contre certains de ces textes s'accompagnent de traductions. Pour le Moyen Âge, la distance avec la langue d'aujourd'hui préserve en quelque sorte les textes de cette modernisation, mais la question des principes d'édition ne s'en pose que davantage (voir Glikman et Verjans 2021 (éds) [ix]). Il semblerait que plus le texte est proche de nous, moins on le traduit, mais plus on le modernise. La question de la traduction intralinguale, y compris pour des périodes plus récentes, est d'ailleurs également un champ d'étude intéressant pour l'étude de la variation que j'aimerais à l'avenir davantage explorer (cf. Karas et Shyldkrot 2021).

⁹⁰ <https://gallica.bnf.fr/>.

moderne que dans la question de la circulation des textes et de la place (des pratiques) de l'écrit (et de la lecture) à l'époque. Pour cela, j'utilise l'édition numérique de la *Queste del saint Graal* de Christiane Marchello-Nizia et Alexei Lavrentiev⁹¹, qui permet d'afficher côte à côte la photo du manuscrit, l'édition diplomatique, l'édition dite normalisée et la traduction. Les possibilités de l'édition électronique nous ouvrent de nouvelles perspectives, dans l'affichage synoptique, l'intégration des variantes, les choix de visualisation, qui peuvent encore beaucoup nous apporter.

Comme je viens de l'évoquer, j'ai également eu la chance d'être très tôt familiarisée avec l'exploitation de grand corpus diachroniques, et d'être intégrée à des équipes qui travaillent à leur constitution et à leur enrichissement. Dès ma thèse, j'ai participé aux réunions du groupe de travail d'Annie Bertin sur les locutions conjonctives, dont l'un des objectifs était de rédiger des entrées du *Dictionnaire de Moyen Français* pour les mots grammaticaux. Nous avons ainsi eu plusieurs échanges avec l'équipe du DMF à l'ATILF (Nancy), en particulier avec Gilles Souvay qui nous a présenté l'architecture d'un article du DMF et les outils du rédacteur. À l'époque, je n'ai finalement apporté aucune contribution au DMF, mais j'ai toujours gardé d'étroits contacts avec Nancy et suivi leurs travaux, notamment le développement des ressources électroniques comme leur lemmatiseur LGeRM (objet de l'article *Lemmatisation et construction automatique de ressources lexicographiques : les développements du lemmatiseur LGeRM* de Sylvie Bazin-Tacchella et Gilles Souvay dans Glikman et Verjans (dir.) (2021) [ix]). J'ai ainsi mis en contact avec Gilles Souvay une étudiante de Master qui faisait son mémoire sur la lemmatisation et l'annotation des verbes sur la *Queste du Graal* sous ma codirection (Clara Bringer, Master technologie des langues, 2017-2019).

J'ai également eu très tôt la chance de travailler avec plusieurs collègues œuvrant à l'élaboration et l'enrichissement de corpus, dont l'équipe de la *Base de Français Médiéval* (désormais BFM). Lors de mes différentes participations à des projets de recherche, de mon post-doctorat dans le projet SRCMF au projet *Les Vocaux*, sans oublier le projet *Democrat* (→ 1.), j'ai également participé à la constitution et à l'enrichissement des corpus. Dans SRCMF, il s'agissait d'annotation syntaxique de textes provenant de la BFM et du *Nouveau Corpus d'Amsterdam*. Le projet *Democrat* (et avant lui, le projet MC4) consistait en l'annotation de textes (dont des textes de la BFM) en chaînes de référence. Dans les deux cas, il s'agissait d'annotation manuelle avec des annotateurs humains. J'ai également participé à la correction

⁹¹ <https://txm-bfm.huma-num.fr/txm/>.

des annotations automatiques en partie du discours de textes de la BFM⁹². J'ai ainsi pu me former aux outils d'annotation et de requêtes, aux problématiques de la constitution de schéma d'annotation, en un mot, à la linguistique sur (grand) corpus. Le développement de la linguistique informatique et, en particulier, des annotations entraînées par des réseaux de neurones, change considérablement les pratiques d'annotation et d'autres évolutions sont à attendre. Aujourd'hui, la plupart des annotations automatiques conservent une part d'erreur à corriger, mais à terme, on peut s'attendre à ce qu'une annotation automatique soit plus cohérente qu'un annotateur humain, dont le degré de vigilance n'est jamais constant. Dans le cadre du projet *Les Vocaux*, nous utilisons ainsi plusieurs outils qui bénéficient de ces avancées, depuis la transcription orthographique automatique⁹³ à l'annotation automatique en partie du discours, lemme et annotation syntaxique⁹⁴. L'idée est de conserver l'alignement avec le son, mais aussi d'aligner la transcription orthographique avec une transcription phonétique. Cela permet à la fois d'opter pour une transcription en orthographe « standardisée », type <tu sais>, qui a l'avantage de faciliter le travail des analyseurs et les requêtes sur les formes, et de conserver l'information de la donnée telle qu'elle a été produite, type [tse]. En effet, si l'on commence à adapter la transcription orthographique à la prononciation, jusqu'à quel point doit-on aller ? Si l'on est habitué à voir se développer (notamment dans des représentations de l'oral) des formes telles que <j'sais pas>, on aura davantage de difficulté à penser à rechercher une forme telle que <pasque> voire <paske> pour <parce que>. Nous retrouvons là des questions que le philologue et l'éditeur de textes médiévaux connaissent bien : la question de la résolution des abréviations. Chose à laquelle je m'attendais moins, travaillant jusqu'ici sur l'écrit, mais que j'ai découverte lors de la préparation de l'édition du corpus *les Vocaux*, la question se pose de la même manière pour la transcription phonétique. Jusqu'à quel point transcrire les prononciations « atypiques » comme *bidon* ou *bretzel* prononcés [bitõ] ou [pretzel] en Alsace ? Là encore, la question de l'exploitation ultérieure du corpus est à prendre en considération, en particulier les questions de recherche qui sont visées par la constitution du corpus. Dans

⁹² *Les Manières de langage* (1396, 1399 et 1415), Philippe de Beaumanoir, *Coutumes de Beauvaisis (volume 1) et L'ornement des dames*.

⁹³ Nous utilisons actuellement la technologie Vocapia via l'éditeur en ligne YobiYoba, grâce à une collaboration avec le LIMSI, en particulier Lori Lamel. Nous sommes également en cours de test pour l'alignement en phones avec le LIMSI (avec Lori Lamel et Philippe Boula de Mareuil). Nous utilisons jusqu'ici pour l'alignement le script EasyAlign développé pour Praat par Jean-Philippe Goldman.

⁹⁴ Pour cet aspect, nous collaborons avec Loïc Grobol (Université Paris Nanterre) avec qui nous avons été mis en contact grâce à Sylvain Kahane. Plusieurs analyseurs sont actuellement en cours de test.

l'optique d'un partage avec la communauté et d'une réutilisation du corpus par d'autres, nous nous dirigeons vers le choix d'une transcription « standard » au marquage le plus neutre possible. Là encore, nous espérons que les informations perdues (ce qui n'est pas transcrit ni annoté est définitivement perdu, cf. Mazziotta 2016) seront en partie au moins récupérables par la superposition des couches d'annotation et l'alignement avec le fichier sonore. Et même en ce qui concerne le fichier sonore, les discussions qui ont suivi ma présentation du projet au congrès des franco-romanistes allemands à Vienne en septembre 2022 m'ont permis de réfléchir à ma procédure d'anonymisation. Pour le respect de l'anonymat des participants, je brouille en effet tout nom et autre élément reconnaissable. Cependant, en faisant ainsi, je perds le peu d'informations disponibles sur l'interlocuteur, en particulier le fait que plusieurs messages recueillis peuvent avoir le même interlocuteur. Suite à ces discussions, j'ai choisi d'adapter ma procédure d'anonymisation et d'intégrer à la transcription orthographique un identifiant pour l'interlocuteur. De même, lors de la phase de transcription orthographique, j'ai dû effectuer des recherches pour reconnaître la mention d'un groupe de K-Pop que je ne connaissais pas, et qui par conséquent a demandé un travail de recherche externe et un appui sur le co-texte⁹⁵, ou encore à un terme technique : un vocal mentionnait notamment les *cellules gliales*, terme peu courant qu'il faut pouvoir retrouver. Dans certains cas, nous avons ajouté une mention « inconnu » pour les termes que nous n'avons pas réussi à identifier.

Les questions qui se posent lors de la constitution de ces corpus et l'établissement des choix d'annotations sont ainsi de vraies questions philologiques, que l'on retrouve quel que soit l'état de langue, comme Nicolas Mazziotta l'a bien montré dans son mémoire de synthèse d'habilitation (2016). Tout choix d'édition et toute annotation est déjà en soi une analyse, qu'il importe de bien documenter, à la fois pour faciliter le travail des annotateurs et l'homogénéité du corpus et, peut-être surtout, l'exploitation ultérieure du corpus par la communauté. C'est la raison pour laquelle nous avons attaché une attention particulière à la rédaction du guide d'annotation SRCMF, qui documente nos choix et permet de retrouver les types de structures incluses dans les étiquettes (Bischof et al. 2012 [10], consultation en ligne uniquement). Le travail de conception des fiches, au format HTML incluant des renvois internes, réalisé par Nicolas Mazziotta, permet une navigation aisée d'une fiche à l'autre et une recherche inversée.

⁹⁵ Le vocal seul ne pouvait pas même permettre de comprendre qu'il s'agissait d'un groupe de K-Pop : « parce que enfin en ce moment je suis grave ... à fond dans ENHYPEN en plus ils sont bien trop com back ... et Given Taken c' est une musique que je j' aurais bien voulu euh ... euh apprendre la choré » (*Les Vocaux*, 07_01), mais la même locutrice évoquait la K-Pop dans un autre message, ce qui a permis d'orienter les recherches.

Là encore, les supports électroniques contribuent à structurer la pensée et à atteindre des visualisations qui n'étaient pas possibles avant. De même, pour toutes les questions de choix d'édition qui se sont posées dans le cadre de la constitution du corpus *Les Vocaux* et dans la mise au point de la chaîne de traitement et du protocole, nous avons travaillé d'abord sur un corpus martyr, préalable nécessaire pour la mise au point des différentes étapes.

Les choix effectués lors de la constitution et de l'annotation d'un corpus (comme, jusqu'à un certain point, lors de l'édition d'un texte médiéval) sont ainsi toujours liés à la question des objectifs visés et à la question de la double lecture : quel public vise-t-on et quelle lecture veut-on donner du corpus. C'est le même type de questions que doit résoudre l'éditeur de textes médiévaux pour faciliter l'accès au texte par le lecteur moderne, sans pouvoir toujours conserver la diversité d'informations utiles à la recherche en linguistique (cf. Glikman et Verjans (dir.) 2021 [ix]). Les outils modernes devraient permettre de multiplier les couches d'informations, les requêtes croisées, tout en permettant de choisir sa visualisation. Il me semble important d'insister sur le fait que les choix seront différents selon les objectifs et qu'il serait vain de faire comme s'il pouvait en être autrement.

Dans tous les cas, il me semble primordial que le chercheur « mette les mains à la pâte ». Il est évidemment important d'obtenir des financements pour des stages et des annotateurs / transcrip-teurs, à la fois pour financer de jeunes chercheurs et pour la formation à la recherche. Cependant, faire soi-même au moins une partie des différentes étapes est le seul moyen de se rendre compte des problèmes effectivement rencontrés lors de la constitution de corpus et de l'annotation, et ainsi de réfléchir aux choix à faire. Il ne faut pas croire qu'il s'agit d'un travail trivial. C'est au contraire un travail d'expert et on ne le dira jamais assez : l'annotation (y compris transcription et édition du corpus), c'est l'analyse. La création de corpus moderne comme ancien est une question de choix, en particulier pour l'annotation, mais pas seulement : construire un corpus, c'est déjà analyser (Mazziotta 2016).

Ces problématiques sont valables quel que soit l'état de langue. Outre les questionnements théoriques, leur application « pratique » est également transversale. En effet, dans la constitution de corpus annoté en syntaxe, la question du découpage en unité maximale se pose d'autant plus que des choix sont nécessaires à faire. En effet, lors de la constitution des deux corpus annotés en syntaxe pour le français oral (*Rhapsodie*⁹⁶, dont j'ai suivi de près l'évolution, et *Les Vocaux*) et ancien français (SRCMF, auquel j'ai contribué), le problème pratique de

⁹⁶ <https://rhapsodie.modyco.fr/> (dernière consultation le 07.12.22).

délimitation de la « phrase » est en substance le même (cf. Benzitoun et al. 2010, Lavrentiev 2010). Ces choix doivent en outre à la fois permettre, par la clarté de leur directive, la plus grande cohérence possible inter-annotateur (voire intra-annotateur...), donc la reproductibilité, et la traçabilité des phénomènes et la réutilisation des données. Pour l'exploitation du corpus, les modalités de son interrogation doivent également parfois être prises en compte. Ces différents paramètres peuvent conduire à faire des choix pratiques, quitte à risquer de s'écarter de son postulat théorique de départ. Là encore, il me semble que la qualité de la documentation doit servir de support. Le développement des outils pourrait aussi, à l'avenir, faciliter certains de ces choix. Notamment, avec Nicolas Mazziotta, nous discutons parfois de la possibilité d'ajouter à l'annotation du corpus *Les Vocaux* une annotation « ambigüe », c'est-à-dire qui conserverait les deux possibilités de lecture. De la même manière, j'avais suggéré à Clara Bringer de conserver les ambiguïtés dans son annotation en temps et mode des verbes de l'édition électronique de *la Queste del saint Graal* (mémoire de Master Technologie des Langues, Bringer 2019).

La taille du corpus peut aussi varier à la fois selon ce qu'on veut en dire, comme je l'ai montré sur l'intérêt d'une étude en système « clos » tel que le texte, et les données à disposition. Actuellement, les grands corpus connaissent un essor considérable, y compris en termes d'enrichissement : par exemple, aussi bien *Frantext* que la BFM comportent des informations sur les mots telles que l'annotation en partie du discours et la lemmatisation. Ces enrichissements facilitent grandement les requêtes ciblées et permettent de limiter le « bruit ». Les enrichissements supplémentaires de la BFM, comme le marquage du discours direct, permettent également des études fructueuses. C'est ainsi que nous avons pu interroger les spécificités syntaxiques du discours direct et de la narration dans Glikman et Mazziotta (2013) [12] et Mazziotta et Glikman (2019) [27]. De même, nos travaux sur l'étude des chaînes de référence portaient sur des sélections plus restreintes de textes (Glikman et al. 2014 [14], Obry et al. (2017) [20], Oberlé et al. (2018) [25], Capin et al. 2021 [33]). Cela est dû à la fois à l'annotation manuelle des chaînes de référence, plus coûteuse en temps, et à la caractéristique de ce type d'étude, qui demande une connaissance fine des textes (personnages, situations, épisodes...).

Bien que les choix opérés dans la constitution et l'annotation d'un corpus puissent toujours être discutés, par exemple pour la BFM, concernant les textes comme les éditions⁹⁷, il me semble d'un intérêt évident de combiner les énergies et les efforts pour continuer l'enrichissement de ces textes. C'est ce qui a été fait dans le cadre de SRCMF et Democrat, dont l'annotation portait sur certains textes de la BFM. De même, j'ai proposé d'intégrer à la BFM le texte *l'Ornement des dames* numérisé par Eudeline Lecluse dans le cadre de son mémoire de master sous ma direction.

C'est aussi pour participer aux efforts visant à l'ouverture de la recherche et à l'interopérabilité que je participe (en tant que représentante de mon laboratoire, LiLPa) au consortium Corli⁹⁸. C'est dans cette optique que nous souhaitons développer le corpus *Les Vocaux*, que nous voulons distribuer sous différents formats. Nous travaillons étroitement avec Cyril Pestel de la plateforme Ortolang⁹⁹ pour la distribution du corpus. Le développement pour une science ouverte passe aussi par la libre distribution des produits de la recherche, tant donc les corpus, que les articles. Je participe ainsi activement aux comités de rédaction de deux revues en accès libre, *Linx*¹⁰⁰ et *Studia Linguistica Romanica*¹⁰¹. C'est aussi dans l'idée du partage et de la mise à disposition que je contribue à la diffusion des informations à travers une veille scientifique faite dans le cadre de mon activité dans la *Société Internationale de Diachronie du Français* (SIDF)¹⁰².

Les corpus constituent donc aujourd'hui une étape nécessaire pour toute recherche diachronique et variationnelle d'une manière générale. Je distingue les ressources qui permettent l'étude des productions et celles qui visent à l'étude des discours sur la langue et des représentations. Pour l'étude des productions, j'utilise ainsi des corpus tels que *Frantext*, la BFM et, notamment, pour des études en diachronie large, le corpus de la GGHF, constitué d'une sélection de textes de *Frantext* et de la BFM visant à une certaine représentativité et ayant reçu un traitement homogénéisé pour faciliter les recherches. Ce corpus comporte en outre une partie

⁹⁷ L'équipe de la BFM travaille cependant en ce sens, tant dans l'ouverture vers des textes non littéraires, que vers l'établissement d'éditions électroniques synoptiques (voir l'édition de la *Queste del saint Graal* de Christiane Marchello-Nizia et Alexey Lavrentiev <https://txm-bfm.huma-num.fr/txm/>).

⁹⁸ Consortium HN CORpus, Langues et Interactions <https://corli.huma-num.fr/>.

⁹⁹ <https://www.ortolang.fr/fr/accueil/>.

¹⁰⁰ <https://journals.openedition.org/linx/>.

¹⁰¹ <https://studialinguisticaromanica.org/>.

¹⁰² <https://diachronie.org/>.

échantillonnée visant à un équilibre par siècles pour des approches plus quantitatives. J'exploite également les corpus que j'ai contribué à enrichir. Le corpus SRCMF est composé de textes déjà présents dans des bases de données BFM et NCA que nous avons enrichi pour l'analyse syntaxique de l'ancien français. Le corpus DEMOCRAT, annoté en chaînes de référence, est en partie composé de textes de la BFM pour la partie médiévale et d'autres bases textuelles, comme *Frantext*, mais aussi d'autres textes, y compris des textes non littéraires comme des textes juridiques et des articles de l'encyclopédie en ligne Wikipédia. Le fait d'avoir participé à la constitution de ces corpus me permet évidemment d'avoir une connaissance fine de leur composition et de leur mode de constitution. De même, ayant contribué à la vérification manuelle de l'étiquetage CATTEX de certains textes de la BFM, j'ai appris à maîtriser le jeu d'étiquettes et les choix d'annotation. Connaître la composition d'un corpus est fondamental pour l'utiliser, en tout cas pour l'utiliser de manière efficace et raisonnée, d'où l'importance de la documentation et explicitation des choix, comme nous l'avons mentionné ci-dessus. Cependant, encourager l'utilisation d'un corpus annoté, en dehors des membres qui y ont contribué est parfois difficile et peut demander un effort de formation auprès de la communauté. La question des outils d'exploitation des corpus peut aussi poser problème. Il me semble également important de diffuser librement le corpus et sous différents formats. Pour *les Vocaux*, nous envisageons ainsi à la fois une distribution sous forme d'un corpus directement exploitable sous TXM, pour les personnes connaissant déjà le fonctionnement de TXM, et au format brut XML ainsi qu'au format CoNLL, pour des utilisateurs préférant utiliser d'autres outils.

Pour les écrits non littéraires, j'utilise le corpus SERMO¹⁰³, le *Corpus 14*¹⁰⁴, les corpus de SMS¹⁰⁵ et certains corpus oraux, en particulier ceux regroupés dans le cadre du projet Orfeo¹⁰⁶. Les corpus French Mixed 2012 de la collection de corpora de Leipzig¹⁰⁷ et Europresse¹⁰⁸ permettent également de faire des exploitations pour les périodes contemporaines. Ils sont composés d'articles de presse et de sites Internet. Europresse est davantage conçu à l'origine pour des exploitations non linguistiques mais peut tout de même être utilisé pour des relevés.

¹⁰³ <http://sermo.unine.ch/>.

¹⁰⁴ <https://www.univ-montp3.fr/corpus14/>.

¹⁰⁵ <http://88milsms.huma-num.fr/corpus.html> et <https://www.whatsup-switzerland.ch/index.php/fr/>.

¹⁰⁶ <https://repository.ortolang.fr/api/content/cefc-orfeo/4/documentation/site-orfeo/corpus-source/index.html>.

¹⁰⁷ https://corpora.uni-leipzig.de/?corpusId=fra_mixed_2012. Le portail présente d'autres corpus français disponibles mais j'ai jusqu'ici utilisé celui-ci.

¹⁰⁸ <https://nouveau.europresse.com/Search/Reading>.

Le corpus de Leipzig est au contraire déjà constitué pour des analyses linguistiques et produit automatiquement quelques calculs de mesures significatives, notamment les fréquences, les cooccurrences significatives (également représentées sous forme de graphe), les voisins de droite et de gauche les plus significatifs, et les mots ayant un contexte similaire. L'utilisation de ce corpus nous a ainsi permis de mettre au jour le fait que, dans ce corpus, *maint* est le mot apparaissant le plus en contexte similaire à *moult*, ce dont nous n'avions pas l'intuition au préalable et qui ouvre une nouvelle perspective de recherche (Glikman et Verjans 2021 [32]). Ces différents corpus combinés nous donnent un aperçu de différents contextes de production, même si les questions de représentativité se posent toujours. Le développement du corpus des *Vocaux* vise à contribuer à l'étude de ces pratiques communicatives nouvelles. J'espère également qu'il permettra d'avoir accès à des productions de type spontanées ou en contexte non surveillé. Comme me l'a fait remarquer mon collègue Vincent Balnat, on peut émettre des doutes quant au degré de spontanéité de ces messages, pour lesquels on fait, a minima, le geste conscient et volontaire de prendre le téléphone, ouvrir l'application et appuyer sur le bouton ! De même, certaines applications permettent la réécoute de l'enregistrement avant envoi, ce point a donc été inclus dans le questionnaire sur les usages. Il me semble qu'on peut cependant s'attendre, pour une partie des messages au moins, à ce qu'ils soient tout de même caractéristiques d'une pratique relevant de la proximité communicative, au sens de Koch et Oesterreicher (2001).

D'autres corpus existent évidemment et je souhaite continuer à me former à les utiliser. En effet, selon les cas, il peut se poser soit des problèmes de partage des données (tous les corpus n'étant pas encore en accès libre), mais aussi de formation à l'utilisation, comme je l'ai évoqué. Tout simplement, connaître leur existence est aussi un premier pas à franchir, dans une époque où la quantité de ressources, corpus, articles, publications, se multiplie à un point tel qu'il est de plus en plus difficile de savoir où chercher. Le travail de veille scientifique que je fais pour l'association SIDF vise à résoudre en partie ce problème, en diffusant les annonces de création de tout nouveau corpus qui peut intéresser la diachronie du français. La plateforme Ortolang vise également une meilleure visibilité des corpus existants et leur partage. Le consortium Corli, au comité de pilotage duquel je participe depuis 2019, a aussi été créé dans le but de permettre une meilleure coordination des différentes initiatives et de créer un réseau autour de l'utilisation de corpus et d'outil tout en partageant des recommandations de bonnes pratiques. En particulier, l'idée est de diffuser les principes FAIR (*Findable, Accessible, Interoperable, Reusable*) et de privilégier la science ouverte. Bien sûr, accéder à toujours plus de données ne résout ni le

problème de leur représentativité ni celui de leur qualité. La question reste de savoir ce qu'on attend des corpus et ce qu'on peut en tirer, en ayant conscience des précautions à prendre.

L'utilisation d'Internet comme corpus est également une question qui se pose depuis longtemps. Il existe des bases composées de données récupérées d'Internet et constituées en corpus, mais, comme il a été déjà mentionné dans la littérature scientifique, Internet en soi n'est pas un corpus, de même qu'une base textuelle n'est pas un corpus, ainsi que Bénédicte Pincemin le rappelle :

Tout ensemble de données textuelles ne fait pas corpus, et les critères traditionnels de bonne formation de corpus visent à vérifier trois types de conditions : des conditions de signifiante (cohérence, pertinence), d'acceptabilité (représentativité, régularité, complétude) et d'exploitabilité (homogénéité, volume) (cf. Pincemin 1999). (Pincemin 2012 : 13-14)

La réutilisation et la vérification des données issues d'Internet pose aussi question. Sans même parler de la disparition ou modification possibles des sites, le commentaire YouTube en exergue de cette section, daté d'il y a plusieurs mois, est déjà difficile à retrouver sur la page où il a été observé. Dans le cadre de Corli, un groupe de travail a été constitué pour réfléchir à établir un outil de citation des corpus (le problème se posant également), pour lequel j'ai suggéré qu'on prenne également en compte la question des citations issues des sites web.

Il n'en reste pas moins qu'Internet peut être pour tout linguiste une source précieuse d'attestations. Plus encore, comme l'ont mis en avant Siouffi, Steuckardt et Wionet (2012), Internet peut constituer une source fondamentale pour l'étude des discours sur la langue et des représentations linguistiques, à travers les nombreux « remarqueurs » modernes. C'est en effet un accès direct à toute une pratique pouvant relever de ce qu'on a aussi appelé *folk linguistics* ou linguistique populaire, comme l'ont bien montré Antoine Gautier et Thomas Verjans lors d'une communication aux journées d'étude Variation(s)¹⁰⁹. Même en dehors des blogues spécialisés sur la langue (y compris par des profanes) et des pages ou sites spécialisés (tel Bescherelle ta mère¹¹⁰), toute zone d'expression peut constituer un commentaire métalinguistique. Ainsi, la citation au début de cette section (3.2.) « Sean Paul is the only human in history that speaks doctor's handwriting » nous donne des indications sur les représentations de l'écrit et de l'oral du locuteur, à travers une comparaison entre les productions orales de Sean

¹⁰⁹ Journées que nous avons coorganisées à Nancy avec Bérengère Bouard en novembre 2022 : <https://www.atilf.fr/recherche/manifestations/colloques/20221017-je-variations-en-francais/>.

¹¹⁰ <https://bescherelletamere.fr/>, dernière consultation le 20.11.2022.

Paul et les représentations de l'écriture manuscrite des médecins¹¹¹, réputée indéchiffrable, ce qui constitue probablement l'élément de comparaison.

Pour l'observation des représentations linguistiques, j'utilise également le corpus Garnier Numérique des grammaires et celui des dictionnaires¹¹² et les dictionnaires anciens comme modernes, notamment les différentes éditions des dictionnaires de l'Académie, dont la consultation est largement facilitée par la plateforme du CNRTL¹¹³. Ces corpus nous permettent d'avoir accès à l'évolution du discours métalinguistique, nous donnant ainsi un autre point de vue sur le statut des variantes.

Combiner ces deux approches permet de mener des analyses comparées de l'évolution des usages et de l'évolution des normes. C'est ce que nous avons fait avec Bérengère Bouard dans l'article sur *pour que*, forme d'abord critiquée par les grammairiens mais qui s'est malgré tout imposée dans l'usage (Glikman et Bouard 2021 [31]) ainsi que dans l'article sur la mise en place de l'accord de *tout* devant adjectif (Bouard et Glikman 2018 [21]).

Enfin, dans ma pratique, un autre élément vient aussi contribuer à l'éclairage des phénomènes et du statut des variantes : le recours à des enquêtes linguistiques. En effet, j'ai eu l'opportunité de collaborer avec Mathieu Avanzi lors du lancement des premières enquêtes en 2015 dans le cadre du projet *Français de Nos Régions* (FRR) (cf. Avanzi et al. 2016 [16]), projet que nous avons poursuivi avec *Donnez votre Français à la Science* (DFS)¹¹⁴ (cf. Glikman et al. 2018 [22]). Les enquêtes que nous avons menées s'avèrent un complément nécessaire pour l'observation de certains phénomènes, qui peuvent être absents des corpus écrits et difficiles à faire produire en contexte d'entretien. C'est bien là tout le problème de l'élicitation des données : comment étudier un phénomène qui n'est pas produit dans nos données ? Ces projets sont basés sur le principe du recours à la foule (en anglais *crowdsourcing*), méthode déjà largement employée dans d'autres disciplines et d'autres langues. C'est cette même méthode d'appel à la foule que j'ai utilisée pour la constitution du corpus Les Vocaux. Dans le cas des

¹¹¹ Image qu'on peut assez facilement se représenter quand on a vécu à une époque où les ordonnances n'étaient pas encore systématiquement préparées par ordinateur puis imprimées, mais qui constitue un type d'écriture manuscrite qui n'est bien évidemment pas réservé aux médecins.

¹¹² *Grand Corpus des grammaires françaises, des remarques et des traités sur la langue (XIV^e-XVII^e s.)*, sous la direction de Bernard Colombat, Jean-Marie Fournier et Wendy Ayres-Bennett, en accès libre dans la plupart des bibliothèques : <https://num.classiques-garnier.com/corpusgrammaires> et le *Grand Corpus des dictionnaires [du 9^e au 20^e siècle]* <https://num.classiques-garnier.com/corpusdictionnaires>.

¹¹³ <https://www.cnrtl.fr/dictionnaires/anciens/>.

¹¹⁴ Voir *Curriculum Vitae*, direction de projets, en annexe.

enquêtes FRR, Mathieu Avanzi était particulièrement intéressé par les variantes régionales, telles que la concurrence entre *pain au chocolat* et *chocolatine*, qui a fait beaucoup parler d'elle dans les journaux et sur les réseaux sociaux, afin de dresser de nouveaux atlas linguistiques avec des données récentes (cf. Avanzi 2017, 2019a, 2020). Dans ce but, il a créé des enquêtes linguistiques à questions fermées, que nous avons diffusées via les réseaux sociaux. La composante régionale des enquêtes était l'occasion d'attirer un grand nombre de personnes, le public étant assez amateur de la mise en avant des spécificités régionales ces dernières années. Cela nous a en outre permis d'insérer dans les enquêtes des éléments moins attractifs, telles que des variantes grammaticales, sur lesquels nous n'avions pas forcément d'hypothèse diatopique, dont nous voulions mesurer la diffusion. Les variantes lexicales étant souvent moins stigmatisées que les variantes grammaticales, j'espérais de ce fait abaisser le taux de surveillance des locuteurs en insérant une question sur *malgré que* entre une question sur *bleuet* et *chocolatine*¹¹⁵. Dans certains cas, le participant pouvait également laisser des commentaires, source complémentaire de discours sur la langue. J'ai exploité ces résultats, en tant que degré d'acceptation des formes, dans différentes contributions : *malgré que* et *à cause que* (Glikman 2019 [26]), *avoir été* pour *être allé* (Glikman & Patard 2022 [35]), *de façon à ce que* (communication au colloque *Diachro X* en 2022), et *après que* + indicatif ou subjonctif dont j'ai commencé l'exploitation pour une communication présentée aux journées d'étude *Variation(s) en français* que j'ai co-organisées à Nancy en novembre 2022¹¹⁶. Dans tous les cas, ces enquêtes nous permettent d'accéder aux représentations des locuteurs¹¹⁷. Bien entendu, dans ce cas, il ne s'agit pas de productions (et encore moins de productions spontanées) des locuteurs, mais bien de représentations, d'attitude linguistique, face aux variantes. Le texte introducteur de l'enquête invitait les locuteurs à choisir la réponse qu'ils utiliseraient en famille ou avec des amis. Je parle de représentations dans ce cas, car les participants font des déclarations sur leur usage, cependant, il ne s'agit pas pour autant forcément de leur usage réel, mais bien de ce qu'ils pensent utiliser comme forme. Pour le lexique régional, on peut au moins admettre que les représentations correspondent plus ou moins à des pratiques, tout du moins à

¹¹⁵ Cela n'a cependant pas évité des commentaires de certains participants du type « mais ce n'est pas régional ça, c'est une faute ! ».

¹¹⁶ <https://www.atilf.fr/recherche/manifestations/colloques/20221017-je-variations-en-francais/>

¹¹⁷ Le moyen d'accéder aux représentations des locuteurs est une préoccupation pour plusieurs chercheurs, comme en témoigne encore la publication récente *Research Methods in Language Attitudes* (Kircher et Zipp 2022), que je n'ai malheureusement pas encore pu consulter au moment de la rédaction de ce mémoire.

des zones de diffusion des termes. Dans le cas des variantes grammaticales, toujours plus stigmatisées¹¹⁸, on peut s'attendre à ce que la pression de la norme scolaire joue sur les déclarations d'usage et que les réponses comportent une part de surveillance. C'est d'ailleurs ce qui est ressorti de l'analyse de certains commentaires. Les participants avaient admis utiliser une variante, tout en reconnaissant que « c'était une faute » (cf. Glikman et Patard 2022 [35]). Dans tous les cas, cela nous dit toujours quelque chose, ne serait-ce que sur l'intégration de cette norme. Les questions étaient soit à choix multiples soit avec des échelles de degré (*utilisez-vous souvent, jamais, etc.*). Le choix des questions doit être fait avec soin afin d'obtenir la réponse sur l'élément voulu, comme je l'ai évoqué à propos de *malgré que*, pour lequel nous n'avons pas proposé la variante construite avec *bien que*. De même, des problèmes de compréhension peuvent se poser. Lorsque cela est nécessaire en complément des autres approches, en particulier sur corpus, il m'arrive de suggérer à mes étudiants de master la mise en place d'une enquête linguistique¹¹⁹. Dans ce cas, je leur demande toujours de procéder à une pré-enquête à tester auprès de quelques personnes pour vérifier à la fois les problèmes de compréhension, notamment dans la formulation des consignes, mais aussi d'élicitation des éléments attendus.

Dans le projet DFS, l'objectif était d'attirer également le public à travers d'autres éléments ludiques et qui contribuent aussi à l'informer, ce que nous avons mis en place à travers la création de quizz et d'un algorithme de géolocalisation (basé sur le principe du « dis-moi comment tu parles et je te dirai d'où tu viens »), développé par Jean-Philippe Goldman et Yves Scherrer sur la base des premiers résultats des enquêtes FRR (Goldman et al. 2018 [24]). Les résultats des quizz ont pu être comparés avec les résultats des enquêtes, comme je l'ai déjà évoqué à la section précédente, ce qui nous a permis de mesurer un autre aspect de la diffusion et des représentations linguistiques : même s'il ne l'emploie pas dans sa pratique quotidienne, tout le monde connaît désormais le mot *chocolatine* (Glikman et al. 2018 [22]).

À partir des projets FRR et DFS, dont le sujet est particulièrement adapté, j'ai plusieurs fois mené des actions de diffusion auprès du grand public, que ce soit dans des interviews afin de faire connaître les projets, mais aussi à travers des animations à la Fête de la Science, à Pint Of

¹¹⁸ Cf. Poplack et Dion (2021 : 84) : « la position normative qu'elle [la norme] incarne reste particulièrement robuste en morphosyntaxe, car c'est le domaine de la grammaire le plus exposé aux jugements de rectitude. »

¹¹⁹ Les étudiants sont formés aux méthodes de recueil de données, y compris l'enquête, dès la licence, mais plusieurs de nos étudiants de master n'ont pas suivi la licence à l'Université de Strasbourg.

Science et même lors d'une animation dans un collège¹²⁰. Ces animations en direction du grand public permettent bien sûr de faire connaître les projets au public, dont nous avons besoin pour recueillir de nouveaux participants. Elles constituent aussi un réel retour des résultats de la recherche vers le grand public. Là encore, j'ai profité de l'attractivité des phénomènes régionaux et des phénomènes phonétiques¹²¹ pour sensibiliser le public à l'histoire de la langue et à la diachronie. J'ai construit une activité sous forme de jeu autour d'une frise chronologique retraçant de grandes étapes de l'histoire du français, à travers quelques exemples d'évolution de mot et même d'évolution des pratiques de l'écrit. Nous avons également préparé une année une activité autour des chaînes de référence en lien avec le projet Democrat. Dans le cadre de ces projets, j'ai été amenée à collaborer avec des acteurs extérieurs au milieu universitaire, la Délégation Générale à la Langue Française et aux Langues de France, qui a soutenu nos projets, d'une part, et le siège français de Wikimedia, avec qui nous avons collaboré autour de l'enregistreur Lingua Libre. J'ai encadré suite à cela Marie Didienne (2019), qui a soutenu un Master de Sciences du Langage sur la diffusion des créations terminologiques de *FranceTerme*¹²². Travailler avec des partenaires extérieurs au monde universitaire, même si cela reste ici des partenaires proches du monde de la recherche, est toujours enrichissant.

Qu'est-ce que ces différentes données nous apprennent ? Les enquêtes diffusées par les réseaux sous forme d'appel à la foule avec libre participation peuvent bien sûr poser un problème du fait d'un possible déséquilibre du type de public qui y répond et de l'échantillon qui, par définition, n'est pas représentatif. Dans le cas de la première enquête, qui a mobilisé plus de 10 000 participants, nous avons fait l'hypothèse que le nombre de réponses pouvait suffire à lisser les différences. Nous n'avons ensuite plus eu autant de participants pour les enquêtes suivantes. Cependant, si le panel n'est pas forcément représentatif, comme nous l'avons évoqué précédemment, prétendre à la représentativité est toujours très difficile, tant pour les enquêtes linguistiques que pour les corpus, comme chacun sait. La question se pose donc systématiquement et doit être considérée avec précautions. Ensuite, je suis convaincue que cela nous apprend tout de même quelque chose. C'est dans tous les cas un indice, une mesure, d'un certain état de fait. Ce qui compte, il me semble, c'est donc surtout de faire attention aux

¹²⁰ Voir *Curriculum Vitae*, communications et activités de vulgarisation, en annexe.

¹²¹ Le stand à la Fête de la science était organisé avec ma collègue Camille Fauth, qui expliquait de son côté l'appareil phonatoire et permettait de visualiser le spectre de la voix des spectateurs dans Praat.

¹²² Organisme du Ministère de la Culture chargé des créations terminologiques pour « le dire en français », tel *autocaravane* pour *camping-car* ou encore *bogue* pour *bug* <http://www.culture.fr/franceterme>.

conclusions qu'on se permet d'en tirer, tant pour les enquêtes que pour l'utilisation des corpus. Dans tous les cas, il faut être conscient de ce dont est constitué le corpus ou l'enquête, de sa composition, y compris formelle, des biais possibles, être conscient des problématiques de bruits et de silence dans toute recherche sur corpus, comprendre ses requêtes, etc. En bref, connaître ses limites. Cela ne veut pas dire pour autant qu'on ne peut rien faire, simplement qu'il faut savoir dans quelles limites on peut tirer des conclusions, être conscient de ce que les données disent ou non. Même partielles, ces différentes données nous apportent un éclairage sur les faits. Il en est de même pour les commentaires métalinguistiques sur Internet, qui vont souvent dans le même sens pour la même raison, qui est que c'est le même genre de personnes qui les produisent (en général pour critiquer la déchéance de « notre belle langue française »). Pour moi, l'analyse de ces données est à considérer comme un éclairage, une lumière portée sur un aspect du phénomène. L'important est alors de combiner les points de vue, les éclairages, pour parvenir à une vision la plus globale possible du phénomène. Tout comme j'ai recours à un faisceau d'indices pour établir la subordination, établir les variantes, j'ai recours à un faisceau de données pour établir le statut de chaque variante.

**4. Comprendre le changement
linguistique. Autour des usages et des
normes : une approche centrée sur le
locuteur**

4. Comprendre le changement linguistique. Autour des usages et des normes : une approche centrée sur le locuteur

*Le temps passe et passe et passe et
beaucoup de choses ont changé
[...]
On fait le Bilan*

Neg'Marrons, *Le Bilan*, 2000.

Après avoir abordé la question de l'établissement des variantes au chapitre 2., j'ai traité au chapitre 3. la question des paramètres et des données pour analyser la variation. Dans ce chapitre, je reviendrai sur l'objectif principal du travail de diachronicien : le changement linguistique, et en particulier sur ce que l'étude de la variation nous apprend sur lui. Ce sera l'occasion de faire le point sur les perspectives de recherche que mes travaux menés jusqu'ici ont ouvertes, sur le champ de recherche où je veux diriger des recherches et sur l'approche du changement que je souhaite continuer à développer. La section 4.1. sera consacrée à la caractérisation des discours, puis nous aborderons en 4.2. le jugement des locuteurs, avant de faire le bilan en 4.3. sur ce que cela nous apprend sur le changement linguistique et les perspectives de recherche que cela ouvre.

4.1. La caractérisation des discours

*Tu sais
Tu sais ce soir j'ai vu
tous les bijoux de la pop*

Eddy de Pretto, *Fête de trop*, 2018.

Comme je l'ai évoqué jusqu'ici, une partie de mes travaux ont porté sur la caractérisation des discours, d'une part sur les particularités de l'oral représenté, d'autre part sur les particularités liées aux genres. Dès ma thèse et mes premiers travaux (cf. notamment les conclusions de Glikman 2008b [3]), j'avais avancé l'hypothèse que la réalisation de la parataxe pouvait être liée à une différence oral / écrit. Sa disparition progressive dans les textes, alors qu'on la retrouve en français contemporain oral, comme le montre bien la citation en exergue ci-dessus sous 4.1. et l'article de presse paru récemment sur la question¹²³ pouvait être liée, selon l'hypothèse que je faisais alors, au développement d'une prose écrite qui s'éloigne de

¹²³ <https://www.leparisien.fr/societe/javoue-je-sais-pas-jai-limpression-il-va-faire-chaud-mais-ou-est-passe-le-que-dans-le-langage-parle-04-09-2022-OFXVNUKI7BCTTIF5VMRFSJK4C4.php>

l'oralité pour aller vers des caractéristiques plus proches de la scripturalité conceptuelle, autrement dit, de la distance communicative selon les paramètres de Koch et Oesterreicher (2001). C'est ce qui m'a amenée à poursuivre mes travaux sur la caractérisation à la fois de l'oral représenté et des genres textuels. Bien que ces hypothèses aient été amenées un peu « naïvement » et peut-être pas « pour les bonnes raisons » ni « avec les bons arguments », comme me l'avait gentiment fait remarquer le regretté Peter Koch lors de ma soutenance de thèse, il y avait quelque chose de juste. La suite de mes travaux et des travaux des collègues qui, comme je l'ai dit en introduction, ont mené des recherches dans la même direction, ont confirmé l'importance de la distinction des discours et ont mis en avant des différences entre les genres, en lien avec les questions de distance et de proximité communicative, ayant une influence sur la réalisation et la distribution des formes. Ainsi, cela peut expliquer que la parataxe, conformément à ce que j'avais pressenti alors, n'a pas disparu du point de vue de la grammaire du français, mais est bien une construction qui relève davantage de la proximité communicative. En cela, on retrouverait le même type de « faux » changement que celui que l'on peut observer pour le *ne* de négation (cf. Stark et Dufter 2007).

Avant même d'étudier les spécificités de l'oral représenté, se pose la question de son statut syntaxique et de la distinction entre le discours rapporté au style direct et indirect, qui va influencer sur la réalisation. Pour cette question, je me suis fondée sur les travaux de Sophie Marnette (1999, 2001), qui a clairement établi la problématique et montré les distinctions possibles pour la période médiévale, en particulier le changement de personne. La distinction entre l'oral représenté au style direct et au style indirect n'est pas toujours facile à faire, tant pour l'ancienne langue que pour les périodes contemporaines, ne serait-ce qu'avec l'existence du discours rapporté au style indirect libre. Le statut syntaxique du discours rapporté au style direct, entre autonomie et subordination, fait également encore débat. Christiane Marchello-Nizia (2012) a ainsi montré que les éléments introducteurs du discours direct sont porteurs de spécificités liées davantage au discours direct qu'à la narration, comme l'emploi du présent. Enfin, la question de l'accès à l'oral pour l'ancienne langue, comme je l'ai évoqué en introduction, a donné lieu à différents types de travaux ces dernières années. Un courant vise à accéder à des traces réelles de l'oral, comme les travaux d'Aude Wirth-Jaillard (2013), qui s'intéresse à l'inscription de l'oral « réel » dans les documents comptables médiévaux, ainsi que les travaux visant à la mise en avant de sources jusqu'ici peu étudiées pour l'histoire de la langue. Un autre courant, auquel je me rattache, considère qu'il s'agit toujours bien de représentation, mais s'intéresse d'une part aux spécificités de l'oral représenté, d'autre part aux marques d'oralité, d'une manière plus générale, qu'on peut observer dans les textes de différents genres.

Les travaux sur l'étude des spécificités de l'oral représenté ont pu bénéficier de l'annotation des frontières du discours direct (établie automatiquement sur la base des guillemets présents dans les éditions modernes), ainsi distingué des passages de narration, dans la Base de Français Médiéval. Cette annotation permet de mener des études outillées et de faire des analyses de type statistiques robustes. Céline Guillot-Barbance, Alexei Lavrentiev, Bénédicte Pincemin et Serge Heiden se sont ainsi penchés sur l'analyse des spécificités du discours direct, en particulier mis en relation avec les parties du discours et le genre textuel (informations toutes deux présentes dans la BFM) (Guillot et al. 2013, 2015, 2017). Leurs travaux ont montré l'intérêt d'une telle approche et de la prise en compte de cette distinction :

Or il est apparu de manière très évidente que c'est toujours l'axe discours direct / non discours direct qui l'emporte sur tous les autres. [...] Par ailleurs, il est apparu très nettement que ce sont à peu près toujours les mêmes catégories qui construisent les oppositions les plus significatives. Nous avons montré que la répartition des catégories morphosyntaxiques s'ordonne selon un principe relativement clair, avec du côté du discours direct les catégories qui relèvent plutôt du domaine verbal et du côté du non discours direct les catégories liées au groupe nominal. Ces premiers résultats nous paraissent ouvrir des pistes de recherche prometteuses, dans la mesure où ils étayaient l'hypothèse d'une grammaire propre au discours direct, dont les grands linéaments semblent se dessiner déjà. [...] Les grandes tendances que nous avons dégagées, et surtout, la netteté avec laquelle le discours direct semble se distinguer du reste, nous invitent à réfléchir aussi à l'importance de ce paramètre dans l'exploitation linguistique que nous faisons des ressources médiévales. Il est très rare qu'on prenne en compte le facteur discours direct / non discours direct dans les recherches diachroniques ou synchroniques sur le français médiéval. Il serait pourtant très utile de l'intégrer de manière plus systématique à l'analyse linguistique, au même titre que la variation vers/prose ou la variation dialectale par exemple. (Guillot et al. 2015)

Parallèlement à leurs travaux, nous avons mené le même genre d'étude avec Nicolas Mazziotta (Glikman et Mazziotta 2013 [12] et Mazziotta et Glikman 2019 [27]) en nous appuyant sur le marquage du discours direct et sur l'annotation syntaxique des textes réalisée dans le cadre du projet SRCMF. Ces travaux sont donc complémentaires des travaux de Céline Guillot-Barbance et ses collègues, en cela que les éléments de comparaison pris ici en considération ne sont plus les parties du discours, mais les fonctions syntaxiques. Nos travaux ont également mis au jour l'existence de spécificités syntaxiques préférentiellement associées à l'oral représenté, tout en conservant une grammaire commune :

À l'exception des quelques constructions que nous venons de traiter (apostrophe et interjection, →3.1), les spécificités syntaxiques mises en évidence ne sont pas réservées à un type discursif. Il est clair que les fonctions qui constituent le noyau dur de la syntaxe se rencontrent à la fois en DD et en narration : il s'agit donc bien dans les deux cas de la même grammaire (comme proposé par Blanche-Benveniste et al. 1990). Toutefois, les tendances repérées mettent en relief des constructions « préférées » par l'un ou l'autre

des deux types, ce qui constitue un indice supplémentaire qu'il s'agit bien de deux types distincts. Par ailleurs, il nous reste à interroger les particularités des constructions préférées de chacun des types : que pouvons-nous en effet déduire de ces associations ?

Il est remarquable que les spécificités associées au DD correspondent au mouvement de l'évolution du français : 1/ le passage de formes verbales simples à davantage de formes complexes (→2.2.1 a) ; 2/ la généralisation de l'expression du sujet obligatoire (→2.2.1 b) ; 3/ et, dans une moindre mesure (voir →2.2.1 d), la généralisation de l'emploi du forclusif dans la négation. [...] Par ailleurs, nous avons constaté que certaines de ces particularités, comme l'association du réfléchi à la narration (→2.2.2 b) et celle de l'objet et de l'attribut au DD (→2.2.1 c), pourraient également être révélatrices de préférences d'ordre lexical. Le détail de ces associations lexicales reste à décrire, mais on pourrait penser ici à des cas comme les verbes épistémiques (voir Glikman à par. [(2012) [8]]). (Glikman et Mazziotta 2013 : 55-57)

Nous avons poursuivi cette étude en élargissant la comparaison à trois textes de genres distincts. L'examen des spécificités syntaxiques nous a permis de montrer des différences entre ces textes, mais également une différence de réalisation entre les propositions indépendantes et subordonnées, différence qui avait déjà été soulevée pour des phénomènes tels que l'ordre des mots ou l'expression du sujet, (cf. Marchello-Nizia et Prévost 2020 : 1062) :

Les analyses ont montré que la principale partition est celle qui oppose les propositions indépendantes (→2.2) aux subordonnées et *roland* aux autres textes. [...]

Toutefois, l'OR [Oral Représenté] de *roland* partage des caractéristiques avec celui de *qgraal* et de *clari* : l'expression pronominale du sujet des phrases et l'emploi de formes infinitives avec auxiliaire modal paraissent des caractéristiques stables de l'OR, alors que l'emploi de *si* ou d'un autre articulateur (en concurrence) est plutôt caractéristique de la narration. [...]

De manière plus générale, la présence d'une opposition OR/narration ne fait aucun doute pour les textes étudiés, mais il ne s'agit pas du seul paramètre influant sur la syntaxe. Malgré une indéniable stabilité des caractéristiques de l'OR et de la narration dans des textes aussi différents que ceux dont nous avons parlé, notre étude montre également qu'il est nécessaire de distinguer l'analyse des subordonnées de celle des principales. Le type de texte doit aussi être pris en considération. L'opposition entre *roland* et les autres textes est plus forte que l'opposition entre oral représenté et narration. Reste encore à étudier si cette particularité de *roland* est liée au genre, à la forme (en vers tandis que les autres textes de l'étude sont en prose), ou encore à l'espace temporel (presque 2 siècles séparant ce texte des deux autres). (Mazziotta et Glikman 2019 : 13-14)

Dans ce même courant, d'autres chercheurs se sont intéressés aux marques d'oralité d'une manière plus générale, comme en témoignent les projets de recherche menés par Gabriella Parussa à Paris¹²⁴ ou par Amalia Rodriguez Somolinos à Madrid et les nombreux volumes

¹²⁴ Gabriella Parussa a ainsi fait soutenir des thèses récentes sur le sujet, comme celle de Pierre Vermander (2020).

collectifs sur la questions (Rodriguez Somolinos (éd) 2013, 2016, Lagorgette et Larrivée (éds) 2013, et Ayres-Bennett et al. (dir.) 2018 [v], Glikman et al. (éds) 2019 [viii] que j'ai contribué à coéditer). Dans ce contexte, en parallèle des travaux que je viens de citer, je me suis intéressée au rôle et à la distribution des verbes parenthétiques (ou épistémiques) dans les textes et leur rapport à l'oralité, dans la continuité de mes travaux sur la parataxe (Glikman 2012 (2009) [8], Schneider et Glikman 2015 [15], Glikman et Schneider 2018 [23]). Là encore, les résultats de ces études ont permis de mettre en avant la stabilité de ces structures. Leur évolution constatée à première vue relèverait en fait davantage, comme pour la parataxe, d'une distribution selon le degré d'oralité. Ces recherches contribuent ainsi à montrer le changement qui s'opère dans le rapport à l'écrit à la fin de la période médiévale :

On voit ainsi à travers l'étude des CP [constructions parenthétiques] ce qui a déjà été observé par ailleurs, un changement qui se produit à partir du moyen français et de la Renaissance dans le rapport aux textes et à l'écrit, et dans le rapport de l'écrit à l'oralité, et le renforcement des marques explicites dans la prose, dès les premiers textes en prose. Ces marques vont se développer dans le même mouvement que la prose se développe elle aussi. La présence de marques explicites dans les CP à l'écrit a pu faire penser que la CP moderne était de création récente. Nos recherches montrent qu'elle existe depuis l'ancien français, mais qu'elle reste cependant fortement associée à l'oralité. Ce serait donc la proximité des textes anciens à l'oralité qui aurait permis d'observer la CP moderne, le développement de l'écrit en prose privilégiant les constructions marquées.

Cette étude montre également l'importance de la prise en compte des genres et de la forme dans la réalisation de constructions, les enjeux stylistiques entrant en compte, et tout l'objet de la recherche diachronique est bien d'évaluer le poids des contraintes grammaticales ou stylistiques dans la construction des CP et leur évolution. Ici, il semblerait que l'évolution des CP est davantage conditionnée par les contraintes stylistiques liées au développement de la prose. (Glikman et Schneider 2018 [23] : 332-333)

Ces travaux m'ont ainsi permis de poursuivre et de confirmer une hypothèse que j'avais posée dans ma thèse. J'y abordais déjà la question des contraintes stylistiques, en regard des contraintes grammaticales. Dans ce cas, il n'y a donc pas eu de changement dans la grammaire, mais bien dans le style. Il faudrait poursuivre les études sur d'autres phénomènes pour continuer à étayer cette hypothèse. On pourrait également mettre en système le développement des marques explicites dans la prose avec le développement des marqueurs conjonctifs observés par ailleurs. Centré ici sur le changement après la période médiévale, il serait intéressant également de voir si d'autres grandes périodes de mutation dans le rapport à l'écrit peuvent transparaître de ce genre d'étude, que ce soit dans le moment de « resserrement » de la langue durant le français classique, ou encore à la naissance du roman réaliste, qui modifie la représentation de l'oral au 19^e siècle (cf. Stark et Dufter 2007), et bien sûr à la « révolution » récente en lien avec

le développement des écrits numériques et des communications médiées par ordinateur. Une perspective qui me semble aussi complémentaire mais que j'ai trop peu prise en compte jusqu'ici est la question de l'évaluation des textes en fonction de leur réception, comme cela a été mis en avant, dans une perspective de sociolinguistique historique, dans le volume dirigé par Anne Carlier et Céline Guillot-Barbance sur le passage du latin au français (Carlier et Guillot (éds) 2018). La question de la circulation des textes et de leur diffusion pourrait être davantage prise en compte pour étudier la diffusion des phénomènes et de la représentation des discours. De même, je souhaite développer dans mes recherches la prise en compte de sources jusqu'ici peu étudiées, grâce aux développements des projets récents de constitution de corpus pour faciliter l'accès à ces sources, comme le *Corpus 14* ou le corpus CONDÉ¹²⁵.

La question de l'oral représenté, comme nous venons de le voir, est ainsi directement en lien avec les spécificités associées aux genres discursifs. Une partie de mes recherches contribuent à l'étude des genres et de leur évolution. Comme je viens de le montrer, mes travaux sur les marques d'oralité ont pris de plus en plus en compte la question du genre. Ma participation aux études sur les chaînes de référence (désormais CR) s'inscrit également dans cette problématique. Avec mes collègues Céline Guillot-Barbance et Vanessa Obry d'abord sur les textes narratifs médiévaux (Glikman et al. 2014 [14]), puis avec Bénédicte Pincemin qui nous a rejoint sur les récits brefs (13^e - 16^e) (Obry et al. 2017 [20]), nous avons analysé la composition des CR dans les textes médiévaux, mettant en avant les spécificités liées aux genres. N'étant moi-même pas spécialiste au départ des questions liées aux CR, j'ai appris beaucoup lors de ces collaborations. De nouvelles pistes se sont ainsi ouvertes. Ces études ont également confirmé que le genre conditionne la réalisation des CR. J'ai eu l'occasion de poursuivre ces recherches avec mes collègues de Strasbourg dans le cadre du projet Democrat, à la fois sur ce que les CR nous apprennent sur l'évolution des genres (cf. Oberlé et al. 2018 [25] sur les textes encyclopédiques) et sur leur rôle dans la structuration textuelle et son évolution (cf. Capin et al. 2021 [33]). Ces travaux contribuent là encore à montrer l'importance de la prise en compte du genre. Une étude que nous avons évoquée alors mais que nous n'avons pas encore mise en place serait de corréler l'étude des CR à l'analyse de la distinction entre oral représenté et narration. Un focus sur les maillons de reprise après le discours direct resterait à faire et pourrait donner des résultats intéressants, à la fois sur la caractérisation de l'oral représenté et sur le

¹²⁵ Projet *La constitution d'un droit européen : Six siècles de coutumiers normands*, <https://www.unicaen.fr/coutumiers/conde/accueil.html>.

statut du discours direct dans la chaîne narrative, à la suite des travaux de Dominique Lagorgette (1998).

Ainsi, mes travaux ont contribué à montrer l'apport que peut avoir la caractérisation des discours selon les passages au discours direct ou en narration, selon les genres discursifs, mais aussi, et cela me paraît complémentaire, selon les paramètres de distance et proximité communicative de Koch et Oesterreicher. Mais nous avons vu que d'autres paramètres pouvaient entrer en jeu, comme la distinction entre proposition indépendante et subordonnée, ou encore selon la réception des textes, constituant un faisceau de paramètres entrant en compte dans la caractérisation des discours. C'est également dans cette perspective que je souhaite contribuer à la description du genre des SMS vocaux dans mon projet actuel. Dans ces vocaux, comme dans les SMS écrits, on peut s'attendre à ce que la prise en compte du destinataire ait un impact important sur les messages, y compris ceux émis par le même locuteur.

La prise en compte de ces différents paramètres nous apprend que nous pouvons observer, à travers l'oral représenté, des traces du changement linguistique. Certains genres discursifs peuvent, à un moment donné de leur évolution historique, avoir des caractéristiques relevant plus ou moins de la proximité communicative et présenter des spécificités liées à ce pôle. Les réalisations langagières peuvent être porteuses à la fois de spécificités liées au médium, qui relèvent pratiquement davantage de contraintes liées au code, oral ou écrit, et de spécificités traditionnellement associées soit au pôle de la proximité, soit de la distance, communicative, et qui sont surtout des associations relevant du style (au sens de la liberté de choix du locuteur), mais constituent cependant des structures « préférées », pouvant varier au cours du temps. Cette distinction, et l'évolution du rapport à l'écrit, a pu faire croire à des changements linguistiques mal compris, qui sont en fait une distribution selon ces paramètres et qui ont pu être masqués par le développement de la prose vers le pôle de la distance communicative. De même, il ne faut pas confondre le développement d'un type d'écrit de la proximité avec des changements dans la langue, comme certains phénomènes dans la représentation de l'oral dans Frantext par exemple peuvent le donner à penser, ainsi que le montrent Elisabeth Stark et Andreas Dufter sur le *ne* de négation :

Par contre, l'augmentation rapide au milieu du XIX^e observée dans les études précédentes pourrait bien s'avérer comme artéfact de l'évolution littéraire du réalisme au naturalisme. Somme toute, ces observations suggèrent que depuis le XVII^e siècle, il y aurait eu moins de changement que ne le reflètent au premier abord les sources disponibles. (Stark et Dufter 2007 : 125)

Il me semblerait intéressant également de poursuivre dans cette direction : le changement dans la représentation de l'oral et dans les genres vis-à-vis du changement dans la langue, où

l'on retrouve ma question initiale, de distinguer ce qui relève de la grammaire, contrainte, et ce qui relève du style, choix.

4.2. Usages, normes et représentations sur la langue

*Je le sais
Sa façon d'être à moi, parfois vous déplaît
Autour d'elle et moi le silence se fait
Mais elle est ma préférence à moi*

Julien Clerc, *Ma préférence*, 1978.

Dans les parties précédentes, nous avons pu voir que mon objet d'étude premier était l'ancien français, avec une prise en compte de l'évolution en français moderne. J'ai ensuite élargi mes périodes d'observation au moyen français, puis au français préclassique, classique et contemporain. Une partie de mes travaux aujourd'hui porte plus spécifiquement sur ce dernier, dans le cadre de mes projets sur les variations contemporaines et les changements en cours, et sur les périodes postérieures à l'ancien français. Cependant, ce choix n'est pas pour m'éloigner de la recherche sur les états de langue anciens. Au contraire, j'ai voulu élargir mon champ d'étude à une diachronie longue, même si je reste persuadée qu'une étude du fonctionnement sur une diachronie courte, voire sur différentes synchronies, pour comprendre la place d'une variante dans le système à un moment T, reste importante¹²⁶. Les points communs entre ce que j'ai pu observer sur l'ancien français et les aboutissements en français moderne, comme je viens de l'évoquer en 4.1. pour la parataxe et les verbes parenthétiques, qui semblent constituer des variations sur un temps long, m'ont conduite à m'intéresser de près à la variation contemporaine, notamment selon d'autres paramètres. J'ai souhaité approfondir mes recherches sur les changements « en cours » ou, au contraire, « au long cours », et ainsi mieux saisir les processus en jeu et les interactions entre les différents paramètres de variations, en particulier oral / écrit et distance / proximité communicative. Enfin, l'élargissement de ma période d'étude me permet également d'intégrer le paramètre du rapport entre usages, norme(s) et représentations, avec la prise en compte des discours sur la langue et du jugement des locuteurs (→ 3.2.). Selon les époques, la situation n'est évidemment pas comparable, avec la place de

¹²⁶ C'est l'une des questions centrales de ce qu'est la diachronie, comme le rappelait encore Bernard Combettes aux journées Variation(s) à Nancy (novembre 2022) : elle peut aussi être une comparaison des différentes synchronies. Voir aussi la position de Kragh et Schøsler (2021 : 97) : « Il faut évidemment procéder en distinguant l'analyse synchronique d'un paradigme de son analyse diachronique, qui ne peut s'effectuer qu'après l'établissement de paradigmes synchroniques successifs. ».

l'écrit dans la société, qui évolue au fil du temps, le développement de l'école ainsi que la zone de diffusion et de contact qui se modifie avec les communications facilitées sur un large périmètre par les médias tels que la radio, la télévision, et maintenant Internet. La prise en compte des représentations, à travers les discours sur la langue, que ce soit un discours normatif ou le jugement du locuteur, constitue un apport important pour comprendre la diffusion des phénomènes linguistiques et mettre au jour le statut des variantes. Ainsi, cette partie de mes recherches vise à intégrer le locuteur dans l'analyse du changement. Cette idée d'intégration du locuteur est présente dès l'article fondateur de Weinreich, Labov et Herzog, à propos de la direction du changement :

These investigations have described changes that continue in the same direction over several generations. Persistence in the direction of change suggests that these changes are variables which have been evaluated in the same way by the speech community over a considerable period of time (Weinreich, Labov et Herzog 1968 : 146)

La prise en compte du locuteur s'est développée dans de nombreux travaux sur le changement, tant dans le courant tenant de la sociolinguistique historique s'intéressant au locuteur et à son rôle, comme les travaux de Lodge, Lusignan, Banniard, Poplack, Ayres-Bennett (2004), entre autres, que dans les approches s'intéressant davantage à la mise en place des normes collectives, comme notamment les travaux de Wendy Ayres-Bennett sur la standardisation et la mise en place du « bon usage » (Ayres-Bennett et Sejjido (éds) 2013, Ayres-Bennett et Bellamy (éds) 2021).

Siouffi (2020) distingue les « représentations » des « attitudes », les unes comme les autres n'étant pas toujours facilement observables (→ 3.2.) et pouvant jouer un rôle dans l'évolution linguistique :

Par « représentations » de la langue, on peut entendre les visions qu'une communauté en développe, les manières dont elle se l'imagine, l'idée qu'elle en a, les valeurs qu'elle y associe, les attentes conscientes ou inconscientes qu'elle a à son égard. Ces représentations éparses peuvent être cristallisées en imaginaires, en attitudes, en idéologies, en discours et en actes. Pour mesurer leur influence sur les usages, il faut en tout cas pouvoir les analyser, et, pour ce faire, disposer de matériaux explicites. (Siouffi 2020 : 150)

[Au 17^e siècle] Les prescriptions n'étant pas toujours explicites, on peut ici juger que c'est le poids des représentations, tant esthétiques que sociales, qui a été le plus grand, comme dans le domaine phonétique, par exemple (valorisation de certains phonèmes comme le [ə], jugé typiquement français, règles d'enchaînement par l'euphonie, élisions, dénasalisation, phénomènes de dissimilation). (Siouffi 2020 : 151)

Par le terme *attitudes*, on peut vouloir englober les opinions, représentations, idéologies, sentiments, évaluations plus ou moins explicites des locuteurs autour des langues, les degrés d'implication psychosociale des locuteurs dans leurs usages, ou, de façon plus

Comprendre le changement linguistique. Autour des usages et des normes : une approche centrée sur le locuteur

restreinte – ce qui sera le cas ici – les attributions de valeur, de prestige, ou encore le choix entre conservatisme et innovation. Ces attitudes sont à comprendre comme moins discursives que les représentations et plus orientées vers les comportements langagiers, mais pour les périodes anciennes, elles sont essentiellement analysables par l'intermédiaire des discours qui en portent la trace. (Siouffi 2020 : 155)

Dans mes travaux en diachronie longue ou en diachronie contemporaine, j'associe ainsi l'analyse sur corpus à l'observation des discours sur la langue. Autant que possible, j'intègre l'évolution des discours métalinguistiques, à travers les discours explicites émanant ou non d'experts (grammaires, dictionnaires, chroniques de langue, commentaires...), mais aussi le jugement linguistique des locuteurs sur les formes, à travers les enquêtes linguistiques (→ 3.2.). Cela permet à la fois de prendre la mesure de la diffusion des formes et celle de l'évaluation du locuteur sur cette variante, complément nécessaire à son étude en corpus. Comme le soulignent Shana Poplack et Nathalie Dion (2021) :

En résumé, nous avançons que le taux d'occurrence, première source de données vers laquelle se tournent les sociolinguistes lorsqu'ils concluent à un changement, n'explique qu'une partie – toute petite de surcroît – de ce qui se passe en réalité. Le conditionnement du choix de variante, indépendant du taux, est plus révélateur. En outre, la fréquence d'une variante ne reflète pas nécessairement sa productivité. (Poplack et Dion 2021 : 111)

Dans une approche sociolinguistique, Poplack et Dion (2021) utilisent la notion de conditionnement du choix comme révélateur ou non du changement :

L'éviction progressive d'une variante parmi ses concurrentes qui expriment la même valeur référentielle constitue évidemment le cas de figure le plus simple et le plus intuitif du changement. Mais, comme le précisent entre autres Poplack et Tagliamonte (2001 : 92), la fluctuation dans la fréquence peut avoir d'autres sources ; même lorsqu'elle signale un changement, il est rare que celui-ci se limite simplement à la montée ou au déclin d'une variante. Bien plus révélatrice est la cartographie des facteurs contextuels affectant la probabilité de choisir une forme donnée plutôt que ses formes concurrentes, ainsi que la direction de leurs effets. Ce conditionnement constitue pour nous la structure, ou grammaire, de la variabilité ; nous nous en servons dans nos comparaisons diachroniques visant à identifier et à caractériser le changement. (Poplack et Dion 2021 : 85)

Le changement linguistique n'est pas toujours ce qu'il paraît être, ne serait-ce que parce qu'il opère en grande partie à un niveau plus profond que la simple augmentation ou diminution superficielle de variantes concurrentes. (Poplack et Dion 2021 : 108)

De la même manière, mes travaux récents m'ont permis de mettre au jour l'importance de l'évolution du statut des variantes par les locuteurs, dans lequel j'inclus ici le discours métalinguistique qui est également représentatif d'un point de vue sur la langue. Le statut des variantes est révélateur de ce qui conditionne ainsi le choix des locuteurs et son rapport au

changement linguistique. C'est une manière d'approcher ce que Siouffi, Steuckardt et Wionet appellent « l'engagement du locuteur dans le choix de la forme » (Siouffi, Steuckardt et Wionet 2018 : 433).

Grâce aux enquêtes linguistiques menées dans le cadre des projets FRR et DFS (→ 3.2.), j'ai pu également cartographier les changements dans les représentations linguistiques en mettant au jour des différences entre la Suisse, la Belgique et la France dans le taux d'acceptation des variantes. Ainsi, pour les locutions *malgré que* et *à cause que*, jugées incorrectes par le discours normatif (Glikman 2019 [26]), comme pour la concurrence *avoir été / être allé* en contexte locatif (Glikman et Patard 2022 [35]), il ressort de l'analyse des taux d'acceptation révélés par les enquêtes linguistiques (Avanzi et al. 2016 [16]) un rejet plus important de ces formes en France que dans les autres pays francophone d'Europe. Encore une fois, cela ne dit rien de l'emploi de ces formes en production, mais peut s'avérer révélateur du poids des représentations et de la norme, ainsi que des attitudes conservatrices des locuteurs de France.

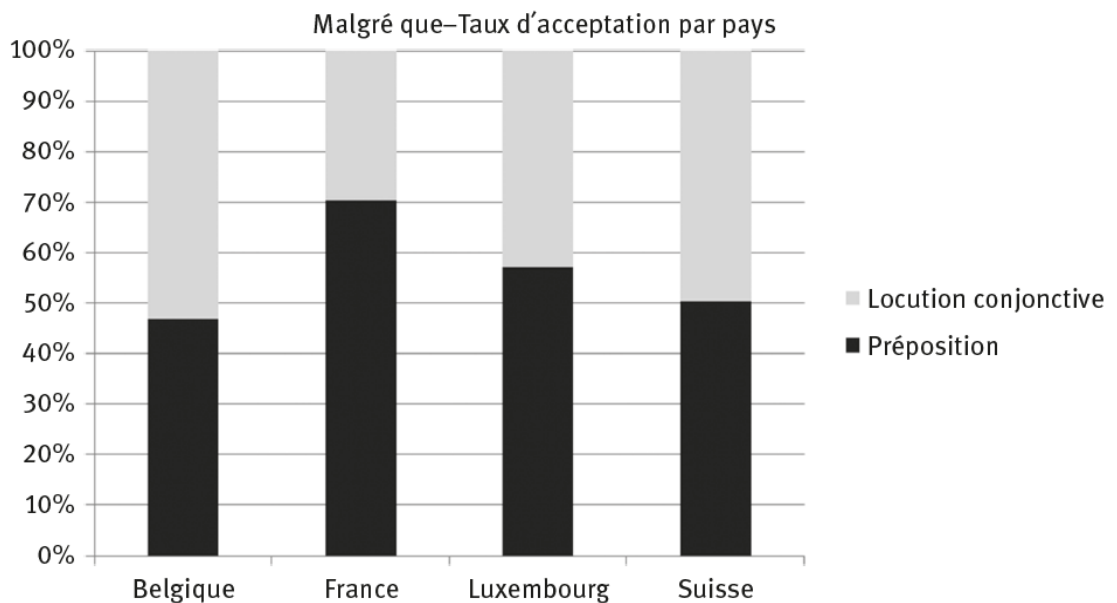
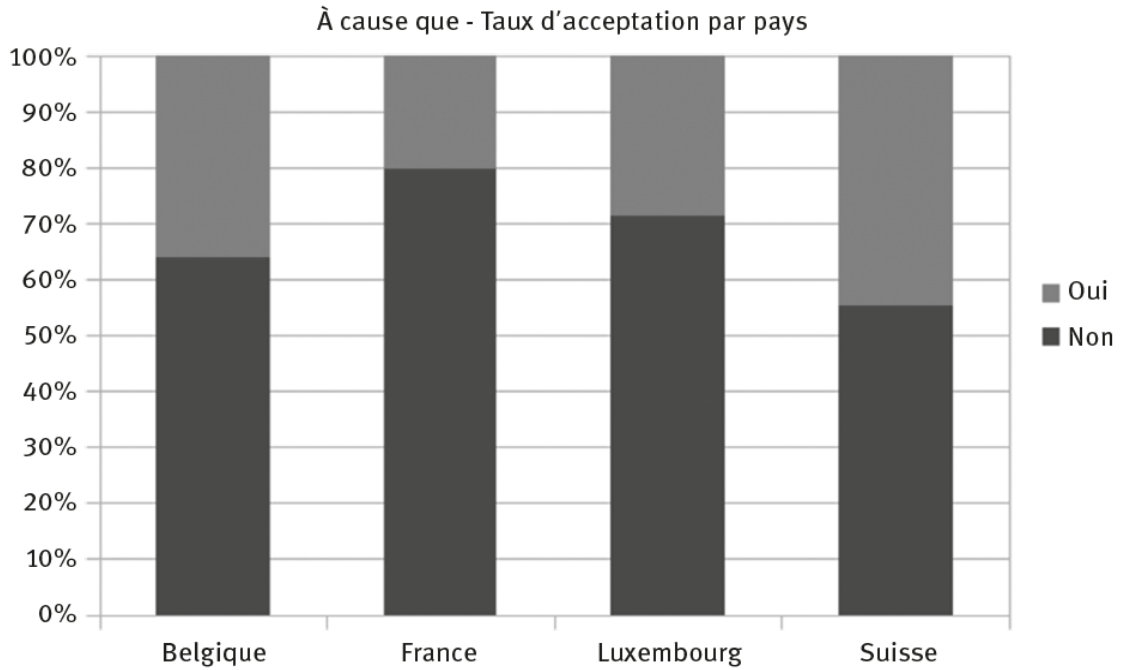
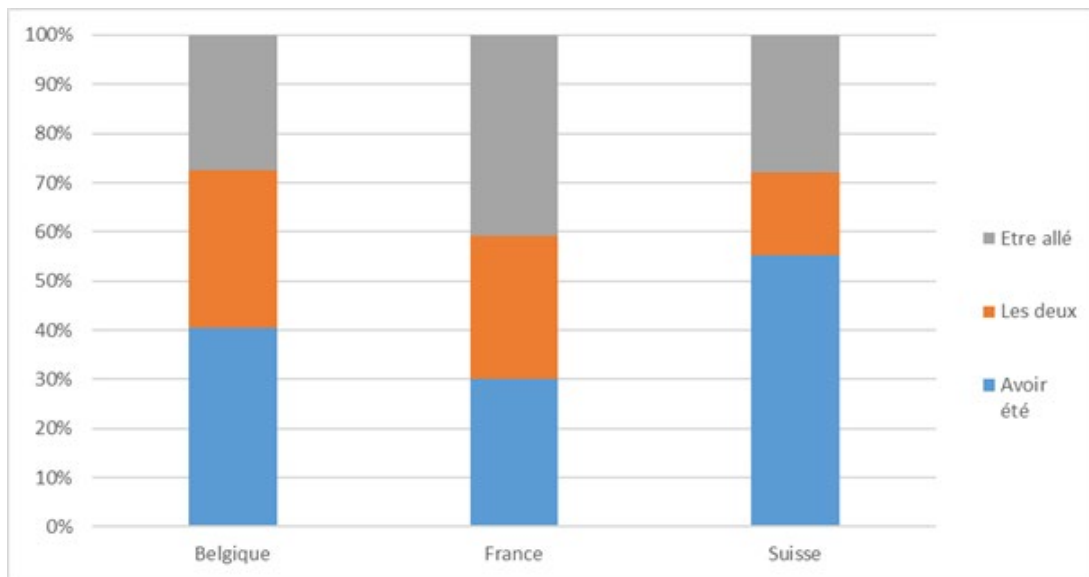


Fig. 2 : Pourcentage d'acceptation des locutions par âge et par pays dans les enquêtes *Le Français de Nos Régions*.

Comprendre le changement linguistique. Autour des usages et des normes : une approche centrée sur le locuteur



(Glikman 2019 [26])



(Glikman et Patard 2022 [35])

Ce rapport plus libre à la norme dans les français hors de France a été également relevé par d'autres études :

De manière générale, les attitudes sont plutôt conservatrices en France, et plus innovantes hors de France (Remysen 2014). (Siouffi 2020 : 156)

Elle peut s'observer également à travers d'autres phénomènes, comme à travers la question de la féminisation, évoquée par Cerquiglini, plus avancée également hors de France :

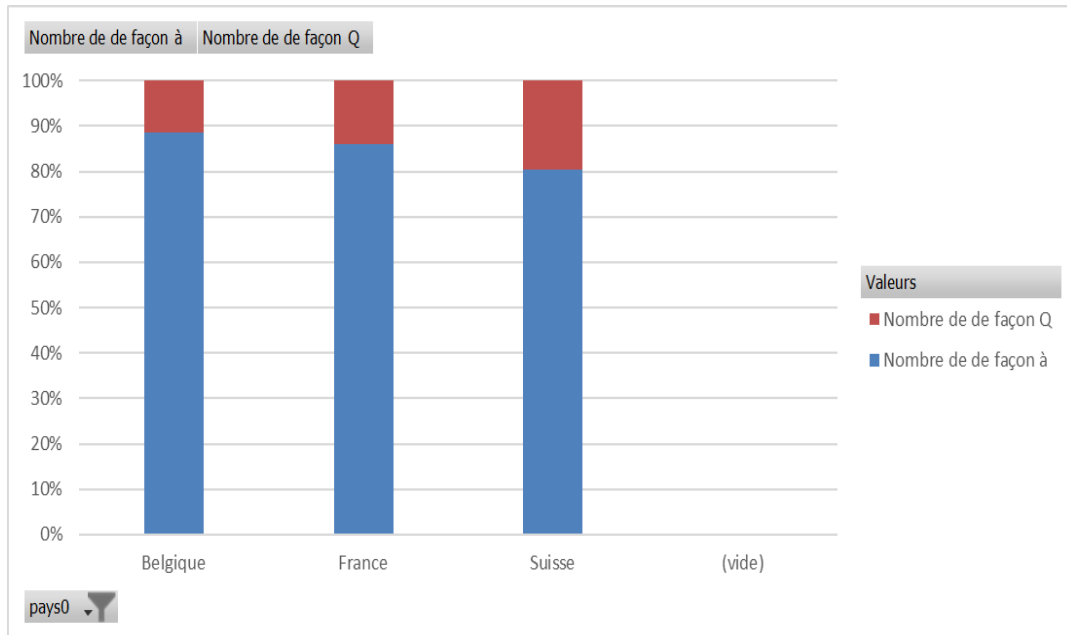
La querelle de la féminisation révèle crûment l'équivoque de notre rapport à cette francophonie. [...] La francophonie, ce sont les autres et leur accent impayable. Or ces « autres » s'accroissent : 220 millions d'humains ont le français pour langue maternelle et l'emploient quotidiennement ; ils rendent désormais les Français minoritaires au sein des usagers. C'est un fait important de l'histoire de la langue, pour ne pas dire un tournant : nouvellement minoritaires, les locuteurs hexagonaux voient leurs représentations mises à mal ; l'arrogance langagière, notamment, est frappée d'obsolescence. Car ces autres, qui plus est, se sentent libres de leur parlure. [...] [on ne se soucie plus guère de ce qui se dit à Paris.]

La féminisation des noms de métiers, titres et fonctions illustre à merveille cette autonomie francophone, et le laborieux deuil français ; car c'est en périphérie que tout a commencé. (Cerquiglini 2018 : 13-15)

Cette différence de représentations entre la France et les pays francophones hors de France se voit également dans l'évolution des discours métalinguistiques. En effet, j'avais déjà remarqué dans l'observation du discours métalinguistique (notamment Glikman 2019 [26]) que les premiers textes admettant dans l'usage des formes préalablement rejetées étaient souvent des grammaires hors de France, comme le *Bon Usage* de Grevisse et Goosse (bien avant l'Académie française, on s'en doutera). Cette intuition s'est vue confirmée par la présentation de Michel Berré, Elisabeth Castadot et Bénédicte Van Gysel lors des journées d'études *Variation(s) en français* à Nancy (novembre 2022), sur l'évolution du discours grammatical en Belgique. Dans leur conférence, ils ont ainsi montré que ce discours, d'abord centré sur une chasse des belgicisms, s'est développé vers une tolérance de plus en plus grande envers la variation.

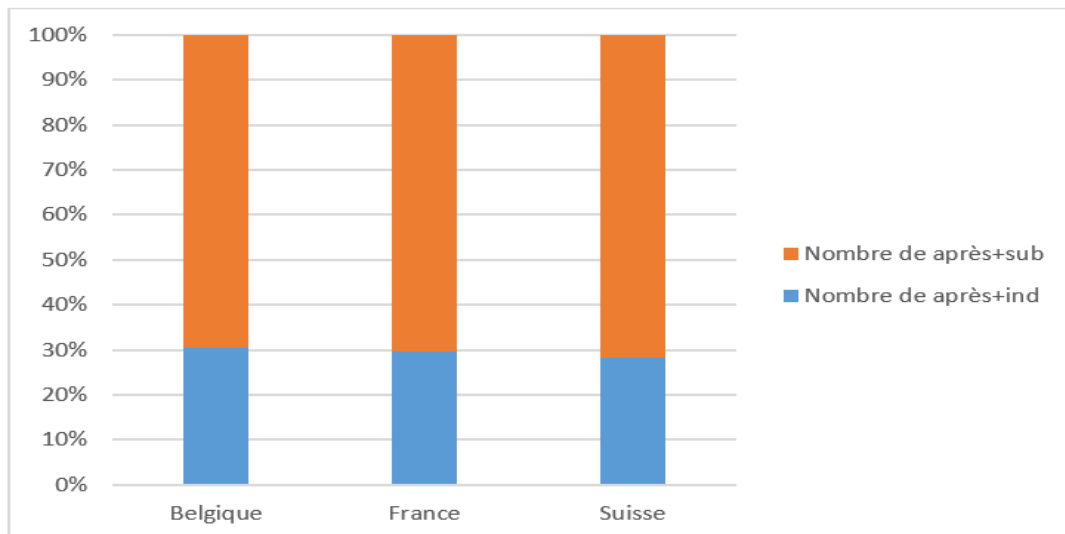
Le jugement des locuteurs peut ainsi être révélateur à la fois de la diffusion des formes et de leur stigmatisation, en lien avec le poids des représentations. Il permet de révéler l'ancrage de ces représentations et le statut accordé aux variantes. À l'inverse, le développement de la locution *de façon à ce que*, que j'ai étudié récemment, montre des représentations inverses. Bien qu'également stigmatisée par la norme, la locution émergente, très peu fréquente dans les corpus au demeurant, et donc probablement assez peu en usage dans des pratiques quotidiennes, n'est pas considérée comme fautive par les locuteurs, qui l'acceptent en grande majorité :

Comprendre le changement linguistique. Autour des usages et des normes : une approche centrée sur le locuteur



(Glikman 2022, *Diachro X*, non publié)

Cela peut ainsi être révélateur du fait que, même s'ils ne l'emploient pas (à en croire la fréquence de l'une ou l'autre variante en corpus), la forme *de façon à ce que* leur paraît plus correcte. Cette représentation peut avoir un lien avec une sorte d'hypercorrection sur la forme « à ce que ». De même, c'est probablement par hypercorrection de l'extension du subjonctif dans les subordonnées que ressort le taux d'acceptation élevé de *après que* + subjonctif, bien étudié par ailleurs, et que j'ai évoqué également dans ma présentation à Nancy lors des Journées Variation(s) :



(Glikman 2022, *Variations*, non publié)

Il est en outre intéressant de voir que pour *malgré que*, dont l'usage est beaucoup plus répandu, le poids de la norme est très présent, tandis que pour *de façon à ce que*, très peu répandu, le jugement normatif est effectivement absent, malgré de fréquentes chroniques linguistiques le fustigeant¹²⁷.

La stigmatisation d'une forme est généralement un élément révélateur du moment où elle émerge, ou en tout cas où elle se diffuse suffisamment pour être remarquée (cf. Barra-Jover 2009). Cependant, comme le précisent Poplack et Dion, ce n'est pas parce que la variante est stigmatisée qu'il s'agit de changement :

La variabilité est largement perçue comme étant inextricablement associée au changement, surtout lorsque l'une des variantes est stigmatisée ou réprouvée. Les sociolinguistes de tradition variationniste reconnaissent qu'il n'en est rien. Ils voient cette association comme une conséquence du « principe de l'âge d'or », que Labov (2001 : 514) a décrit comme la croyance populaire la plus profondément ancrée en matière de langue. (Poplack et Dion 2021 : 83)

Plus encore, la stigmatisation d'une forme par le discours grammatical et le jugement des locuteurs peut au contraire être révélateur de situations différentes. Ainsi, pour *malgré que* et à *cause que*, l'évolution des discours métalinguistiques suit un chemin différent, à *cause que* ayant été la variante recommandée avant de prendre l'étiquette « vieillie » dans la 8^e édition du dictionnaire de l'Académie Française (1932-1935) puis « familier » au cours du 20^e siècle, parallèlement (avec un peu de retard) à son recul dans les usages (entamé depuis le 17^e siècle), tandis que *malgré que*, stigmatisée à ses débuts, commence à être acceptée dans certains discours (Glikman 2019 [26] : 84-89). De même, la concurrence entre *avoir été* et *être allé* a d'abord été considérée comme sémantique, avant d'être qualifiée de familier (Glikman et Patard 2022 [35], voir aussi Steuckardt et Dal Bo 2018). Ainsi, la stigmatisation peut être le signe de différents cas de figure : de l'apparition et de la diffusion d'une forme, mais aussi d'une disparition, entraînant un déplacement du statut de la forme, ou encore d'une variation stable sur un temps long. Les représentations linguistiques et le sentiment linguistique peuvent révéler une zone d'hésitation dans la langue, le degré de diffusion d'une forme. Ils constituent dans tous les cas un marqueur du statut des variantes et de son évolution.

Comme l'ont montré Poplack et Dion (2021), il peut y avoir un changement dans le taux observé d'une variante, sans qu'il s'agisse d'un changement dans la langue, si le

¹²⁷ <https://www.lefigaro.fr/langue-francaise/expressions-francaises/2017/05/27/37003-20170527ARTFIG00003-de-maniere-a-ce-que-ne-faites-plus-la-faute.php>.

conditionnement de la variante ne change pas. De la même manière, le changement se fait lorsqu'il y a changement dans le statut des variantes, même si cela n'est pas encore visible dans les taux d'usage. Inversement, cela permettrait aussi de comprendre les « faux changements » et les variations sur un temps long : tant que la variante ne change pas de statut, la variation reste stable, comme pour la parataxe, dont l'emploi dans le pôle de la proximité communicative (et peut-être également dans des registres oraux) semble finalement assez stable. Ainsi, il me semble important d'inclure à l'étude du changement l'analyse des variations et du changement dans les représentations.

Un autre paramètre important est ressorti de mes travaux récents, qui concerne également le rôle du locuteur dans le changement linguistique. Il s'agit de ce que j'ai appelé la préférence ou construction préférée. Cette notion de préférence dans la langue se situe à un autre niveau que la grammaire et même, il me semble, que les notions de prestige et de représentations. En effet, dans mes recherches sur les conjonctives en fonction sujet (non publiées, → 2.3.), je m'étais rendue compte que la construction, bien que totalement grammaticale, n'était que peu réalisée en corpus, donc peu choisie. Cette notion de préférence, que j'avais déjà touché du doigt dans ma thèse avec la question des contextes restreints de réalisation de la parataxe, et que je retrouve dans mes derniers travaux en cours sur les complétives en fonction sujet, est centrale : si la forme est possible dans la langue, dans la grammaire, et s'il n'existe aucune contrainte de distribution, pourquoi, alors, une structure se réalise-t-elle ou non ? Cela nous ramène à la question fondamentale de ce chapitre sur normes, usages et perception : quel rapport le locuteur entretient-il avec les structures de la langue ? On pourrait en effet également analyser en termes de préférence le développement des tournures redondantes en *le fait que* (utilisées justement préférentiellement dans le cas des complétives sujet), ou la préférence pour les tournures à l'infinitif plutôt que conjonctive (par exemples avec *après que* vs *après* + infinitif). Dans ce dernier cas, le développement de stratégie d'évitement permet évidemment de contourner des zones de flou et, probablement, d'insécurité linguistique. Les stratégies d'évitement se mettent cependant en place suivant un même schéma, vers un marquage explicite et redondant. Cette notion de préférence pourrait ainsi, semble-t-il, être rapprochée des notions de vecteur et de direction du changement proposées par Labov à propos de la transmission du changement, mais pourrait aussi s'appliquer à la diffusion des changements à travers la notion de préférence :

Comment se fait-il que les enfants n'imitent pas leurs parents ou leurs aînés, qu'ils produisent des formes différentes - non pas simplement différentes, mais **différentes de manière prévisible** ? Ce qui est transmis n'est donc pas une forme, mais un vecteur, ou, autrement dit, une instruction de modifier dans une certaine direction » (Labov 1992 : 22)

La question de la direction des changements était déjà directement mise en relation avec l'évaluation du locuteur sur les formes dans le modèle de Weinreich, Labov et Herzog :

These investigations have described changes that continue in the same direction over several generations. Persistence in the direction of change suggests that these changes are variables which have been evaluated in the same way by the speech community over a considerable period of time) (Weinreich, Labov et Herzog 1968 : 146)

La répartition des variantes, qui peut rester stable sur un temps long, peut ainsi ne pas être un changement dans la grammaire, comme on peut le voir avec l'exemple de la parataxe, toujours possible aujourd'hui, voire revenant sur le devant de la scène comme le montre le récent article de presse qui la fustige¹²⁸. Dans le cas de la parataxe, elle est toutefois révélatrice d'une évolution dans les préférences, ou du moins dans les usages de l'écrit tels qu'ils se mettent en place à la fin de la période médiévale en opérant une distance de plus en plus grande avec les pratiques orales. J'ai montré dans Glikman (2013 [11]) que la « disparition » de la parataxe et le développement de la préférence pour la version explicitement marquée va de pair avec le développement de la répétition du subordonnant dans la coordination (cf. Marcotte 1997 ; Graeme-Ritchie 1907) :

L'absence de subordonnant semble ainsi connaître une évolution inverse de celle de la répétition de *que* en coordination de subordonnées et dans les phrases longues, ce qui montre une tendance générale à l'explicitation, davantage marquée en prose, à travers la systématisation de la réalisation du morphème subordonnant.

Pour vérifier cette hypothèse, nous avons procédé à un test comparatif sur *La Chanson de Roland*, chanson de geste en vers du 11^e siècle (abrégé *Roland*), texte connu par ailleurs pour le nombre important de cas d'absence de subordonnant qu'il contient, et *La Conquête de Constantinople* de Robert de Clari, récit historique en prose du 13^e siècle (abrégé *Clari*). Ce test a été réalisé sur une version préliminaire du *Syntactic Reference Corpus of Medieval French* (Prévost/Stein à par.), corpus ayant été annoté syntaxiquement. Ceci nous a permis de relever de manière automatique le nombre de propositions subordonnées (soit de type complétif, pouvant être sujet, objet ou régime, dans la terminologie du projet d'annotation, soit de type relatif), et parmi celles-ci le nombre de conjonctions exprimées, et d'effectuer le même test pour les subordonnées entrant dans une coordination. Les résultats préliminaires de ce test montrent qu'il y a bien évolution, tant dans la diminution de l'absence de subordonnant dans l'ensemble des subordonnées (avec un passage de 84% de subordonnées avec subordonnant dans *Roland*, pour 96% dans *Clari*), que dans l'augmentation de la répétition du subordonnant dans les coordinations: tandis que dans *Roland*, la répétition du subordonnant dans les

¹²⁸ <https://www.leparisien.fr/societe/javoue-je-sais-pas-jai-limpression-il-va-faire-chaud-mais-ou-est-passe-le-que-dans-le-langage-parle-04-09-2022-OFXVNUKI7BCTTIF5VMRFSJK4C4.php>

subordonnées coordonnées ne dépasse pas 40% des cas, dans *Clari*, elle est de plus de 80%. (Glikman 2013 [11] : 11-12)

Ce genre de parallèle gagnerait à être étudié à plus large échelle pour mieux comprendre cette tendance générale. J'ai également évoqué dans cet article le fait que le développement des usages de *que* peut être mis en parallèle avec son développement dans des contextes où il n'est pas attendu, là encore sorte de redondance, comme dans les subordonnées inverses (« Mais le trépidant trio pourrait aussi bien se faire courser dans la forêt de Brocéliande *qu'*on s'en ficherait totalement. » [Télérama, juillet 2009] cf. Corminboeuf 2009 : 267 et Béguelin & Corminboeuf 2005, dont est tiré l'exemple), dans les incises de discours rapporté (« évidemment *qu'*elle a dit l'infirmière » [VALIBEL, bfaSJ1] (cité par Avanzi 2011)), dans les constructions à adverbe en tête (« heureusement *que...* », « peut-être *que...* ») et les formes dites « populaires » de l'interrogation (« où *qu'*il est ? ») (cf. Muller (1997 : 19)), voire avec le développement des emplois de *que* d'une manière générale (cf. Deulofeu 1999). On peut également comparer cette tendance à un marquage plus explicite avec le développement et la multiplication des conjonctions de subordination en moyen français (Verjans 2009), qui semble aller de pair avec le développement de la prose, mais aussi avec le développement du marquage explicite du complément du nom, signe que cette tendance ne touche pas que les composantes phrastiques mais également des niveaux « inférieurs » de lien syntaxique.

Dans ce que j'ai pu observer, cette tendance au développement du marquage explicite est encore actuelle dans la langue et se poursuit dans des périodes plus récentes par le développement de structures redondantes évoqué *supra*, comme pour *de façon que* remplacé par *de façon à ce que*¹²⁹, *malgré que* toujours rejeté par la norme au profit de *malgré le fait que*, de même que *à cause que* pour *à cause du fait que* (Glikman 2019 [26]). Le même mouvement peut être observé dans les propositions complétives en position sujet du type « que tu me lises me fait plaisir » pour lesquelles on préfère « le fait que tu me lises... », sorte de double nominalisation¹³⁰. Parallèlement à cela, j'ai également observé (notamment dans Glikman 2020a [29] et Glikman et Bouard 2021 [31]) que dès qu'une construction infinitive concurrente est présente, elle est majoritaire, voire très largement majoritaire, comme je l'ai montré pour les

¹²⁹ Communication orale Glikman J. (2022) « Variation et mécanismes d'analogie dans l'évolution des locutions conjonctives formées sur *façon*, *manière* et *sorte* », Colloque international *DIACHRO X Le français en diachronie*, 31 mai au 2 juin 2022, Sorbonne Université.

¹³⁰ Communication orale Glikman J. (2019) « Les propositions complétives sujet dans l'histoire du français » *DIACHRO IX · le français en diachronie*, Université de Salamanque (Espagne), les 28, 29 et 30 mars 2019, Facultad de Filología, Palacio de Anaya, Plaza de Anaya, Salamanca.

subordonnées finales, ici *pour* + infinitif dominant en outre l'ensemble des conjonctions possibles :

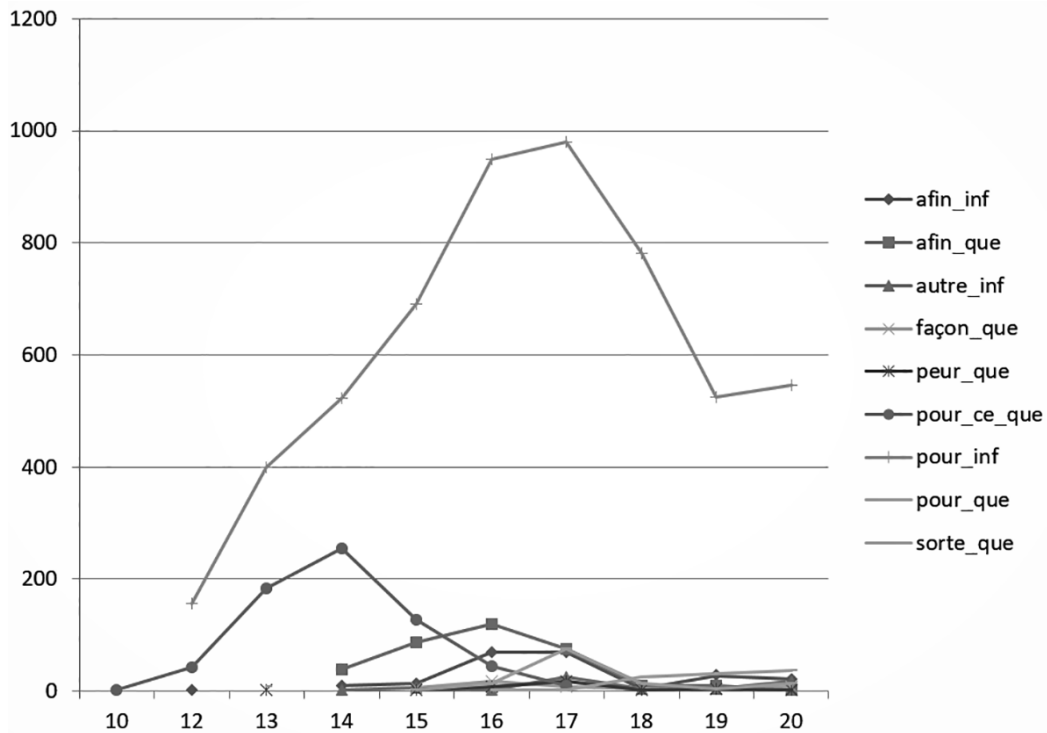


Tableau 2 : les différentes locutions conjonctives et leur équivalent en constructions infinitives (en nombre d'occurrences par siècle) (Glikman et Bouard 2021 : 157)

La construction infinitive étant, comme nous l'avons évoqué, non introduite par un marquage explicite, la question se pose alors de savoir si cette préférence pour les constructions infinitives est à considérer comme un mouvement contradictoire à la préférence pour un marquage explicite, ou si la présence d'une forme verbale infinitive suffit à constituer un marquage explicite. La forme infinitive est en effet considérée comme une marque forte d'intégration (cf. Lehmann 1988). La préférence pour les infinitives doit également être considérée dans l'optique de la tendance naturelle à la recherche d'invariabilité et de simplicité dans la langue (à l'œuvre dans de nombreux changements, cf. Frei 1929), la forme infinitive étant invariable, et évitant ainsi l'incertitude de la conjugaison et de l'emploi du mode pour le locuteur. Une exception à cette prévalence des infinitives est cependant ressortie dans mes travaux préparatoires à une étude sur *après que*, que je souhaite développer prochainement. Une rapide recherche dans le corpus de la GGHF 17-20 sur *après* montre bien la même tendance : la séquence *après* + infinitif apparaît trois fois plus (fréquence brute) que la séquence de la locution conjonctive *après que*. Cependant, pour *avant* (le développement du subjonctif avec *après que* étant souvent présenté comme analogique à la présence du subjonctif dans *avant que*), on ne retrouve pas

cette prévalence de l'infinitive, la proportion de *avant (que) (de) + inf.* et *avant que* étant presque identique.

L'idée selon laquelle une construction peut être à un moment donné toujours grammaticale, mais non « préférée », peut également être rapprochée de la notion de productivité, en particulier en syntaxe : un élément peut être grammatical sans être productif (« la fréquence d'une variante ne reflète pas nécessairement sa productivité. », Poplack et Dion 2021 : 111). Cette notion de productivité, étudiée notamment par Peter Lauwers dans le cadre d'un projet de recherche en cours¹³¹, peut également fournir des perspectives de recherche intéressantes pour l'évaluation des choix des locuteurs.

4.3. Qu'est-ce que cela nous apprend sur le changement linguistique ?

*Même si j'troque mes Rangos
Pour des Mephistos
J'kifferai toujours une guitare disto*

Svinkels, *Réveille le punk*, 1999.

Après avoir évoqué ce que l'analyse de la variation nous apprend sur la caractérisation des discours (→ 4.1.) et sur l'évolution des représentations (→ 4.2.), je souhaite faire ici un rapide bilan sur la place de la variation dans la compréhension du changement linguistique. Là aussi, je repartirai de la question du rôle du locuteur dans le changement linguistique. Comme l'annonce Alain Rey en ouverture de son article « Usages, jugements et prescriptions linguistiques » : « Aucune langue n'échappe à ceux qui l'utilisent. » (Rey 1972 : 4). La prise en compte du locuteur est, de fait, fondamentale dans l'analyse du changement linguistique : les langues changent parce que des gens les parlent. Le rôle du locuteur interroge en outre plusieurs facettes du changement. On pourrait commencer par évoquer la question des causes du changement, depuis les facteurs historiques et sociaux, traditionnellement traités en histoire externe, aux besoins de la langue (Frei 1929), dont le besoin d'expressivité (Meillet 1912). Ce dernier point est bien connu comme moteur du changement :

Dans les communautés linguistiques, on observe souvent la concurrence de deux dynamiques contraires : une dynamique de conformité et une dynamique de différenciation (Siouffi 2016). Les deux vont parfois de pair pour créer de micro-parlers. (Siouffi 2020 : 154)

¹³¹ <https://research.flw.ugent.be/en/projects/language-productivity-work>.

Pour Siouffi (2020), le « sujet parlant » constitue même justement le croisement où s'articulent histoire externe et histoire interne, le point de passage du fait individuel au fait collectif, et l'étude du changement passe par l'analyse de la conscience que le locuteur a du changement :

Peut-on faire une histoire interne des langues sans prendre en compte la conscience même de ce qu'est la langue par le sujet ? Saussure considérait « la collectivité sociale et ses lois comme un de ses éléments internes [au système de signes] et non externes » (Saussure 2002 : 290), puisque, selon lui, le système de signes n'a vocation à avoir du sens que dans la collectivité, parfois appelée « masse parlante », ou communauté des locuteurs. Cette déclaration permet de comprendre sous un autre angle la distinction diachronie / synchronie telle que Saussure l'a proposée, et telle qu'on l'utilise encore majoritairement. Si, en synchronie, « ce qui est réel », pour Saussure, « c'est ce dont les sujets parlants ont conscience à un degré quelconque ; tout ce dont ils ont conscience et rien que ce dont ils peuvent avoir conscience » (Saussure 2002 : 183), en diachronie, dès lors, **on pourra choisir d'articuler les changements à la conscience, ou au sentiment que ce sujet a des usages. [...] La compréhension du fait variationnel permet alors de passer, par la considération de paramètres traditionnellement considérés comme externes, de l'échelle réduite du sujet à celle plus large de communautés ou d'usages.** (Siouffi 2020 : 68-69, nous soulignons)

Le rôle du locuteur amène aussi à interroger le lieu du changement, entre innovation, diffusion et aboutissement du changement :

Dans ce cas, il est patent que l'intention communicative des locuteurs ne saurait être confondue avec une quelconque intention de modifier la langue, ni même avec une « connaissance » systématique de celle-ci. Le caractère local et discursif du changement apparaît particulièrement bien, et la question de savoir quel est le lieu du changement (Verleyen 2008 : 471) particulièrement épineuse. On pourra, alors, être intéressé par les propositions de Keller (1994), qui prolongent celles de Coseriu, et qui mettent en avant, entre le locuteur et le système, les actes de communication concrets. « *The changes of tomorrow are the consequences of our acts of communication today* » (Keller 1994 : 14). (Siouffi, Steuckardt et Wionet 2018: 433-434)

À ces questions s'ajoutent celles du rythme du changement, et de comment intégrer, dans une théorie du changement linguistique, les variations sur un temps long, ce qui change ou ne change pas, ou encore les effets de mode, plus ou moins temporaires (Siouffi 2016) : faut-il les qualifier de « faux changements » (Siouffi, Steuckardt et Wionet 2018) ou de changements « mal compris » (Stark et Dufter 2007) ? Le changement se situe-t-il dans l'innovation, dans la diffusion, dans la disparition d'une variante, dont on sait par ailleurs qu'elle n'arrive pas toujours ? Peut-on toujours parler de changement quand aucune des variantes ne disparaît mais que la variation s'installe ? Comme le rappellent Siouffi, Steuckardt et Wionet (2018), tout dépend de la définition du changement qu'on adopte :

Pour reprendre la question posée sous forme de plaisanterie par notre titre, ce que nous avons essayé de montrer, c'est que l'étude de certains phénomènes diachroniques liés à des zones spécifiques de la langue posait des questions à la théorie actuelle du changement. S'agissant des modalisateurs, peut-on parler de changement linguistique ? Oui, sans doute si on donne à *changement* la définition très englobante que propose par exemple Shapiro : « une innovation qui a cessé d'être un trait individuel et est devenue un fait social » (Shapiro 1991 : 11-13 ; nous traduisons). Mais non si on associe au changement le modèle de la grammaticalisation. Et pas, non plus, si on conserve la définition que donne Posner du change : « *completion of an innovatory process, involving loss of older variants* » (Posner 1997 : 126). Dans le cas des modalisateurs, même si une étude systémique (voir Dostie 2004) peut avoir son intérêt pour le linguiste, il nous semble que, au plan de la diachronie, il faut accepter de renoncer à la variation, et regarder en face le phénomène de la mode, autrement dit du marquage privilégié d'une forme plutôt qu'une autre qui ne serait plus nécessairement à voir en concurrence avec d'autres, mais qui, tout en coexistant avec d'autres formes voisines, occupe momentanément le devant de la scène – au moins dans certains espaces linguistiques. Ce phénomène de circulation du marquage signale peut-être, dans la zone spécifique qu'est la modalisation du discours, l'existence de modifications des usages qui ne relèvent pas réellement du *changement*.

Plus généralement, l'exemple des modalisateurs met bien évidence la difficulté de conjuguer une perspective évolutionniste ou déterministe et une perspective émergentiste sur le langage. Pour Badiou-Monferran et Verjans (2015), la position émergentiste se caractérise par la postulation (i) d'une discontinuité entre l'état de choses initial et l'état de choses immédiatement subséquent, (ii) d'un accroissement de la complexité. Les éditeurs de ce collectif sur les « disparitions » soulignent également que l'appréhension du changement linguistique par le biais des disparitions fait mieux apparaître la question de l'imprédictibilité des nouvelles structures et celle du saut qualitatif entre les étapes. Une approche émergentiste du changement linguistique serait peut-être mieux à même d'intégrer en son sein l'étude des faits de parole à la mode en ce qu'elle ne suppose pas la nécessité d'une continuité de processus, comme le fait la théorie de la grammaticalisation, privilégiant le simple ajout d'une complexité (ici l'ajout d'un fonctionnement « en modalisateur » aux fonctionnements antérieurs), que cette complexité soit amenée à s'installer dans la langue ou pas. (Siouffi, Steuckardt et Wionet 2018 : 436-37)

Trois points me semblent importants à retenir des résultats que j'ai dégagés de mes travaux jusqu'ici. Premièrement, il me semble que le lieu du changement se situe aussi dans le changement de statut des variantes et que l'étude de ce changement de statut contribue à notre compréhension du changement linguistique. Tout changement commence ainsi par un changement de statut des formes, y compris la routinisation et la perte du caractère expressif des formes, à l'origine de leur renouvellement. Par changement de statut, j'inclus la variation dans les représentations, mais aussi dans les contextes d'emploi et les paramètres de variation. Deuxièmement, la notion de préférence, distincte de la grammaire, pourrait constituer le point de passage entre la contrainte et le choix. En cela, on gagnerait également à intégrer son étude dans les approches du changement intégrant le rôle du locuteur. La prise en compte de la

variation constitue à elle seule un pas vers la recherche de ce qui relève du style, de la grammaire, ou d'un autre paramètre de variation, parmi lesquels la préférence peut constituer une piste d'explication. Enfin, dans ce dernier chapitre encore, j'ai essayé de mettre en avant l'importance de la prise en compte des différents paramètres, comme un faisceau d'indices, dont l'étude doit être menée conjointement, comme le conclut Lene Schøsler :

Cela m'amène à une observation conclusive de nature méthodologique sur la variation selon les paramètres diasystématiques, qui pourrait contribuer à expliquer les cas d'absence d'oppositions claires entre l'ACR [anticausative réflexive] et l'AC [anticausative non marquée]. J'ai la conviction que l'étude systématique de la distribution, selon les variables temps, espace, niveau social, style et médium, combinée avec l'étude de la propagation des changements selon la théorie d'*actualisation* de H. Andersen (2001a-c), permet de mieux comprendre le processus des changements en cours. J'ai esquissé une telle piste *supra*, en me référant aux nouvelles alternances causatives du type *on divorce la femme – la femme divorce*, cf. (9b), provenant de l'étude sur le français avancé de M. Krötsch et W. Oesterreicher (2002). À cela s'ajoute, si possible, l'utilisation de métadonnées, par exemple les textes des Remarqueurs de l'époque classique et de l'époque des Lumières. Par contre, je n'ai pas ici la possibilité de procéder à une étude extra-linguistique, à savoir la possible influence de l'italien, pour comprendre les causes de l'essor de la structure ACR après 1500. (Schøsler 2021 : 116)

L'analyse en faisceau est ainsi importante à tous les niveaux : pour l'établissement du lien syntaxique, qui permet de poser deux réalisations comme variantes syntaxiques, tout comme pour l'analyse de la variation. Pour mener à bien cette dernière, les différents paramètres dia-variationnels doivent être pris en considération. Dans ces différents paramètres, j'inclus le genre, le code, la distance communicative, la visée, la programmation et l'adaptation au contexte. Ces différents facteurs influent sur la production, mais aussi sur la manière dont elle est perçue. Cette perception de l'objet langagier produit influera à son tour sur la manière dont les locuteurs voudront l'imiter ou au contraire, s'en distinguer.

**En guise d'ouverture : transmission et
diffusion**

En guise d'ouverture : transmission et diffusion

*Given in the light
Written on the pages
Is the answer to a neverending story*

Limahl, *Never Ending Story*, 1984.

Cette citation, à l'heure d'apporter les derniers mots de ce mémoire, est vraiment représentative de ce que doit constituer ce moment de clôture : un rappel du fait que ce n'est pas une fin¹³². Au-delà de l'exercice formel, j'ai encore pris conscience de nouveaux éléments et certains thèmes ont émergé lors de la rédaction de cette synthèse, et c'est sur cela que je voudrais mettre l'accent pour finir. De même que les langues continuent d'évoluer, la posture de chercheur est elle-même en constante évolution. On peut toujours continuer à approfondir, apprendre, découvrir, évoluer et, surtout, pour emprunter le titre de Labov (2007), transmettre et diffuser¹³³. La transmission passe évidemment par nos missions d'enseignement et d'encadrement. Là encore, pour poursuivre la transposition, peut-être que l'on peut retenir parmi les pistes avancées par Labov (1992) l'idée selon laquelle notre rôle est de transmettre non pas des objets de recherche, mais des directions vers lesquelles chercher. La diffusion, elle, se fait non seulement dans l'animation de la recherche, le partage des données et des produits de la recherche, mais doit aussi se mener à travers un retour envers le grand public.

En guise de conclusion, je reviendrai ici brièvement sur les points qui me semblent importants à retenir et les perspectives que cela ouvre, puisqu'il ne s'agit pas d'une fin. En réfléchissant à l'évolution de mes idées, il me semble que l'une des évolutions se situe notamment dans ma conception des domaines. Dans ma thèse, j'avais insisté sur une représentation indépendante des domaines (syntaxe, sémantique, etc.). Aujourd'hui je continue de considérer qu'il existe un avantage, en termes de conceptualisation, à définir des unités minimales et maximales propres à chaque domaine et à établir des tests adaptés pour chacun d'eux. Cependant, dans l'analyse du fait langagier, tout est lié : lexique, syntaxe, discours, contraintes et choix liés à la situation, effet pragmatique, etc., et tout doit être pris en compte.

Derrière les concepts partagés (variation, variante, variétés...), parfois avec certaines incompréhensions, il me semble que le point de rencontre entre la sociolinguistique et la

¹³² Comme me le rappelait Dominique Lagorgette, qui le tenait elle-même de Diane Vincent, un mémoire, ça ne se termine pas, ça s'arrête.

¹³³ Utilisés par Labov selon l'idée que la transmission se fait de manière « verticale » (i.e. des parents vers les enfants) alors que la diffusion se fait de manière « horizontale » (d'un groupe linguistique à un autre).

linguistique variationnelle est encore plus important avec la prise en compte du sujet parlant. Ma position, en plaçant ainsi l'étude des représentations linguistiques au centre des analyses du changement linguistique, me permet de faire le pont entre les deux approches. De la même manière, il me semble que le rôle du diachronicien peut aussi être de faire le pont entre le passé et le présent et de faire s'enrichir les approches diachroniques et synchroniques. Indépendamment de mon envie de « rendre moderne » (pour ne pas dire « sexy ») les études diachroniques, il me semble que, dans notre société actuelle, autour de la cristallisation sur une idée du français idéalisée, sur les débats de société tels que la féminisation ou l'orthographe, on peut se demander quelle est la place pour le linguiste qui observe et pour le diachronicien dont la perspective est plus large. Peut-être cette place justement, celle de mettre de la perspective dans les représentations, comme le peintre sur sa toile.

Bibliographie

Bibliographie

- Auer, P. (1998). Zwischen Parataxe und Hypotaxe : ‘Abhängige Hauptsätze’ im gesprochenen und geschriebenen Deutsch. *Zeitschrift für Germanistische Linguistik*, 26(3), 284-307. <https://doi.org/10.1515/zfgl.1998.26.3.284>
- Avanzi, M. (2009 [2012]). La prosodie des verbes parenthétiques en français parlé. *Linx*, 61, 131-144. <https://doi.org/10.4000/linx.1344>
- Avanzi, M. (2011). *Prosodie des constructions extraposées et asyndétiques en français parlé* [Thèse de doctorat]. Université de Neuchâtel, Université de Paris Oest/Nanterre.
- Avanzi, M. (2012). *L’interface prosodie/syntaxe en français : Dislocations, incises et asyndètes*. Peter Lang.
- Avanzi, M. (2017). *Atlas du français de nos régions*. Armand Colin.
- Avanzi, M. (2019a). *Parlez-vous (les) français ? Atlas des expressions de nos régions*. Armand Colin.
- Avanzi, M. (2019b). Cartographier les régionalismes du français en Suisse romande et en France voisine à l’ère des sciences participatives. *Cahiers internationaux de sociolinguistique*, 14(2), 43-103. <https://doi.org/10.3917/cisl.1802.0043>
- Avanzi, M. (2020). *Comme on dit chez nous : Le grand livre du français de nos régions*. Le Robert.
- Ayres-Bennett, W. (2004). *Sociolinguistic variation in seventeenth-century France : Methodology and case studies*. Cambridge University Press.
- Ayres-Bennett, W. (éd.). (2018). *Nouvelles voies d’accès au changement linguistique*. Classiques Garnier.
- Ayres-Bennett, W. & Bellamy, J. (éds). (2021). *The Cambridge handbook of language standardization*. Cambridge University Press.
- Ayres-Bennett, W. & Rainsford, T. M. (éds). (2014). *L’histoire du français : État des lieux et perspectives*. Classiques Garnier.
- Ayres-Bennett, W. & Seijido, M. (2013). *Bon usage et variation sociolinguistique : Perspectives diachroniques et traditions nationales*. ENS éditions.
- Badiou-Monferran, C. & Verjans, T. (éds). (2015). *Disparitions : Contributions à l’étude du changement linguistique*. Honoré Champion.
- Bailly, A. (1895). *Dictionnaire grec-français*. Hachette.
- Banniard, M. (1997). *Du latin aux langues romanes*. Nathan.
- Barra Jover, M. (2000). Constatation et induction face aux corpus diachroniques : Le problème du n⁺¹ texte. *Cahiers FoReLLIS*, 14, 7-20.
- Barra Jover, M. (2009). Des variantes invisibles à la fragmentation des langues romanes. *Recherches linguistiques de Vincennes*, 38, 105-137. <https://doi.org/10.4000/rlv.1784>
- Béguelin, M.-J. & Corminboeuf, G. (2005). De la question à l’hypothèse : Aspects d’un phénomène de coalescence. In C. Rossari, A. Beaulieu, A. Cojocariu & A. Razgouliaeva (éds), *Les États de la question* (p. 67-89). Nota Bene.

- Béguelin, M.-J., Corminboeuf, G. & Avanzi, M. (2011). *La Parataxe. Tome 1. Entre dépendance et intégration*. Peter Lang.
- Béguelin, M.-J., Corminboeuf, G. & Lefeuvre, F. (éds). (2019). *Types d'unités et procédures de segmentation*. Lambert Lucas.
- Benzitoun, C. (2006a). *Description morphosyntaxique du mot quand en français contemporain* [Thèse de doctorat]. Université de Provence, Aix-Marseille.
- Benzitoun, C. (2006b). Examen de la notion de “subordination” Le cas des *quand* “insubordonnés”. *Faits de Langues*, 28(1), 35-46. <https://doi.org/10.1163/19589514-028-01-900000005>
- Benzitoun, C., Dister, A., Gerdes, K., Kahane, S., Pietrandrea, P., Sabio, F. & Debaisieux, J.-M. (2010). *tu veux couper là faut dire pourquoi*. Propositions pour une segmentation syntaxique du français parlé. *2^e Congrès Mondial de Linguistique Française*, 2075-2090. <https://doi.org/10.1051/cmlf/2010201>
- Benzitoun, C. & Sabio, F. (2010). Où finit la phrase ? Où commence le texte ? L'exemple des regroupements de constructions verbales. *Discours*, 7, 3-25. <https://doi.org/10.4000/discours.7966>
- Bertin, A. (1997). *L'expression de la cause en ancien français*. Droz.
- Blanche-Benveniste, C. (1989). Constructions verbales “en incisives” et rection faible des verbes. *Recherches sur le français parlé*, 9, 53-73.
- Blanche-Benveniste, C., Bilger, M., Rouget, C. & van den Eynde, K. (1990). *Le français parlé : Études grammaticales*. CNRS éditions.
- Blanche-Benveniste, C. & Willems, D. (2007). Un nouveau regard sur les verbes “faibles”. *Bulletin de la Société de Linguistique de Paris*, 102(1), 217-254. <https://doi.org/10.2143/BSL.102.1.2028205>
- Blasco, M. & Bodelot, C. (2018). Langue parlée/langue écrite, du latin au français : Un clivage dans l'histoire de la langue? *Langages*, 208(4), 5-16. <https://doi.org/10.3917/lang.208.0005>
- Bouard, B. (2007). *Structure de la proposition et construction verbale : Régime, complément et transitivité dans les grammaires françaises, 1651-1863* [Thèse de doctorat]. Paris 7.
- Bringer, C. (2019). *Lemmatisation et annotation des formes verbales du français médiéval* [Mémoire de maîtrise]. Université de Strasbourg.
- Buridant, C. (2019). *Grammaire du français médiéval : XI^e-XIV^e siècles*. ELiPhi.
- Canut, C. (1998). Pour une analyse des productions épilinguistiques. *Cahiers de praxématique*, 31, 69-90. <https://doi.org/10.4000/praxématique.1230>
- Capin, D. (2019). Présence du démonstratif dans les contextes subordonnants en français médiéval. In D. Capin, V. Obry, J. Glikman & T. Revol (éds), *Le français en diachronie* (p. 87-106). ELiPhi.
- Carlier, A. & Guillot-Barbance, C. (éds). (2018). *Latin tardif, français ancien : Continuités et ruptures*. De Gruyter. <https://doi.org/10.1515/9783110551716>
- Carlier, A. & Prévost, S. (2021). Constructions, constructionnalisation et changement linguistique. Présentation. *Langue française*, 209(1), 9-22. <https://doi.org/10.3917/lf.209.0009>

- Cerquiglioni, B. (1989). *Éloge de la variante : Histoire critique de la philologie*. Le Seuil.
- Cerquiglioni, B. (2018). *Le ministre est enceinte, ou la grande querelle de la féminisation des noms*. Le Seuil.
- Choi-Jonin, I. (2005). Les subordinations spatio-temporelles sans marque segmentale. In F. Lamermet & H. Nølke (éds), *La syntaxe au cœur de la grammaire. Recueil offert en hommage pour le 60^e anniversaire de Claude Muller* (p. 55-64). Presses Universitaires de Rennes.
- Combettes, B. (1996). L'intégration syntaxique de la subordonnée conjonctive topicalisée en français : Approche historique. In C. Muller (éd.), *Dépendance et intégration syntaxique. Subordination, coordination, connexion* (p. 89-96). De Gruyter.
- Combettes, B. (2017). Facteurs discursifs et contraintes syntaxiques : Aspects diachroniques de la relation de cataphore. *Langue française*, 195(3), 111-130. <https://doi.org/10.3917/lf.195.0111>
- Combettes, B. (2018). La subordination dans la Chronique de Frédégaire : Les propositions non régies. In A. Carlier & C. Guillot-Barbance (éds), *Latin tardif, français ancien : Continuités et ruptures* (p. 323-412). De Gruyter. <https://doi.org/10.1515/9783110551716>
- Conti, V. & Béguelin, M. (2010). Le statut des concessives en avoir beau du français: Considérations synchroniques et diachroniques. *Journal of French Language Studies*, 20(3), 271-288. doi:10.1017/S095926951000027X
- Corminboeuf, G. (2009). *L'expression de l'hypothèse en français : Entre hypotaxe et parataxe*. De Boeck Duculot.
- Corminboeuf, G. (2010). Une description des constructions introduites par un *si* dit austinien. *Travaux de linguistique*, 60(1), 127-144. <https://doi.org/10.3917/tl.060.0127>
- Corpus French Mixed 2012 French mixed corpus based on material from 2012. Leipzig Corpora Collection. Dataset. https://corpora.uni-leipzig.de?corpusId=fra_mixed_2012.
- Coseriu, E. (1998). Le double problème des unités dia-s. *Les Cahiers dia. Études sur la diachronie et la variation linguistique*, 1, 9-16.
- Creissels, D. (1995). *Éléments de syntaxe générale*. PUF.
- Creissels, D. (2006a). *Syntaxe générale, une introduction typologique. Vol. 1. Catégories et constructions*. Hermès Science, Lavoisier.
- Creissels, D. (2006b). *Syntaxe générale, une introduction typologique. Vol. 2. La phrase*. Hermès Science, Lavoisier.
- Culicover, P. W. & Jackendoff, R. (1997). Semantic subordination despite syntactic coordination. *Linguistic Inquiry*, 28, 195-217.
- Culicover, P. W. & Jackendoff, R. (2005). *Simpler syntax*. Oxford University Press.
- Debaisieux, J.-M. (1994). *Le fonctionnement de parce que en français parlé contemporain : Description linguistique et implications didactiques* [Thèse de doctorat]. Université de Nancy II.
- Debaisieux, J.-M. (2007). La distinction entre dépendance grammaticale et dépendance macro-syntaxique comme moyen de résoudre les paradoxes de la subordination. *Faits de langues*, 28, 119-132.

- Debaisieux, J.-M. (2013). *Analyses linguistiques sur corpus*. Hermès Science, Lavoisier.
- Dehé, N. & Wichmann, A. (2010). Sentence-initial *I think (that)* and *I believe (that)* : Prosodic evidence for uses as main clause, comment clause and discourse marker. *Studies in Language*, 34(1), 36-74. <https://doi.org/10.1075/sl.34.1.02deh>
- Deulofeu, H.-J. (1989). Les couplages de constructions verbales en français parlé : Effet de cohésion discursive ou syntaxe de l'énoncé. *Recherches sur le français parlé*, 9, 111-141.
- Deulofeu, H.-J. (1999). Questions de méthode dans la description morphosyntaxique de l'élément *que* en français contemporain. *Recherches sur le français parlé*, 15, 163-198.
- Deulofeu, H.-J. (2001). La notion de construction corrélatrice en français : Typologie et limites. *Recherches sur le français parlé*, 16, 103-124.
- Deulofeu, H.-J., Debaisieux, J.-M. & Martin, P. (2019). Apparent insubordination as discourse patterns in French. In K. Beijering, G. Kaltenböck & M. Sol Sansiñena (éds), *Insubordination Theoretical and empirical issues* (p. 349-383). De Gruyter. <https://hal.archives-ouvertes.fr/hal-03172591>
- Didienne, M. (2019). *Néologismes et enrichissement de la langue française* [Mémoire de maîtrise]. Université de Strasbourg.
- Diewald, G. (2011). Pragmaticalization (defined) as grammaticalization of discourse functions. *Linguistics*, 49(2), 365-390. <https://doi.org/10.1515/ling.2011.011>
- Dostie, G. (2004). *Pragmaticalisation et marqueurs discursifs : Analyse sémantique et traitement lexicographique*. De Boeck Ducleot.
- Dostie, G., Diwersy, S. & Steuckardt, A. (2021). Présentation du numéro. *Linx*, 82, 1-6. <https://doi.org/10.4000/linx.8005>
- Ducrot, O. & Todorov, T. (1972). *Dictionnaire encyclopédique des Sciences du langage*. Le Seuil.
- Dufter, A. (2018). Repenser la 'spatialisation' de la linguistique variationnelle. In M. Glessgen, J. Kabatek & H. Völker (éds), *Repenser la variation linguistique. Actes du Colloque DIA IV à Zurich (12-14 sept. 2016)* (p. 63-73). ELiPhi.
- Dufter, A., Fleischer, J. & Seiler, G. (éds). (2009). *Describing and Modeling Variation in Grammar*. Mouton de Gruyter. <https://doi.org/10.1515/9783110216097>
- Dufter, A., Fleischer, J. & Seiler, G. (2009). Introduction. In A. Dufter, J. Fleischer & G. Seiler (éds), *Describing and Modeling Variation in Grammar* (p. 1-20). Mouton de Gruyter. <https://doi.org/10.1515/9783110216097>
- Dufter, A. & Stark, E. (2003). La variété des variétés : Combien de dimensions pour la description? *Romanistisches Jahrbuch*, 53(1), 81-102. <https://doi.org/10.1515/9783110179590.81>
- Dufter, A. & Stark, E. (2007). La linguistique variationnelle et les changements linguistiques 'mal compris' : Le cas du *ne* de négation. In B. Combettes & C. Marchello-Nizia (éds), *Études sur le changement linguistique en français* (p. 115-128). Presses Universitaires de Nancy.
- Frei, H. (2003). *La grammaire des fautes*. Ennoia. (Original work published 1929)

- Gachet, F. (2010). Les structures temporelles en *à peine* : Évolution diachronique et fonctionnement syntaxique. *2e Congrès Mondial de Linguistique Française*, 207-221. <https://doi.org/10.1051/cmlf/2010126>
- Gadet, F. (1992). Variation et hétérogénéité. *Langages*, 26(108), 5-15. <https://doi.org/10.3406/lgge.1992.1647>
- Gadet, F. (2017). Variatio delectat : Variation et dialinguistique. *Langage et société*, 160-161(2), 75-91. <https://doi.org/10.3917/lc.160.0075>
- Gadet, F. (2018). L'oralité ordinaire à l'épreuve de la mise en écrit : Ce que montre la proximité. *Langages*, 208(4), 113-129. <https://doi.org/10.3917/lang.208.0113>
- Gadet, F. (2020). Langue et variation. In *Encyclopédie grammaticale du français* (2^e éd.). <http://encyclogram.fr>
- Gadet, F. (2021). Variation. *Langage et société*, HSI, 331-336. <https://doi.org/10.3917/lc.hs01.0332>
- Garagnon, A.-M. & Calas, F. (2002). *La phrase complexe : De l'analyse logique à l'analyse structurale*. Hachette.
- Gautier, A. (2006). *Unité et discontinuité : Une approche épistémologique et systématique de la phrase* [Thèse de doctorat]. Sorbonne Université.
- Gérard, C. (2019). Linguistique des genres : Objet et méthode. Statut culturel des genres et variétés génériques. *Linx*, 78, 1-59. <https://doi.org/10.4000/linx.3030>
- Gerdes, K. & Kahane, S. (2006). L'amas verbal au cœur d'une modélisation topologique du français. *Linguisticae Investigationes*, 29(1), 75-89. <https://doi.org/10.1075/li.29.1.07ger>
- Glessgen, M.-D., Kabatek, J. & Völker, H. (éds). (2018). *Repenser la variation linguistique : Actes du Colloque DIA IV à Zurich (12-14 sept. 2016)*. ELiPhi.
- Godard, D. (1992). Le programme labovien et la variation syntaxique. *Langages*, 26(108), 51-65. <https://doi.org/10.3406/lgge.1992.1650>
- Graeme-Ritchie, R. L. (1907). *Recherches sur la syntaxe de la conjonction "que" dans l'ancien français*. Champion.
- Gregov, N. (2022). Analyse critique du traitement de l'apposition dans la grammaticographie contemporaine. In F. Neveu, S. Prévost, A. Steuckardt, G. Bergounioux & B. Hamma (éds), *8^e Congrès Mondial de Linguistique Française*. EDP Sciences. <https://doi.org/10.1051/shsconf/202213803002>
- Gross, M. (1984). Lexicon-Grammar and the Syntactic Analysis of French. *10th International Conference on Computational Linguistics and 22nd Annual Meeting of the Association for Computational Linguistics*, 275-282. <https://doi.org/10.3115/980491.980549>
- Groupe de Fribourg. (2017). *Grammaire de la période*. Peter Lang.
- Guillot, C., Heiden, S., Lavrentiev, A. & Pincemin, B. (2015). L'oral représenté dans un corpus de français médiéval (9^e-15^e) : Approche contrastive et outillée de la variation diasystémique. In K. A. J. Kragh & J. J. Lindschouw (éds), *Les variations diasystémiques et leurs interdépendances dans les langues romanes. Actes du Colloque DIA II à Copenhague (19-21 nov. 2012)* (p. 15-28). ELiPhi.

- Guillot, C., Lavrentiev, A. & Heiden, S. (2013). Oral représenté au Moyen Âge : Vers une définition et une méthodologie d'analyse. In D. Lagorgette & P. Larrivée (éds), *Représentations du sens linguistique 5 : Actes de la conférence internationale organisée à Chambéry, 25-27 mai 2011* (p. 17-41). Presses Universitaires de Savoie.
- Guillot-Barbance, C. (2013). *Diachronie du latin au français médiéval : Pour une approche empirique et contextuelle (corpus, outils et philologie numérique)* [Mémoire de synthèse d'Habilitation à diriger des recherches]. Université de Strasbourg.
- Guillot-Barbance, C., Heiden, S. & Lavrentiev, A. (2017). Base de français médiéval : Une base de référence de sources médiévales ouverte et libre au service de la communauté scientifique. *Diachroniques*, 7, 168-184.
- Guillot-Barbance, C., Pincemin, B. & Lavrentiev, A. (2018). Représentation de l'oral en français médiéval et genres textuels. *Langages*, 208(4), 53-68. <https://doi.org/10.3917/lang.208.0053>
- Kahane, S. & Gerdes, K. (2022). *Syntaxe théorique et formelle. Vol. 1. Modélisation, unités, structures*. Language Science Press. <https://zenodo.org/record/6446067>
- Kahane, S. & Mazziotta, N. (2015). Syntactic Polygraphs. A Formalism Extending Both Constituency and Dependency. *Proceedings of the 14th Meeting on the Mathematics of Language (MoL 2015)*, 152-164. <https://doi.org/10.3115/v1/W15-2313>
- Karas, H. & Shyldkrot, H. B. Z. (2021). Intralingual Diachronic Translation and Transfer: The Case of Old French. *Romance Studies*, 39(4), 189-207, DOI: 10.1080/02639904.2021.2002553.
- Kircher, R. & Zipp, L. (éds). (2022). *Research methods in language attitudes*. Cambridge University Press.
- Koch, P. (2015). Disparition lexicale, variétés linguistiques et traditions discursives. In C. Badiou-Monferran & T. Verjans (éds), *Disparitions : Contributions à l'étude du changement linguistique* (p. 68-88). Honoré Champion.
- Koch, P. & Oesterreicher, W. (2001). Gesprochene Sprache und geschriebene Sprache/Langage parlé et langage écrit. In G. Holtus, M. Metzeltin & C. Schmitt (éds), *Lexikon der Romanistischen Linguistik: Vol. 1(2)* (p. 584-627). Niermeyer.
- Kragh, K. A. J. (2021). Proposition d'une classification des marqueurs discursifs comme membres d'un paradigme. *Langue française*, 209(1), 119-140. <https://doi.org/10.3917/lf.209.0119>
- Kragh, K. A. J. & Schøsler, L. (2021). Analyse constructionnelle et paradigmatique illustrée par le verbe *voir*. *Langue française*, 209(1), 83-100. <https://doi.org/10.3917/lf.209.0083>
- Kuyumcuyan, A. & Theissen, A. (éds). (2020). *La polyvalence de ce en français : De la syntaxe à la cohésion textuelle*. Editura Universităţii din Oradea.
- Kuyumcuyan, A. & Theissen, A. (2020). *Ce*, syntaxe et sémantique : Présentation. *Langue française*, 205(1), 7-12. <https://doi.org/10.3917/lf.205.0007>
- La corrélation en français. Étude morphosyntaxique*. (2015). Classiques Garnier.
- Labov, W. (1972). *Sociolinguistic patterns*. University of Pennsylvania Press.

Bibliographie

- Labov, W. (1976). *Sociolinguistique* (A. Kihm, Trad.). éditions de Minuit. (Original work published 1972)
- Labov, W. (1992). La transmission des changements linguistiques. *Langages*, 26(108), 16-33. <https://doi.org/10.3406/lgge.1992.1648>
- Labov, W. (2007). Transmission and Diffusion. *Language*, 83(2), 344-387. <https://doi.org/10.1353/lan.2007.0082>
- Laferrière, A. (2018). *Les incisives dans les genres narratifs : « certaines formules des plus prometteuses »*. Classiques Garnier.
- Lagorgette, D. (1998). *Désignatifs et termes d'adresse dans quelques textes en moyen français*. [Thèse de doctorat]. Université Paris 10-Nanterre.
- Lagorgette, D. (2002). Autour de la référence : Désignatifs et termes d'adresse en moyen français. *Linx*, 12, 106-117. <https://doi.org/10.4000/linx.1289>
- Lagorgette, D. (2006). Du vocatif à l'apostrophe : Problèmes terminologiques et théoriques, termes d'adresse et détachement en diachronie en français. *L'information grammaticale*, 109(1), 38-44. <https://doi.org/10.3406/igram.2006.3814>
- Lagorgette, D. (2019). Paul Zumthor et la mouvance. In I. Muzart-Fonseca Dos Santos & J.-R. Valette (éds), *Poétiques de Paul Zumthor (1915-2015)* (p. 113-129). Classiques Garnier. https://classiques-garnier.com/doi/garnier?article=JvtMS01_113
- Lagorgette, D. & Larrivée, P. (éds). (2013). *Représentations du sens linguistique 5 : Actes de la conférence internationale organisée à Chambéry, 25-27 mai 2011*. Presses Universitaires de Savoie.
- Lavrentiev, A. (2010). La "phrase" en français médiéval : Une réalité ou une reconstruction artificielle? 2^e Congrès Mondial de Linguistique Française. <https://doi.org/10.1051/cmlf/2010125>
- Le Goffic, P. (1993). *Grammaire de la phrase française*. Hachette supérieur.
- Lefevre, F. (1999). *La phrase averbale en français*. L'Harmattan.
- Lehmann, C. (1988). Towards a typology of clause linkage. In J. Haiman & S. A. Thompson (éds), *Clause combining in grammar and discourse* (p. 181-226). John Benjamins.
- Lehmann, C. (2002). *Thoughts on grammaticalization* (2^e éd.). Seminar für Sprachwissenschaft.
- Lodge, R. A. (1997). *Le français : Histoire d'un dialecte devenu langue*. Fayard.
- Lodge, R. A. (2018). Fondements empiriques d'une théorie du changement linguistique. In P. Blumenthal & D. Vigier (éds), *Études diachroniques du français et perspectives sociétales* (p. 195-214). Peter Lang. <https://doi.org/10.3726/b11459>
- Marchello-Nizia, C. (1978). Un problème de linguistique textuelle : La classe des éléments joncteurs de propositions. In R. Martin (éd.), *Études de syntaxe du moyen-français, Actes du Colloque de Metz* (p. 33-42). Klincksieck.
- Marchello-Nizia, C. (1993). Parataxe et syntaxe du vers en ancien français. In I. Perrin (éd.), *Approches énonciatives de l'énoncé complexe* (p. 63-71). Peeters.
- Marchello-Nizia, C. (1997). *La langue française aux XIV^e et XV^e siècles*. Nathan.

Bibliographie

- Marchello-Nizia, C. (2006). *Grammaticalisation et changement linguistique*. De Boeck & Larrier.
- Marchello-Nizia, C. (2010). *L'évolution du français*. Vigdor. (Original work published 1995)
- Marchello-Nizia, C. (2012). L'oral représenté : Un accès construit à une face cachée des langues 'mortes'. In C. Guillot, B. Combettes, A. Lavrentiev, E. Oppermann-Marsaux & S. Prévost (éds), *Le changement en français : Études de linguistique diachronique* (p. 247-264). Peter Lang.
- Marchello-Nizia, C., Combettes, B., Prévost, S. & Scheer, T. (éds). (2020). *Grande grammaire historique du français* (Vol. 1-2). De Gruyter Mouton.
- Marcotte, S. (1997). *La coordination des propositions subordonnées en moyen français*. Droz.
- Marnette, S. (1999). *Il le vos mande, ge sui qui le vos di* : Les stratégies du dire dans les chansons de geste. *Revue de linguistique romane*, 63, 387-417.
- Marnette, S. (2001). Du discours insolite : Le discours indirect sans *que*. *French studies*, 55(3), 297-313.
- Martin, R. & Wilmet, M. (1980). *Syntaxe du moyen français*. Bière.
- Martinet, A. (1985). *Syntaxe générale*. Armand Colin.
- Mazziotta, N. (2012). Le *Syntactic Reference Corpus of Medieval French* : Structure, outils et exploitation. *11^{es} Journées internationales d'Analyse statistique des Données Textuelles (JADT)*, 701-713. <https://hdl.handle.net/2268/141540>
- Mazziotta, N. (2016). *Représenter la connaissance en linguistique. Observations sur l'édition de matériaux et sur l'analyse syntaxique* [Mémoire de synthèse d'Habilitation à diriger des recherches, Paris Nanterre]. <https://hdl.handle.net/2268/204408>
- Meillet, A. (1912). L'évolution des formes grammaticales. *Scientia*, 6(12), 384-400.
- Muller, C. (1997). *La subordination en français : Le schème corrélatif*. Armand Colin.
- Panckhurst, R., Détrie, C., Lopez, C., Moïse, C., Roche, M. & Verine, B. (2013). Sud4science, de l'acquisition d'un grand corpus de SMS en français à l'analyse de l'écriture SMS. *Épistémè - revue internationale de sciences sociales appliquées*, 9, 107-138.
- Patard, A. (2014). Réflexions sur l'origine de l'insubordination. Le cas de trois insubordonnées hypothétiques du français. *Langages*, 196(4), 109-130. <https://doi.org/10.3917/lang.196.0109>
- Pierrard, M. (1996). *Ce que* en subordination propositionnelle : Un stabilisateur catégorie. In C. Muller (éd.) *Dépendance et intégration syntaxique. Subordination, coordination, connexion* (p. 77-88). De Gruyter.
- Pincemin, B. (2012). Hétérogénéité des corpus et textométrie. *Langages*, 187(3), 13-26. <https://doi.org/10.3917/lang.187.0013>
- Poplack, S. & Dion, N. (2021). Cartographie de la variation et du changement morphosyntaxique en français : Leçons à retenir. *Cahiers internationaux de sociolinguistique*, 18(1), 83-115. <https://doi.org/10.3917/cisl.2101.0083>
- Praxiling - UMR 5267 (2019). *Corpus 14* [Corpus]. ORTOLANG (Open Resources and TOols for LANGuage) - www.ortolang.fr, v2, <https://hdl.handle.net/11403/corpus14/v2>. ou

- version en ligne : Corpus 14 [En ligne]. Montpellier, Laboratoire PRAXILING (UMR 5267, Université Paul-Valéry Montpellier-CNRS), <http://textometrie.univ-montp3.fr/>
- Prévost, S. (2011a). *Expression et position du sujet pronominal du 12^{ème} au 14^{ème} siècle : Une approche quantitative* [Recherche inédite en vue de l'obtention de l'Habilitation à Diriger des Recherches]. École Normale Supérieure de Lyon.
- Prévost, S. (2011b). *Français médiéval en diachronie : Du corpus à la langue* [Mémoire de synthèse d'Habilitation à diriger des recherches]. École Normale Supérieure de Lyon.
- Prévost, S. (2020). Partie 1 : Fondements théoriques et méthodologie. In C. Marchello-Nizia, B. Combettes, S. Prévost & T. Scheer (éds), *Grande grammaire historique du français* (Vol. 1-2). De Gruyter Mouton.
- Prévost, S. & Marchello-Nizia, C. (2020). Le sujet. In C. Marchello-Nizia, B. Combettes, S. Prévost & T. Scheer (éds), *Grande grammaire historique du français* (Vol. 1-2, p. 1055-1126). De Gruyter Mouton.
- Rey, A. (1972). Usages, jugements et prescriptions linguistiques. *Langue française*, 16(1), 4-28. <https://doi.org/10.3406/lfr.1972.5701>
- Riegel, M., Pellat, J.-C. & Rioul, R. (2008). *Grammaire méthodique du français* (3^e éd.). PUF. (Original work published 1994)
- Riegel, M., Pellat, J.-C. & Rioul, R. (2018). *Grammaire méthodique du français* (7^e éd.). PUF. (Original work published 1994)
- Rodríguez Somolinos, A. (éd.). (2013). *Marques d'oralité en français médiéval*. PUPS.
- Rodríguez Somolinos, A. (2016). *Énonciation et marques d'oralité dans l'évolution du français*. Presses universitaires de Paris Nanterre.
- Rouquier, M. (1990). Le terme *ce que* régissant une complétive en ancien français. *Revue Romanique*, 25, 47-72.
- Rouquier, M. (2007). Les constructions clivées en ancien français et en moyen français. *Romanica*, 125(1), 167-212.
- Rouquier, M. (2014). *L'émergence des constructions clivées, pseudo-clivées et liées en français*. Classiques Garnier.
- Schnedecker, C. & Vaguer, C. (éds). (2020). *L'oral représenté en diachronie et en synchronie : Une voie d'accès à l'oral spontané?* Armand Colin.
- Schneider, S. (2007). *Reduced Parenthetical Clauses as Mitigators : A corpus study of spoken French, Italian and Spanish*. John Benjamins. <https://doi.org/10.1075/sc1.27>
- Schneider, S. (2019). *Du verbe au marqueur verbal : L'évolution des réseaux de constructions* [Communication non publiée]. Colloque international DIACHRO IX: le français en diachronie, Salamanque, 29 mars 2019.
- Schneider, S. (2022). *Phases d'évolution linguistique. Une analyse comparative* [Communication non publiée]. Colloque international DIACHRO X : le français en diachronie, Paris, 31 mai 2022.
- Schøsler, L. (1984). *La déclinaison bicasuelle de l'ancien français : Son rôle dans la syntaxe de la phrase, les causes de sa disparition*. Odense University Press.

- Schøsler, L. (2021). L'analyse constructionnelle des structures à réduction du sujet. *Langue française*, 209, 101-118.
- Schøsler, L. & Glessgen, M. (2018). Repenser les axes diasystématiques : Nature et statut ontologique. In M.-D. Glessgen, J. Kabatek & H. Völker (éds), *Repenser la variation linguistique : Actes du colloque DIA IV à Zurich (12-14 sept. 2016)* (p. 11-52). ELiPhi.
- Serbat, G. (2003). Les complétives en *quod*. In C. Bodelot (éd.), *Grammaire fondamentale du latin, Tome X, Les propositions complétives en latin* (p. 528-747). Peeters.
- Siouffi, G. (éd.). (2012). *Sentiment de la langue et diachronie*. PUPS.
- Siouffi, G. (éd.). (2016). *Modes langagières dans l'histoire*. Honoré Champion.
- Siouffi, G. (2016). Présentation. In G. Siouffi (éd.), *Modes langagières dans l'histoire* (p. 7-30). Honoré Champion.
- Siouffi, G. (2020). Partie 2 : Histoire externe. In C. Marchello-Nizia, B. Combettes, S. Prévost & T. Scheer (éds), *Grande grammaire historique du français* (Vol. 1-2). De Gruyter Mouton.
- Siouffi, G. (éd.). (2020). *Une histoire de la phrase française : Des Serments de Strasbourg aux écritures numériques*. Actes Sud.
- Siouffi, G., Steuckardt, A. & Wionet, C. (2012). Comment enquêter sur les diachronies courtes et contemporaines? *3^e Congrès Mondial de Linguistique Française, 1*, 215-226. <https://doi.org/10.1051/shsconf/20120100214>
- Siouffi, G., Steuckardt, A. & Wionet, C. (2018). Y a-t-il de faux changements linguistiques ? L'exemple des modalisateurs. In W. Ayres-Bennett (éd.), *Nouvelles voies d'accès au changement linguistique* (p. 421-437). Classiques Garnier.
- Soutet, O. (1990). *La concession en français. Des origines au XVI^e siècle*. Droz.
- Stark, E. (2016-2018). Projet FSN « What's up, Switzerland? » (Sinergia: CRSII1_160714). Université de Zurich. www.whatsup-switzerland.ch.
- Stark, E. (2017). Pertinence de l'analyse grammaticale en linguistique variationnelle. *Langage & société*, 160-161(2), 283-298.
- Steuckardt, A. & Dal Bo, B. (2018). *Avoir été* ou *être allé* ? Évolution d'une concurrence, d'après des corpus lettrés et peu lettrés. In P. Blumenthal & D. Vigier (éds), *Études diachroniques du français et perspectives sociétales* (p. 295-310). Peter Lang. <https://doi.org/10.3726/b11459>
- Tanguy, N. (2009). *Les segments averbaux, unités syntaxiques de l'oral* [Thèse de doctorat]. Université Paris 3.
- Tesnière, L. (1959). *Éléments de syntaxe structurale*. Klincksieck.
- Thompson, S. A. & Mulac, A. (1991a). A quantitative perspective on the grammaticization of epistemic parentheticals in English. In E. C. Traugott & B. Heine (éds), *Approaches to grammaticalization. Vol. 2 : Focus on types of grammatical markers* (p. 313-329). John Benjamins.
- Thompson, S. A. & Mulac, A. (1991b). The discourse conditions for the use of the complementizer *that* in conversational English. *Journal of pragmatics*, 15, 237-251.

Bibliographie

- Urmson, J. O. (1952). Parenthetical Verbs. *Mind*, 61(244), 480-496. <https://doi.org/10.1093/mind/lxi.244.480>
- Verjans, T. (2009). *Essai de systématique diachronique : Genèse des conjonctions dans l'histoire du français* [Thèse de doctorat]. Université Paris-Sorbonne.
- Verjans, T. (2012). Le linguiste et la variante : quelle (s) leçon (s) en tirer ? In A. Salamon, A. Rochebouet & C. Lecornec-Rochelois (dir.) *Le texte médiéval. De la variante à la re-création* (p. 89-99). Presses de l'Université Paris-Sorbonne.
- Verjans, T. (2014). "Système de possibilités" et changement linguistique. In W. Ayres-Bennett & T. M. Rainsford (éds), *L'histoire du français : État des lieux et perspectives* (p. 305-320). Classiques Garnier.
- Vermander, P. (2020). *Les reliefs de la voix. Écriture et oralité en moyen français* [Thèse de doctorat]. Université Sorbonne Nouvelle.
- Völker, H. (2009). La linguistique variationnelle et la perspective intralinguistique. *Revue de linguistique romane*, 73, 7-76. <https://doi.org/10.5167/UZH-28298>
- Weinreich, U. & Herzog, M. (1968). Empirical foundations for a theory of language change. In W. P. Lehmann & Y. Malkiel (éds), *Directions for Historical Linguistics* (p. 95-195). University of Texas Press.
- Winter-Froemel, E. (2020). Traditions discursives. *Publictionnaire. Dictionnaire encyclopédique et critique des publics*. Mis en ligne le 04 décembre 2020. Accès : <http://publictionnaire.huma-num.fr/notice/traditions-discursives/>.
- Wirth-Jaillard, A. (2013). De Estienne Husson pour ce qu'il dit à Jannon Morelot : 'Un filz de bastarde ne me puet valoir !'. In D. Lagorgette & P. Larrivée (éds), *Représentations du sens linguistique 5 : Actes de la conférence internationale organisée à Chambéry, 25-27 mai 2011* (p. 65-80). Presses Universitaires de Savoie.
- Wüest, J. (2010). La notion de diamésie est-elle nécessaire? *Travaux de linguistique*, 59(2), 147-162. <https://doi.org/10.3917/tl.059.0147>
- Zumthor, P. (1981). Intertextualité et mouvance. *Littérature*, 41, 8-16.

Annexes

Annexe 1 : Curriculum Vitae

Nom : Julie GLIKMAN
Âge : 43 ans
Date de naissance : 12/12/1979
Situation actuelle : Maître de Conférences en Linguistique française diachronique
Corps : MCF depuis 01/09/2012
Grade : CN - échelon 6 depuis octobre 2019
Établissement : Université de Strasbourg, Faculté des Lettres, Institut de Linguistique et Langue Française
Laboratoire : Linguistique, Langues, Parole (LiLPa), UR 1339
ID ORCID : <https://orcid.org/0000-0002-9332-2986>

DIPLÔMES

-
- 2010 **Qualification** en section 7 du CNU.
 2009 **Thèse de doctorat en cotutelle** en Sciences du Langage (Université Paris Ouest Nanterre – La Défense, France et Université de Potsdam, Allemagne) : *Parataxe et Subordination en Ancien Français, Système syntaxique, variantes et variation*, mention très honorable avec les félicitations du jury, financement : allocataire de recherche. Soutenue publiquement le 5 décembre 2009.
 2005 **DEA de Sciences du Langage** : *La parataxe en diachronie*, sous la direction de A. Bertin, Paris X Nanterre, mention très bien, financement : bourse de mérite.
 2004 **Maîtrise de Sciences du Langage** : *La parataxe dans La Chanson de Roland*, Paris X Nanterre, sous la direction de A. Bertin et I. Weill, mention bien.
 2002 **Licence de Lettres Modernes**, Paris X Nanterre.
 1999 **DEUG de Lettres Classiques**, Paris X Nanterre.
 1997 **Baccalauréat général série Littéraire** : mention assez bien.

EXPÉRIENCES PROFESSIONNELLES

-
- 2022-23 Renouvellement de la **Délégation CNRS**, laboratoire ATILF, Nancy (1 an)
 2021-22 Obtention d'un **CRCT** au titre du CNU (6 mois) et d'une **Délégation CNRS**, laboratoire ATILF, Nancy (6 mois)
 depuis 2012 **Maître de Conférences en Linguistique française diachronique**, Université de Strasbourg, Faculté des Lettres, département de Linguistique Française
 2010-12 **Post-Doctorat**, LaTTiCe, Montrouge, Projet ANR-DFG SRCMF et GGHF
 2009-10 **A.T.E.R.**, Université de Savoie, Chambéry (U. Savoie) (demi-poste)
 2008-09 **A.T.E.R.**, Université de Paris Ouest Nanterre – La Défense (UPX) (demi-poste)
 2005-08 **Allocataire de recherche**, Université de Paris Ouest Nanterre – La Défense
 2010 Vacances de recherche, projet SRCMF (Syntactic Reference Corpus of Medieval French), (S. Prévost & A. Stein resp.), annotation syntaxique de corpus.
 2008-09 Vacances de recherche, projet GGHF (Grande Grammaire Historique du Français), (C. Marchello-Nizia, B. Combettes, T. Scheer et S. Prévost resp.)
 2008 Vacances de recherche, projet ANR ELICO (L. Tovena resp.) (annotation de corpus d'ancien français)
 2005-06 Chargée de cours à l'Université de Paris Ouest Nanterre – La Défense

PRIMES ET DISTINCTIONS

2018 Attribution de la PEDR

FONCTIONS D'INTÉRÊT COLLECTIF

Responsabilités pédagogiques

2016-2020 et 2022-2024 : **Directrice de l'Institut de Linguistique Française** (3 mandats)

2014-2016 : **Directrice adjointe de l'Institut de Linguistique Française**

2018- 2024 : **Membre du Comité de pilotage du Master Sciences du Langage**

2018- 2024 : **Membre du Conseil de Perfectionnement de la Licence Sciences du Langage**

2013-2018 : **Responsable de la Licence Lettres mention Sciences du Langage**

Responsabilités scientifiques

2021- : **Membre du comité de pilotage** de la Fédération De Recherche Langage Et Communication (FRLC)

2019- : **Membre du bureau** de Fonctionnement discursif et traduction

2019- : **Membre du comité de pilotage** du consortium Corli.

2018- : **Membre du groupe de travail** sur la restructuration thématique du laboratoire LiLPa

Responsabilités éditoriales

2018- : Comité éditorial de la revue *Studia Linguistica Romanica* (<https://studialinguisticaromanica.org/index.php/slr>)

2014- : Comité de rédaction de la revue *Linx* (<https://journals.openedition.org/linx/>)

Autres responsabilités - sociétés savantes

2018-2022 : Membre du bureau et webmestre de la Société Internationale de Diachronie du Français (SIDF, <https://diachronie.org/>)

2014-2018 : Membre du bureau de la Société Internationale de Diachronie du Français (SIDF)

Membre de comité d'évaluation national

2019- : Membre nommée de la section 07 Sciences du langage du Conseil National des Universités (CNU)

Jury nationaux

2013, 2014, 2015, 2017, 2018 : MEMBRE DU JURY DU CAPES EXTERNE DE LETTRES

ENSEIGNEMENTS DISPENSÉS DEPUIS 2015 :

<i>LINGUISTIQUE DIACHRONIQUE</i>			
Intitulé	Volume	Public	Effectifs moyens
Fonctionnement et histoire de l'orthographe	24h	L1 SDL	25 étudiants
Naissance de la norme et constitution du français moderne : le français préclassique et classique	24h	L1 SDL	25 étudiants
Interface synchronie / diachronie : variation et changement en cours	12h	L2 SDL	50 étudiants
Introduction à l'ancien français	24h	L2 SDL, LM, LC	35 étudiants
Introduction à l'ancien et au moyen français	24h	L2 SDL, LM, LC	35 étudiants
Constitution du système orthographique du français	9h	L3 LM	90 étudiants
Introduction à la recherche en linguistique diachronique	24h	L3 SDL, LM, LC	25 étudiants
Le changement linguistique	12h	Master SDL	15 étudiants
Histoire de la langue française	24h	L1 SDL	80 étudiants
Moyen Français	24h	L3 SDL, LM, LC	35 étudiants
Français Préclassique et Classique	24h	L3 SDL, LM, LC	25 étudiants
<i>Linguistique synchronique et générale</i>			
Langue, codes et communication : le rapport oral/écrit	12h	L1 SDL	70 étudiants
Recueil et traitement des données	18h	L2 SDL	50 étudiants
Méthodologie disciplinaire	18h	L1 SDL	30 étudiants
Stylistique - TD	12h	L3 LM, LC	30 étudiants
Histoire de la linguistique – <i>avant Saussure</i>	12h	L1 SDL	80 étudiants
Histoire de la linguistique – <i>après Saussure</i>	12h	L1 SDL	80 étudiants
<i>Professionnalisation</i>			
Projet Personnel et Professionnel de l'étudiant	12h	L3 SDL, LM, LC	15 étudiants

CHARGE ANNUELLE D'ENSEIGNEMENT DEPUIS 2012 :

ANNÉE UNIVERSITAIRE	ÉTABLISSEMENT D'ENSEIGNEMENT	NOMBRE D'HEURES ENSEIGNÉES (HETD)	DÉCHARGES / CRCT / DÉLÉGATION CNRS OBTE-NUS... (PRÉCISEZ)
2022-23	Université de Strasbourg	0	Délégation CNRS
2021-22	Université de Strasbourg	0	CRCT d'un semestre accordé pour projet de recherche par le CNU et délégation CNRS
2020-21	Université de Strasbourg	200,25	
2019-20	Université de Strasbourg	220,25	
2018-19	Université de Strasbourg	227,75	
2017-18	Université de Strasbourg	215	
2016-17	Université de Strasbourg	207,50	
2015-16	Université de Strasbourg	117	Décharge au terme de MCF nouvel arrivant en première nomination à l'Université de Strasbourg (décharge à prendre dans les 5 ans de la nomination)
2014-15	Université de Strasbourg	216,5	
2013-14	Université de Strasbourg	204	
2012-13	Université de Strasbourg	204,5	

DIRECTION DE MÉMOIRES DE MASTER (M1/M2) SOUTENUS (5)

Léonie BOUCHWALDER, *Détection et caractérisation des discours sur la langue dans la presse*, Master 1 (2018-2019) et Master 2 (2019-2021) Technologie des Langues, codirection avec Delphine Bernhard (50%).

Marie DIDIENNE, *Néologismes et enrichissement de la langue française.*, Master 1 (2017-2018) & Master 2 (2018-2019) Sciences du langage (encadrement 100%).

Clara BRINGER, *Lemmatisation et annotation des formes verbales du français médiéval.*, Master 1 (2017-2018) & Master 2 (2018-2019) Technologie des Langues, codirection avec Delphine Bernhard (50%).

Estelle ROBERT, *Étude de la vitalité des mots et expressions du français régional.* Master 1 (2017-2018) & Master 2 (2018-2019) Technologie des Langues, codirection avec Delphine Bernhard (50%).

Evangelia BALOUTA, *Étude de la vitalité des mots régionaux du français.*, Master 1 (2015-2016) et Master 2 (2016-2017) Technologie des Langues, codirection avec Delphine Bernhard (50%).

DIRECTION DE MÉMOIRES DE MASTER EN COURS (2)

Jonathan FONTAINE, Full : *l'émergence d'un nouveau mot, ses sens et son emploi*, Master Recherche Sciences du langage (M1 2020-2021 soutenu), Master 2 en cours 2021-2023 (encadrement 100%).

Elia VERTUEUX : sujet en cours de définition, Master Recherche Sciences du langage, M1 2022-2023 (encadrement 100%).

COMITÉ DE SUIVI DE THÈSE (2)

Timothée CELEYRON, *Phraséologie et genres textuels en diachronie longue. Étude du lexique de l'armement guerrier*. sous la direction de J. Sorba, Université Grenoble Alpes, première inscription 09.2022.

Charlène WEYH, *Évolution du système verbal français, entre régularisations et normes (1300-1700) : le cas du présent de l'indicatif*, sous la direction de S. Bazin & B. Bouard, Université de Lorraine, ATILF, première inscription 09.2019.

AUTRES TÂCHES D'ENCADREMENT SCIENTIFIQUE

Encadrement régulier de stages de Master dans le cadre des projets de recherche DFS et Les Vocaux

Membre du comité de lecture des rapports de stage, Master Métiers de l'Édition.

Accueil de stagiaires dans le cadre du programme d'études « Strasbourg Field Study & Internship » d'IFE (Internships in Francophone Europe) et suivi de leur mémoire de fin de stage.

Accueil d'un stagiaire lycéen dans le cadre d'un stage d'observation de milieu professionnel.

DIRECTION DE PROGRAMMES NATIONAUX FINANCÉS (5)

2022-2024 : ORALIDIA : Oralité et diachronie : une voie d'accès au changement linguistique.

Lauréat de l'appel à projet IDEX Recherche exploratoire, Université de Strasbourg. Durée du projet : deux ans, octobre 2022 à décembre 2024. Montant de la subvention reçue : 18 000 euros. Objectifs : constitution d'un corpus inédit de français parlé spontané : les « sms vocaux », en vue de l'étude de la diffusion des formes émergentes ou de leur disparition, et du français parlé dans différents contextes, en particulier informels. Mise en avant du patrimoine et des particularités linguistiques de la région Grand Est, en privilégiant le recueil auprès des populations de Nancy et Strasbourg. Recherche exploratoire en vue du dépôt de projet d'envergure.

2022 : ORALITÉ ET DIACHRONIE : UNE VOIE D'ACCÈS AUX FORMES ÉMERGENTES, en codirection avec Christophe Benzitoun. Lauréat de l'appel à projet Recherche, laboratoire ATILF et Université de Lorraine, co-financement avec le laboratoire LiLPa de l'Université de Strasbourg. Montant obtenu dans le cadre de l'appel à projet : 4 100 euros. Objectifs : étudier les dia-variations en français contemporain et les mécanismes en œuvre dans la création, diffusion de formes émergentes et les disparitions de formes vieillissantes, afin de mieux comprendre le changement linguistique, à travers l'étude des nouveaux modes de communication. Projet exploratoire en vue de la préparation du projet de constitution de corpus Les Vocaux.

2017 : DONNEZ VOTRE FRANÇAIS À LA SCIENCE ! DFS2. Lauréat de l'appel à projets national *Langues et numérique 2017* « réalisation », Délégation Générale à la Langue Française et aux Langues de France (DGLFLF), Ministère de la Culture et Communication. Durée du projet : 12 mois. Montant de la subvention reçue : 37 000 euros. Objectifs : développement de la plateforme de recueil de données linguistiques avec intégration d'un volet enregistrement de productions orales et optimisation du générateur de formulaire pour mise à la disposition de la communauté de l'outil de construction d'enquête en ligne.

- 2016 : DONNEZ VOTRE FRANÇAIS À LA SCIENCE ! Lauréat de l'appel à projets national *Langues et numérique 2016*, Délégation Générale à la Langue Française et aux Langues de France (DGLFLF), Ministère de la Culture et Communication. Durée du projet : 12 mois. Montant de la subvention reçue : 10 000 euros. Objectifs : réalisation d'une plateforme de recueil de données linguistiques et de retour au grand public (quiz de vulgarisation scientifique, jeu de géolocalisation et enquêtes linguistiques en ligne) : www.donnezvotre-francais.fr.
- 2015 : LE FRANÇAIS DE NOS RÉGIONS. Soutien financier de Délégation Générale à la Langue Française et aux Langues de France (DGLFLF), Ministère de la Culture et Communication, hors appel. Durée du projet : 12 mois. Montant de la subvention reçue : 10 000 euros. Objectif : création d'un consortium de recherche autour de la variation en français régional.

IMPLICATION DANS DES PROJETS DE RECHERCHE FINANCÉS (HORS DIRECTION)

- 2021-2024 : Implication à 70% dans l'Institut du langage et de la communication (Université de Strasbourg, projet bénéficiant d'un soutien spécifique de 4 ans dans le cadre d'un accompagnement à l'effort de structuration (2021-2024) - Création d'une fédération de recherche, <http://savoirs.unistra.fr/recherche/linterdisciplinarite-nouvelle-generation/les-instituts-thematiques-interdisciplinaires-iti/>)
- 2016-2020 : Implication à 20% dans le projet ANR-15-CE38-0008 DEMOCRAT (DEscription et MODélisation des Chaînes de Référence : outils pour l'Annotation de corpus (en diachronie et en langues comparées) et le Traitement automatique), F. Landragin, C. Guillot & C. Schnedecker (resp.) - Analyse linguistique des chaînes de référence en diachronie et constitution d'un corpus diachronique annoté en chaînes de référence (<https://www.lattice.cnrs.fr/democrat/>). Annotation syntaxique, participation à l'élaboration du modèle, rédaction du guide d'annotation
- 2011-12 : Participation au PEPS INS2I - INSHS *Modélisation Contrastive et Computationnelle des Chaînes de Coréférence* (MC4), volet diachronique (F. Landragin dir.).
- 2010-2012 : Vacations puis post-doctorat dans le projet ANR-DFG Syntactic Reference Corpus of Medieval French (SRCMF), S. Prévost & A. Stein (resp.) - Constitution d'un corpus de textes médiévaux annoté en syntaxe (<http://srcmf.org/>).
- 2008-2010 : Vacations pour le projet GGHF (Grande Grammaire Historique du Français), C. Marchello-Nizia, B. Combettes, T. Scheer et S. Prévost (resp.) - Veille bibliographique et constitution d'une bibliographie commune.
- 2008 : Vacations pour le projet ANR ELICO, L. Tovena (resp.) - Construction d'une database diachronique sur les déterminants du français (annotation pour la partie ancien français) (<http://elico.linguist.jussieu.fr/index.php>).
- 2008 – 2015 : Participation au groupe de travail *Les Conjonctions en diachronie* (A. Bertin resp.) (projet financé par l'ILF pour la période 2008/2009).
- 2006-07 : Participation au projet *Locutions, langues, cultures, cognition*, coordonné par D. Lee-man.

ORGANISATION DE CONGRÈS NATIONAUX/INTERNATIONAUX (3)

2017 :

Colloque international *Diachro VIII*, Strasbourg, 2-4 février 2017, comité d'organisation : D. Capin (Université de Strasbourg), J. Glikman (Université de Strasbourg), A. Kuyumcuyan (Université de Strasbourg), V. Obry (Université de Haute Alsace), T. Revol (Université de Strasbourg), C. Schnedecker (Université de Strasbourg).

2012 :

Organisation en collaboration avec M. Avanzi, S. de Pontonx et S. Schneider du colloque international « Les « Verbes Parenthétiques » : Hypotaxe, Parataxe ou Parenthèse ? » (<http://sites.google.com/site/parenthetiques2012/>)

2007 :

Organisation en collaboration avec L. Mansour, Y. Schochenmeier et S. Weiser du quatrième colloque des doctorants et jeunes chercheurs de Paris X Nanterre (Coldoc 07) « Le vocabulaire scientifique et technique en Sciences du Langage » Nanterre, 20-21 juin.

ORGANISATION DE JOURNÉES D'ÉTUDE (10)

2022 :

Journée d'étude *Sciences participatives et nouvelles données pour la recherche en linguistique*, Nancy, Vendredi 30 Septembre 2022, organisée par Christophe Benzitoun et Julie Glikman.

Journées d'étude *Variation(s) en français*, Nancy, 17 et 18 novembre 2022, organisée par Bérengère Bouard et Julie Glikman

2017 :

Journée d'étude *La linguistique des genres en questions*, Strasbourg, 9 Juin 2017, organisée par Christophe Gérard et J. Glikman.

Journée d'étude *Réécritures nouvelles : Sentiment et conscience de la réécriture*, 3^e Journée Réécritures médiévales, Strasbourg, 8 décembre 2017, organisée par D. Capin, J. Glikman et T. Revol.

2016 :

Journée d'étude *Des dialectes galloromans aux français régionaux : Cartographier et interpréter les données*, Strasbourg, 27 mai 2016, organisée par M. Avanzi et J. Glikman.

2015 :

Journée d'étude *Réécritures médiévales : domaine juridique, scientifique, religieux, littéraire*, Strasbourg, 23 janvier 2015, organisée par Daniéla Capin, Julie Glikman, Pierre Nobel, Thierry Revol.

Journée d'étude *Chaînes de Référence et phénomènes de variation*, Strasbourg, 5 juin 2015, organisée par Julie Glikman et Catherine Schnedecker.

Journée d'étude *Réécritures génériques médiévales*, 2^e Journée Réécritures médiévales, Strasbourg, 25 septembre 2015, organisée par Daniéla Capin, Julie Glikman, Pierre Nobel, Thierry Revol.

Journée d'étude *Variation linguistique et crowdsourcing : Étudier la variation au 21^e siècle*, Strasbourg, 23 octobre 2015, organisée par Mathieu Avanzi et Julie Glikman.

2010 :

Organisation en collaboration avec M. Avanzi d'une journée d'étude « Entre rection et incidence : des constructions verbales atypiques ? », 26 mars 2010, Nanterre (<http://sites.google.com/site/cvriparis2010/>)

Organisation de séances thématiques de séminaire du laboratoire d'accueil

- 2021 Organisation d'une séance thématique du séminaire LiLPa « L'annotation », 22 janvier 2021.
- 2019-21 Co-responsable du séminaire de l'équipe *Fonctionnement discursif et traduction*, laboratoire LiLPa.
- 2018 Organisation d'une séance thématique du séminaire du laboratoire LiLPa « Normes et variation », 9 mars 2018.
- 2009 Organisation d'une séance thématique du séminaire du laboratoire Modyco « Quand la syntaxe dépasse les bornes, les phrases l'imitent : la question de l'unité maximale de la syntaxe », 18 Novembre 2009.
- 2009 Organisation d'une séance thématique du séminaire du laboratoire Modyco « L'interface prosodie-syntaxe-pragmatique », 21 Janvier 2009.
- 2006-07 Organisation du groupe de travail *Syntaxe*, Nanterre.

EXPERTISES

Membres de comités scientifiques de colloques

Congrès Mondial de Linguistique Française (2022, 2020, 2014), *Colloque de la Société Internationale de Diachronie du Français (SIDF)* (2020), *La postmodalité et les cycles de vie des expressions modales* (2022), « *Langues minoritaires* » : *quels acteurs pour quel avenir ?* (2019), *Syntaxe et Discours III : Types d'unités et procédures de segmentation* (2017), *DIACHRO VII Le français en diachronie* (2015), *Consécutivité et Simultanéité en linguistique, langues et parole* (2015), *DéterminationS* (2014), *Disparitions et changements linguistiques* (2011)

Expertises ponctuelles pour un ouvrage

Expertise d'un article à l'issue du 10^e Colloque international « *Christine de Pizan* », Sorbonne Nouvelle, 18 juin 2019 au 21 juin 2019, organisé par D. Demartini, C. Le Ninan, G. Parussa et A. Valentini en vue de la publication.

Le français innovant, Collection « Sciences pour la communication », F. Diémoz, G. Dostie, P. Hadermann et F. Lefeuve (dir.), Peter Lang, 2020.

Types d'unités et procédures de segmentation, M.-J. Béguelin, G. Corminboeuf, F. Lefeuve (dir.), Lambert Lucas, collection *Études linguistiques et textuelles (ELT)*, 2020.

Le français en diachronie, Dépendances syntaxiques, morphosyntaxe verbale, grammaticalisation, S. Prévost & B. Fagard (éd.), Peter Lang, 2017.,

Lexique, Grammaire, Discours. Les marqueurs discursifs, G. Dostie & F. Lefeuve (dir.) Paris, Honoré Champion, 2017.

Expertises pour des revues scientifiques :

Travaux de Linguistique, Journal of Pragmatics, Discours, Linx. Çédille, Les Cahiers du GEPE

LISTE COMPLÈTE DES COMMUNICATIONS ORALES SANS ACTES (C-COM)

Les communications ayant donné lieu à publication ne sont pas mentionnées ici

Communications dans des colloques internationaux (15)

- Glikman J., Camille Fauth, Nicolas Mazziotta, Christophe Benzitoun (2022) « Une nouvelle voie d'accès au français populaire : les Vocaux », *13^e congrès des francoromanistes*, 21-24 septembre 2022, Université de Vienne.
- Glikman J., Landragin F., Schnedecker C. et Todirascu A. (2022) Un exemple de corpus annoté en diachronie longue : le corpus Democrat, enjeux et exploitations. *ConCorDiaL*, 14 octobre 2022, Grenoble.
- Glikman J. (2022) « Variation et mécanismes d'analogie dans l'évolution des locutions conjonctives formées sur *façon*, *manière* et *sorte* » Colloque international *DIACHRO X Le français en diachronie*, 17 au 19 janvier 2022, Sorbonne Université.
- Glikman J. (2019) « Les propositions complétives sujet dans l'histoire du français » *DIACHRO IX · LE FRANÇAIS EN DIACHRONIE*, Université de Salamanque (Espagne), les 28, 29 et 30 mars 2019, Facultad de Filología, Palacio de Anaya, Plaza de Anaya, Salamanca
- Glikman J., N. Mazziotta (2016) « Oral représenté et narration dans la Chanson de Roland et La Conquête de Constantinople de Robert de Clari. » *Colloque International Représentation de l'oral et marques d'oralité dans la diachronie du français*, Facultad de Filología – Universidad Complutense de Madrid, Madrid 12-13 mai 2016
- Mazziotta N., J. Glikman (2014) « Spécificités syntaxiques des types de discours dans les *Quatre Livres des Rois* (12^e s.) », Colloque *Les états anciens de langue à l'heure du numérique*, Paris, Novembre 2014.
- Glikman J. et T. M. Rainsford (2012) *Ce objet ou attribut : étude diachronique*, colloque Diachro 6, Louvain, 17-20 octobre 2012.
- Glikman J. et S. Prévost (2012) *Non-expression du sujet dans les subordonnées en ancien français*, colloque DIA II, Copenhague, 19-21 novembre 2012.
- Glikman J., Fagard B. (2011) « Des traits d'oralité en ancien français ? L'exemple de *savoir* », colloque de la Société Internationale de Diachronie du Français, Nancy, Septembre 2011.
- Glikman J. (2010) « La double complémentation : double marquage ou dislocation ? », Colloque ComplémentationS, Santiago de Compostela, Espagne, 20-23 oct. 2010.
- Glikman J., Avanzi M., Benzitoun C. (2010) « Essai de typologie comparée des modes d'intégration phrastique en ancien français et en français moderne », 26^e Congrès International de Linguistique et Philologie Romane (CILFR 2010), Valence, Espagne, 6-11 septembre 2010.
- Tanguy, N., J. Glikman (2009) « *Deux énoncés averbaux en parataxe, et hop, une prédication complexe*. Les énoncés averbaux peuvent-ils construire des prédications complexes paratactiques en français parlé ? », Conférence de l'Association For French Language Studies (AFLS), Neuchâtel, Suisse, 3-5 septembre 2009.
- Glikman J. (2008) « Les constructions à Verbe en 1^{ère} position : structures interrogatives ou système hypothétique ? », colloque international Diachro IV, Madrid, Espagne, 22-24 octobre 2008 (<http://www.ucm.es/info/diachroIV/archivos/resumen.html>).
- Glikman J. (2008) « Comparative and diachronic approach of the non introduced subordinates in French and German », Fifth International Contrastive Linguistics Conference (ICLC5), Leuven, Belgique, 7-9 juillet 2008.

Glikman J. (2008) « Modalisation vs subordination », Colloque du Collège doctoral franco-allemand : *Locutions : les langues entre cultures et cognition*, Potsdam, Allemagne, 22-23 mai 2008.

Communications à des journées d'étude (14)

- Glikman J. et Mazziotta N. (2022) « Projet « Les Vocaux » : Outils et formats », *TraSoGal*, Liège, 24 juin 2022.
- Glikman J., Mazziotta N., Fauth C., Benzitoun C. (2022) « Le projet Les Vocaux : bilan d'étape. » *Sciences participatives et nouvelles données*, Nancy, 30 sept. 2022.
- Glikman J. (2022) « Sentiment linguistique et perception de la variation comme indice du changement linguistique » *Variation(s) en français*, Nancy, 17 et 18 novembre 2022.
- Glikman J. (2021) « Variation et analogie dans la formation des locutions conjonctives », *TraSoGal 2021*, journée d'études sur le traitement des sources galloromanes, Université de Liège, 12 novembre 2021.
- Glikman J. (2021) « Corpus et enquête. Recueillir des données avec DFS », *Plate-forme collaborative et participative de transcription, relecture et annotation de corpus*, Workshop interdisciplinaire organisé par Huma-Num, la MSH Lorraine et CORLI, en ligne, 18 mai 2021.
- Glikman J. (2019) « Stratégies d'évitement dans l'histoire des complétives sujet en français », *Les constructions verbales (non) canoniques : De la réalisation argumentale à la structure propositionnelle*, Université de Fribourg, Suisse, 23 mai 2019.
- Glikman J. (2018) « Création d'une plateforme collaborative : Donnez votre français à la Science ! », colloque *Des données aux savoirs : les enjeux du numérique dans la recherche en Sciences Humaines et Sociales*, colloque MISHA 2018, 1 et 2 octobre 2018, Strasbourg.
- Glikman J. (2017) « Les propositions subordonnées de but en diachronie du français : travail pour la Grande Grammaire Historique du Français » *TraSoGal 2017*, 9e journée d'études sur le traitement des sources galloromanes, 4 mai 2017, Université de Liège.
- Glikman J. (2016) « Les variations au long cours : un cas de non-évolution ? », Journée d'étude *Le français médiéval : quel(s) rythme(s) d'évolution ?*, 11 mars 2016, Paris Sorbonne.
- Mazziotta N. & J. Glikman (2016) « Approche statistique exploratoire des spécificités syntaxiques de l'oral représenté en ancien français » *TraSoGal 2016* - 8e journée d'étude sur le traitement des sources galloromanes, 19 mai 2016, Université de Liège.
- Avanzi M. & J. Glikman (2016) « Le français de nos régions », *Journée d'étude Des dialectes galloromans aux français régionaux : Cartographier et interpréter les données*, Université de Strasbourg, le 27 mai 2016.
- Guillot C., Glikman J., Obry V. (2012) « Chaînes de coréférence dans *Les Miracles* d'Adgar », Journée d'étude du projet *Modélisation Contrastive et Computationnelle des Chaînes de Coréférence* (MC4), Lattice, 20 mars 2012.
- Glikman J. (2011) « Présentation du projet *Syntactic Reference Corpus of Medieval French* », Journée d'étude « Nouvelles technologies et français ancien », A. Bertin (org.), Nanterre, 1^{er} Avril 2011.
- Glikman J. (2010) « Méthodologie et problèmes de corpus en diachronie du français », Journée d'étude « Travaux sur Corpus », D. Le Pesant (org.), Nanterre, 14 Mai 2010.

Communication à des séminaires externes sur invitation (8)

- Glikman J. (2021) « Jugement des locuteurs et diachronie contemporaine : la perception des variantes comme indice du changement linguistique », Conférence au séminaire de master de G. Siouffi, Paris Sorbonne, 17 novembre 2021.
- Glikman J. (2021) « L'analyse des propositions complétives sujet dans l'histoire du français : problèmes théoriques et pratiques », séminaire de recherche de l'équipe de l'Institut für Linguistik/Romanistik, Université de Stuttgart, Allemagne, 19 mai 2021.
- Glikman J. (2017) « La subordination : quels critères pour l'intégration phrastique en diachronie ? » Séminaire du GEHLF (Groupe d'Étude en Histoire de la Langue Française), 20 octobre 2017, Paris Sorbonne.
- Glikman J. (2017) « L'évolution de la variation autour de « malgré que » et « à cause que » », Séminaire de l'Institut für Linguistik/Romanistik, Stuttgart, juillet 2017
- Bouard B. & J. Glikman (2017) « Histoire de TOUT devant adjectif : variation linguistique et élaboration d'une règle (XVIe-XIXe siècles). » ATILF, Réunion de l'équipe « linguistique historique française et romane », 25/01/17.
- Glikman J. (2009) « La parataxe en ancien français », Invitée par la chaire de linguistique française au séminaire de recherche en linguistique, Faculté des Lettres et Sciences Humaines, Université de Neuchâtel, Suisse, 22 avril.
- Glikman J. (2008) « Remarques sur les verbes recteurs faibles en diachronie », module M3 du Programme doctoral en Sciences du langage organisé sous l'égide de la Commission universitaire de suisse occidentale (CUSO), Fribourg, Suisse, 18-21 février.
- Glikman J. (2007) « Die asynthetische Subordination in Französisch, diachronische Aspekten », séminaire de G. Hassler, Potsdam, Allemagne, 5 juin.

Communications internes aux séminaires des laboratoires d'accueil (11)

- Glikman J. (2022) « Étude de diachronie contemporaine: la perception des variantes comme indice du changement linguistique. » Séminaire de l'équipe Discours, ATILF, 27 avril 2022.
- Glikman J., A. Todirascu (2021) « La constitution du guide d'annotation : deux retours d'expérience : coréférence et expressions polylexicales », Séminaire thématique LiLPa, 22 janvier.
- Glikman J. *et al.* (2017) « Donnez votre français à la science 2 », séminaire Projets LiLPa, 17 novembre.
- Glikman J., C. Schnedecker *et al.* (2017) « Projet Democrat : bilan à mi-parcours », séminaire Projets LiLPa, 17 novembre.
- Glikman J. (2015) « Constructions parenthétiques, marques d'oralité et type de textes en diachronie du français. Présentation de travaux récents menés en collaboration avec Stefan Schneider », séminaire de l'équipe Fonctionnement Discursif et Traduction, LiLPa, 10 avril.
- Glikman J. (2015) « Français de nos régions », séminaire Projets LiLPa, 20 novembre.
- Glikman J. (2015) « Democrat », séminaire Projets LiLPa, 20 novembre.
- Glikman J. (2012) « Structures syntaxiques et variation en ancien français », séminaire LiLPa, 14 décembre.
- Glikman J. (2009) « Phrase et subordination : *Crime et châtement* », séminaire du laboratoire Modyco, Nanterre, 18 novembre.
- Glikman J. (2009) « Les rapports prosodie – syntaxe en ancien français », séminaire du laboratoire Modyco, Nanterre, 21 janvier.

Glikman J. (2008) « Y a-t-il de la subordination en ancien français ? », séminaire du laboratoire Modyco, Nanterre, 16 janvier.

PV-COM : Communications et activités de vulgarisation (12)

2022 – présentation du projet Les Vocaux, entretien pour le magazine *Savoir(s)* de l'Université de Strasbourg : <https://savoirs.unistra.fr/eclairage/les-enregistrements-vocaux-passes-a-la-loupe>

2018 – présentation du projet à l'évènement public de restitution de la DGLFLF : Glikman J., J.-P. Goldman, C. Fauth, C. Benzitoun, M. Avanzi, P. Boula de Mareüil, Y. Scherrer, R. Gerbet, A. Lamielle. (2018) « Donnez votre français à la science ! », *Langues et Numérique 2018*, organisé par la DGLFLF. <https://www.culture.gouv.fr/Sites-thematiques/Langue-francaise-et-langues-de-France/Politiques-de-la-langue/Langues-et-numerique/Langues-et-numerique-2018-en-video>

2018, 2017, 2016 – Animation du stand « Bonjour, comment parlez-vous ? » avec C. Fauth, Fête de la science, Strasbourg (animation linguistique autour de la voix et de la variation linguistique diatopique et diachronique).

2017 – Festival *Pint of Science*, 17 mai 2017 : animation « Bonjour, comment parlez-vous ? » avec Camille Fauth (quiz linguistique autour de la voix et de la variation linguistique diatopique et diachronique).

2017 – Café scientifique junior au Lycée Foch, Haguenau, juin 2017 : « Langage d'hier, d'aujourd'hui et de demain », Enseignante : M. Diebolt, Intervenantes : J. Glikman, M.-A. Carvalho, M. Siemushyna, (Jardin des Sciences).

2017 – Interview pour le site Recherche de l'Université de Strasbourg : <http://www.recherche.unistra.fr/index.php?id=26259#c119058>

2016 – Interview pour le journal 20minutes : <http://www.20minutes.fr/magazine/economie-col-laborative/services-collaboratif/on-ne-sattendait-pas-a-cet-engouement-des-francais-pour-leur-langue-153039/>

2015 - Interview sur France Bleu Alsace (présentation des enquêtes du français de nos régions) <https://www.francebleu.fr/infos/societe/quel-francais-regional-parlez-vous-1444193234>

2015 - Interview pour la chaîne *Alsace20*, émission *6 minutes Eurométropole* (présentation du travail sur le français régional) (la vidéo n'est actuellement plus disponible en ligne).

2015 - Interview pour le journal d'actualité de l'Université de Strasbourg (présentation des enquêtes du français de nos régions) <http://www.lactu.unistra.fr/index.php?id=23200#c106862>

LISTE COMPLÈTE DES PUBLICATIONS

Le classement adopté ici est celui proposé par l'HCERES

ACL : Articles dans des revues à comité de lecture (11)

Les articles consistant en l'introduction de numéros coédités sont signalés par un astérisque.

Capin D., J. **Glikman**, C. Schnedecker & A. Todirascu (2021) « Le rôle des chaînes de référence dans la structuration textuelle : étude diachronique de l'ancien français au français moderne », *Langages* n°224, p. 87-107. DOI : 10.3917/lang.224.0087.

***Glikman** J. & T. Verjans (2021) « Introduction », *Regards linguistiques sur l'édition des textes médiévaux, Diachroniques* n°8, Paris, Sorbonne

***Glikman** J., C. Gérard (2019) « Présentation », *La linguistique des genres, en actes et en questions, Linx*, 78 | 2019, <https://journals.openedition.org/linx/3358>.

Mazziotta N. et J. **Glikman** (2019) « Oral représenté et narration en ancien français. Spécificités syntaxiques dans trois textes de genres distincts », *Linx*, 78 | 2019, <https://journals.openedition.org/linx/3151>.

Oberlé B., C. Schnedecker, E. Baumer, D. Capin, J. **Glikman**, C. Guo, Th. Revol, A. Todirascu, J. Tushkova (2018) « Les chaînes de référence dans les textes encyclopédiques du 12^e au 21^e siècle : étude longitudinale. », *Travaux de Linguistique*, 2018/2 n° 77, p. 67 – 141, <https://www.cairn.info/revue-travaux-de-linguistique-2018-2-page-67.htm>.

*Schnedecker C., J. **Glikman**, F. Landragin (2017) « Les chaînes de référence : annotation, application et questions théoriques. » *Langue française*, 2017/3 (N° 195), p. 5-16. <https://www.cairn.info/revue-langue-francaise-2017-3-page-5.htm>

Obry V., J. **Glikman**, C. Guillot-Barbance, B. Pincemin (2017) « Les chaînes de référence dans les récits brefs en français : étude diachronique (XIII^e – XVI^e s.). » *Langue française*, 2017/3 (N° 195), p.91 - 110. <https://www.cairn.info/revue-langue-francaise-2017-3-page-91.htm?contenu=resume>

Glikman J., C. Guillot-Barbance, V. Obry (2014) « Les chaînes de référence dans un corpus de textes narratifs médiévaux : traits généraux et facteurs de variation », in *Langages* 195, 43-60. <https://www.cairn.info/revue-langages-2014-3-page-43.htm>

Glikman J. (2012 [2009]) « Les incises en *croire, cuidier, penser* en Ancien Français », in numéro thématique « Entre rection et incidence : des constructions verbales atypiques ? Etudes sur je crois, je pense et autres parenthétiques », M. Avanzi & J. Glikman éd., *Linx* 61, 71-85. <https://journals.openedition.org/linx/1337>

*Avanzi M., J. **Glikman** (2012 [2009]) *Entre rection et incidence : des constructions verbales atypiques ? Études sur je crois, je pense et autres parenthétiques*, Introduction, *Linx* 61, 7-11. <https://journals.openedition.org/linx/1325>

Glikman J. (2010) « Rattachement et rection en Ancien Français », in « La rection verbale entre micro- et macro-syntaxe », A. Gautier, L. Pino Serrano & D. Van Raemdonck éd., *Travaux de Linguistique* 60, 29-44. <https://www.cairn.info/revue-travaux-de-linguistique-2010-1-page-29.htm>

ACLN : Articles de revue à comité de lecture non répertoriées par HCERES ou dans BDD internationales (4)

***Glikman**, J. G. Parussa, R. Waltereit (2019) « Introduction », *Les marqueurs du discours en diachronie du français : nouvelles perspectives. Studia Linguistica Romanica*, n° 2, 2019, <https://studialinguisticaromanica.org/index.php/slr/article/view/33>.

- Mazziotta N., Bischof B., **Glikman J.**, Rainsford T. (2012) « Ce sujet dans les « constructions impersonnelles » du *Roman de Tristan* de Bérout », *L'Information Grammaticale*, 132, janvier 2012, 48-52.
- Glikman J.** (2009) « Le rapport entre frontières de vers et frontières prosodiques en ancien français », in « Grammaire et Prosodie », *Les Travaux du CerLiCO* n°22, D. Roulland (dir.), 23-37.
- Hassler G., P. Hümmer, J. **Glikman** (2008) « La traduction en allemand de locutions prépositives avec *sous* », *L'Information grammaticale*, n° 117, 35-47.

C-ACTI : Actes de Colloque international (6)

- Glikman J.** & A. Patard (2022) *Être pour aller* en français d'Europe, *SHS Web Conf.* 138 02001 (2022). DOI: 10.1051/shsconf/202213802001
- Glikman J.**, C. Fauth (2022) « Un nouvel accès à la parole spontanée : les vocaux » *34^e Journées d'Études sur la Parole, JEP2022*, Noirmoutier, 13-17 juin 2022.
- Glikman J.**, C. Benzitoun, J.-P. Goldman, Y. Scherrer, M. Avanzi et P. Boula de Mareüil (2018) *Donnez votre Français à la Science ! Internet et la documentation de la diversité linguistique : présentation de la plateforme et premiers résultats*, SHS Web of Conferences 46, 02003 (2018), *Congrès Mondial de Linguistique Française - CMLF 2018*, 23p. <https://doi.org/10.1051/shsconf/20184602003>
- Goldman, J.-P., Y. Scherrer, J. **Glikman**, M. Avanzi, C. Benzitoun, P. Boula de Mareüil (2018) « Crowdsourcing Regional Variables and Automatic Geolocalisation of Speakers of European French. » *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Nicoletta Calzolari, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Koiti Hasida, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hélène Mazo, Asuncion Moreno, Jan Odijk, Stelios Piperidis and Takenobu Tokunaga (ed.), European Language Resources Association (ELRA), Miyasaki, Japon, p. 3336-3342. <http://www.lrec-conf.org/proceedings/lrec2018/summaries/517.html>
- Avanzi, M., C. Barbet, J. **Glikman** & J. Peuvergne (2016) « Présentation d'une enquête pour l'étude des régionalismes du français. » SHS Web of Conferences (Vol. 27, 03001), *5^e Congrès Mondial de Linguistique Française (CMLF)*, F. Neveu, G. Bergounioux, M.-H. Côté, J.-M. Fournier, L. Hriba et S. Prévost (éd.), 14 pages. <https://doi.org/10.1051/shsconf/20162703001>
- Glikman J.** (2008) « Les subordonnées asyndétiques en ancien français », Durand J., Habert B., Laks B. (éds.), *Congrès Mondial de Linguistique Française - CMLF'08*, Paris, Institut de Langue Française, 225-240. <https://doi.org/10.1051/cmlf08206>

ACTN : Communications avec actes dans un congrès national (2)

- Avanzi M., C. Benzitoun, J. **Glikman** (2007) « Comment se comprendre sans se méprendre ? L'exemple de trois termes problématiques : *période, parataxe, subordination inverse* », actes de Coldoc 07, Glikman J., L. Mansour & S. Weiser éds.
- Glikman J.** (2006) « Les propositions subordonnées sans mot subordonnant, étude diachronique », actes du séminaire thématique de sciences du langage *syntaxe*, Nanterre, 5 mai, in *Les Cahiers de l'ED 139, Sciences du langage, 2005-2006*, Publication de l'Université Paris X Nanterre, Nanterre, p. 55-69.

D0 : Directions d'ouvrage ou de revue (9) + (1) CJC

Direction de numéro de revue (5)

- Glikman J.** & T. Verjans (2021) *Regards linguistiques sur l'édition des textes médiévaux, Diachroniques* n°8, Paris, Sorbonne.
- Glikman J.**, G. Parussa & R. Waltereit (éds) (2019) *Les marqueurs du discours en diachronie du français : nouvelles perspectives. Studia Linguistica Romanica*, n° 2, 2019, <https://studialinguisticaromanica.org/index.php/slr/issue/view/2>.
- Glikman J.** & C. Gérard (dir.) (2019) *La linguistique des genres, en actes et en questions, Linx*, 78 | 2019, <https://journals.openedition.org/linx/2892>.
- Schnedecker C., J. **Glikman** & F. Landragin (dir.) (2017) *Les chaînes de référence en corpus, Langue française*, 2017/3 (N° 195). <https://www.cairn.info/revue-langue-francaise-2017-3.htm>
- Glikman J.** & M. Avanzi (dir.) (2012 [2009]) *Entre rection et incidence : des constructions verbales atypiques ? Etudes sur je crois, je pense et autres parenthétiques, Linx* 61-2009. <https://journals.openedition.org/linx/1322>

Direction d'ouvrages collectifs (4)

- Glikman J.**, H. Bat Zeev Shyldkrot, S. Lehmann, F. Sitri & T. Verjans (dir.) (2021) *De la diachronie à la synchronie et vice versa. Mélanges offerts à Annie Bertin par ses étudiants et ses amis en remerciement pour son apport aux études diachroniques*. Chambéry, Presses Universitaires Savoie Mont Blanc.
- Capin D, J. **Glikman**, V. Obry & T. Revol (éds) (2019) *Le Français en diachronie : Moyen français, Segmentation des énoncés, Linguistique textuelle*, Eliphi, TraLiRo, Morphologie, syntaxe, grammaticographie, Strasbourg.
- Ayres Bennet, W., A. Carlier, J. **Glikman**, T. M. Rainsford, G. Siouffi & C. Skupien Dekens (dir.) (2018) *Nouvelles voies d'accès au changement linguistique*, Classiques Garnier. <https://classiques-garnier.com/nouvelles-voies-d-acces-au-changement-linguistique.html>
- Schneider S., **Glikman J.** & M. Avanzi (2015) *Parenthetical Verbs, Linguistische Arbeiten* 557, De Gruyter. <https://doi.org/10.1515/9783110376142>

Direction d'actes de CJC (1)

- Glikman J.**, Mansour L., Weiser S. (2007) *Actes du 4^e quatrième colloque des doctorants et jeunes chercheurs de Paris X Nanterre (Coldoc 07) « Le vocabulaire scientifique et technique en Sciences du Langage »*, actes en ligne sur le site du laboratoire Modyco.

OS-CH : Chapitres d'ouvrage scientifique (22)

Les articles consistant en l'introduction de volumes coédités sont signalés par un astérisque.

- Verjans T. & J. **Glikman** (2022) « Un ou une espèce de changement ? », *Synchronie et diachronie : l'enjeu du sens*, A. Bertin, O. Soutet et T. Ponchon (éds), Paris, Champion, collection Bibliothèque de grammaire et de linguistique n° 67, 423-434.
- ***Glikman J.**, H. Bat Zeev Shyldkrot, S. Lehmann, F. Sitri & T. Verjans (2021) « Introduction », *De la diachronie à la synchronie et vice versa. Mélanges offerts à Annie Bertin par ses étudiants et ses amis en remerciement pour son apport aux études diachroniques*. Chambéry, Presses Universitaires Savoie Mont Blanc, 11-20.

- Glikman J. & T. Verjans** (2021) « *Après moult rebondissements : sur les emplois modernes de moult* », *De la diachronie à la synchronie et vice versa. Mélanges offerts à Annie Bertin par ses étudiants et ses amis en remerciement pour son apport aux études diachroniques*. Glikman J., H. Bat Zeev Shyldkrot, S. Lehmann, F. Sitri & T. Verjans (éds), Chambéry, Presses Universitaires Savoie Mont Blanc, 375-404.
- Glikman J. & B. Bouard** (2021) « La locution conjonctive *pour que* dans l’histoire du français entre usage et construction de la norme : un exemple d’échec des recommandations des remarqueurs », Aquino-Weber, Dorothée, Cotelli Kureth, Sara et Skupien Dekens, Carine (éd.), *La norme du français et sa diffusion dans l’histoire*, Paris, Honoré Champion, p. 143-161.
- Combettes, B. & J. **Glikman** (2020) « Évolution de la phrase complexe », *Grande Grammaire Historique du Français*, C. Marchello-Nizia, B. Combettes, S. Prévost and T. Scheer (dir.), Berlin, De Gruyter, p. 1338-1342.
- Glikman, J.** (2020) « Les subordonnées régies », *Grande Grammaire Historique du Français*, C. Marchello-Nizia, B. Combettes, S. Prévost and T. Scheer (dir.), Berlin, De Gruyter, p. 1342-1359.
- Glikman, J.** (2020) « Les subordonnées finales », *Grande Grammaire Historique du Français*, C. Marchello-Nizia, B. Combettes, S. Prévost and T. Scheer (dir.), Berlin, De Gruyter, p. 1360-1368.
- *Capin D., J. **Glikman**, V. Obry et T. Revol (2019) *Présentation : Le Français en diachronie : Moyen français, Segmentation des énoncés, Linguistique textuelle*, Eliphi, TraLiRo, Strasbourg, p. XI-XVIII.
- Glikman J.** (2019) « Les locutions conjonctives *malgré que* et *à cause que* : normes et usages en diachronie », *Des parlars d’oil à la francophonie. Contact, variation et changement linguistiques*. A. Dufter, K. Grübl, T. Scharinger (eds), *Beihefte zur Zeitschrift für romanische Philologie 440*, De Gruyter, p. 75-96. <https://www.degruyter.com/view/book/9783110541816/10.1515/9783110541816-004.xml>
- Glikman J. & S. Schneider** (2018) « Constructions parenthétiques, marques d’oralité et type de textes en diachronie du français », *Nouvelles voies d’accès au changement linguistique*, W. Ayres Bennet, A. Carlier, J. Glikman, T. M. Rainsford, G. Siouffi et C. Skupien Dekens (dir.), Classiques Garnier, p. 317-334.
- *Ayres Bennet W., A. Carlier, J. **Glikman**, T. M. Rainsford, G. Siouffi & C. Skupien Dekens (2018) « Introduction », *Nouvelles voies d’accès au changement linguistique*, W. Ayres Bennet, A. Carlier, J. Glikman, T. M. Rainsford, G. Siouffi et C. Skupien Dekens (dir.), Classiques Garnier,
- Bouard B. & J. **Glikman** (2018) « Histoire de tout devant adjectif : variation linguistique et élaboration d’une règle (XVIe-XIXe siècles). », *Histoire des langues et histoire des représentations linguistiques*, Sous la direction de Bernard COLOMBAT, Bernard COMBETTES, Valerie RABY et Gilles SIOUFFI, Paris : Honoré Champion Éditeur, p. 293-332.
- Capin D. & J. **Glikman** (2017) « De l’emploi interjectif des conjonctions : le cas de CAR introduisant l’impératif dans les textes médiévaux », *Le français en diachronie, Dépendances syntaxiques, morphosyntaxe verbale, grammaticalisation*, Sophie Prévost & Benjamin Fagard (éd.), Peter Lang, p. 295-316.
- Fagard B., P. Pietrandrea & J. **Glikman** (2016) « Syntactic and semantic aspects of Romance complementizers. », *Complementizers in European Languages*, Kasper Boye & Petar Kehayov (éd.), De Gruyter, p.75-130.

- Glikman J.** (2016) « Subordination et absence de subordonnant en Ancien Français : le cas des relatives », *Subordonnants et subordination à travers les langues*, Hava Bat-Zeev Shyldkrot, Annie Bertin & Olivier Soutet (éd.), Paris, Honoré Champion, p. 85-112.
- Schneider S. & J. **Glikman** (2015) « Origin and development of French parenthetical verbs », in Stefan Schneider, Julie Glikman & Mathieu Avanzi (éd.) *Parenthetical Verbs, Linguistische Arbeiten 557*, De Gruyter, p. 163-188.
- *Schneider S., J. **Glikman** & M. Avanzi (2015) « Introduction », in S. Schneider, J. Glikman & M. Avanzi (éd.) *Parenthetical Verbs, Linguistische Arbeiten 557*, De Gruyter, p. 1-9.
- Glikman J.** & T. Verjans (2013) « Old French Parataxis: syntactic variant or stylistic variation? », *Research on Old French: The state of the research*, D. Arteaga (ed.), *Studies in Natural Language and Linguistic Theory* vol. 88, Springer Netherlands, 243-260.
- Glikman J.** & N. Mazziotta (2013) « Représentation de l'oral et structures syntaxiques dans la prose de *la Queste del saint Graal* (1225-1230) », *Représentation du Sens Linguistique V*, Chambéry, Mai 2011, Editions de l'Université de Savoie, 69-90.
- Glikman J.** (2013) « L'absence de subordonnant en français », *L'absence au niveau syntagmatique*, L. Fesenmeier, A. Grutschus & C. Patzelt eds, *Analecta Romanica Band 80*, 181-199.
- Glikman J.** (2010) « Peut-on établir des critères formels de reconnaissance de la parataxe : l'apport de l'ancien français », *La Parataxe*, tome 1 « Entre dépendance et autonomie », Béguelin M.-J., M. Avanzi & G. Corninboeuf (éds), Berne : Peter Lang, 355-371.
- Glikman J.** (2008) « Les complétives non introduites en ancien français », in Fagard, Benjamin / Prévost, Sophie / Combettes, Bernard / Bertrand, Olivier (éds), *Évolutions en français. Études de linguistique diachronique*, Berne : Peter Lang. 105-118.

PV-A : Articles de vulgarisation (2)

- 2015 - Billet « Des spécialités alsaciennes » sur le blog *Français de nos régions* <https://francaisdenosregions.com/2015/10/13/des-specialites-alsaciennes/>
- 2015 - Billet « Les hommes, les femmes, et la bière » sur le blog *Français de nos régions* <https://francaisdenosregions.com/2015/06/25/les-hommes-les-femmes-et-la-biere/>

CR : Comptes rendus et recensions (3)

- Glikman, J.** 2022. Compte rendu. Anne Carlier, Céline Guillot-Barbance (éds.) 2018. Latin tardif, français ancien. Continuités et ruptures. Berlin: De Gruyter. *Studia linguistica romanica* 2022.7, 126-140. <https://doi.org/10.25364/19.2022.7.6>.
- Glikman J.** (2016) *Le français en diachronie: nouveaux objets et méthodes*. Sous la direction d'Anne Carlier, Michèle Goyens et Béatrice Lamiroy. , *French Studies*, Volume 70, Issue 4, October 2016, Pages 629–630, <https://doi.org/10.1093/fs/knw205>.
- Glikman J.** (2008) « Perret, Michèle, *Introduction à l'histoire de la langue française* », *Cahiers de praxématique* [En ligne], 50 | 2008, document 11, mis en ligne le 01 janvier 2013, consulté le 21 septembre 2020. URL : <http://journals.openedition.org/praxematique/979> ; DOI : <https://doi.org/10.4000/praxematique.979>

Annexe 2 : Table du volume 2 (volume des travaux)

1. Avanzi M., C. Benzitoun, J. Glikman (2007) « Comment se comprendre sans se méprendre ? L'exemple de trois termes problématiques : période, parataxe, subordination inverse », *actes de Coldoc 07*, Glikman J., L. Mansour & S. Weiser (éds).
2. Glikman J. (2008a) « Les complétives non introduites en ancien français », in Fagard, Benjamin / Prévost, Sophie / Combettes, Bernard / Bertrand, Olivier (éds), *Évolutions en français. Études de linguistique diachronique*, Berne : Peter Lang. 105-118.
3. Glikman J. (2008b) « Les subordonnées asyndétiques en ancien français », Durand J., Habert B., Laks B. (éds), *Congrès Mondial de Linguistique Française - CMLF'08*, Paris, Institut de Langue Française, 225-240. <https://doi.org/10.1051/cmlf08206>
4. Hassler G., P. Hümmer, J. Glikman (2008) « La traduction en allemand de locutions prépositives avec sous », *L'Information grammaticale*, n° 117, 35-47.
5. Glikman J. (2009b) « Le rapport entre frontières de vers et frontières prosodiques en ancien français », in « Grammaire et Prosodie », *Les Travaux du CerLiCO n°22*, D. Roulland (dir.), 23-37.
6. Glikman J. (2010a) « Peut-on établir des critères formels de reconnaissance de la parataxe : l'apport de l'ancien français », *La Parataxe, tome 1* « Entre dépendance et autonomie », Béguelin M.-J., M. Avanzi & G. Corminboeuf (éds), Berne : Peter Lang, 355-371.
7. Glikman J. (2010b) « Rattachement et rection en Ancien Français », in « La rection verbale entre micro- et macro-syntaxe », A. Gautier, L. Pino Serrano & D. Van Raemdonck (éds), *Travaux de Linguistique 60*, 29-44. <https://www.cairn.info/revue-travaux-de-linguistique-2010-1-page-29.htm>
8. Glikman J. (2012 [2009]) « Les incisives en croire, cuidier, penser en Ancien Français », in numéro thématique « Entre rection et incidence : des constructions verbales atypiques ? Études sur je crois, je pense et autres parenthétiques », M. Avanzi & J. Glikman (éds), *Linx 61*, 71-85. <https://journals.openedition.org/linx/1337>
9. Mazziotta N., Bischof B., Glikman J., Rainsford T. (2012) « Ce sujet dans les « constructions impersonnelles » du Roman de Tristan de Béroul », *L'Information Grammaticale*, 132, janvier 2012, 48-52.
10. Bischof, B.-B., Glikman, J., Mazziotta, N., Prévost, S., Stein, A. (2012) Guide d'annotation du Syntactic Reference Corpus of Medieval French (SRCMF). Consultation en ligne uniquement: <http://srcmf.org/fiches>
11. Glikman J. (2013) « L'absence de subordonnant en français », *L'absence au niveau syntagmatique*, L. Fesenmeier, A. Grutschus & C. Patzelt (éds), *Analecta Romanica Band 80*, 181-199.
12. Glikman J. & N. Mazziotta (2013) « Représentation de l'oral et syntaxe dans la prose de la Queste del saint Graal (1225-1230) », D. Lagorgette, P. Larrivée (éds) *Représentation du Sens Linguistique 5*, Chambéry, Presses de l'Université de Savoie, 43-64.

13. Glikman J. & T. Verjans (2013) « Old French Parataxis: syntactic variant or stylistic variation? », *Research on Old French: The state of the research*, D. Arteaga (ed.), Studies in Natural Language and Linguistic Theory vol. 88, Springer Netherlands, 243-260.
14. Glikman J., C. Guillot-Barbance, V. Obry (2014) « Les chaînes de référence dans un corpus de textes narratifs médiévaux : traits généraux et facteurs de variation », in *Langages* 195, 43-60. <https://www.cairn.info/revue-langages-2014-3-page-43.htm>
15. Schneider S. & J. Glikman (2015) « Origin and development of French parenthetical verbs », in Stefan Schneider, Julie Glikman & Mathieu Avanzi (éds) *Parenthetical Verbs*, Linguistische Arbeiten 557, De Gruyter, p. 163-188.
16. Avanzi, M., C. Barbet, J. Glikman & J. Peuvergne (2016) « Présentation d'une enquête pour l'étude des régionalismes du français. » SHS Web of Conferences (Vol. 27, 03001), *5e Congrès Mondial de Linguistique Française (CMLF)*, F. Neveu, G. Bergounioux, M.-H. Côté, J.-M. Fournier, L. Hriba et S. Prévost (éds), 14 pages. <https://doi.org/10.1051/shsconf/20162703001>
17. Fagard B., P. Pietrandrea & J. Glikman (2016) « Syntactic and semantic aspects of Romance complementizers. », *Complementizers in European Languages*, Kasper Boye & Petar Kehayov (éds), De Gruyter, p.75-130.
18. Glikman J. (2016) « Subordination et absence de subordonnant en Ancien Français : le cas des relatives », *Subordonnants et subordination à travers les langues*, Hava Bat-Zeev Shyldkrot, Annie Bertin & Olivier Soutet (éds), Paris, Honoré Champion, p. 85-112.
19. Capin D. & J. Glikman (2017) « De l'emploi interjectif des conjonctions : le cas de CAR introduisant l'impératif dans les textes médiévaux », *Le français en diachronie, Dépendances syntaxiques, morphosyntaxe verbale, grammaticalisation*, Sophie Prévost & Benjamin Fagard (éds), Peter Lang, p. 295-316.
20. Obry V., J. Glikman, C. Guillot-Barbance, B. Pincemin (2017) « Les chaînes de référence dans les récits brefs en français : étude diachronique (XIIIe – XVIe s.). » *Langue française*, 2017/3 (N° 195), p.91 - 110. <https://www.cairn.info/revue-langue-francaise-2017-3-page-91.htm?contenu=resume>
21. Bouard B. & J. Glikman (2018) « Histoire de tout devant adjectif : variation linguistique et élaboration d'une règle (XVIe-XIXe siècles). », *Histoire des langues et histoire des représentations linguistiques*, Sous la direction de Bernard Colombat, Bernard Combettes, Valérie Raby et Gilles Siouffi, Paris : Honoré Champion Éditeur, p. 293-332.
22. Glikman J., C. Benzitoun, J.-P. Goldman, Y. Scherrer, M. Avanzi et P. Boula de Mareüil (2018) Donnez votre Français à la Science ! Internet et la documentation de la diversité linguistique : présentation de la plateforme et premiers résultats, *SHS Web of Conferences* 46, 02003 (2018), *Congrès Mondial de Linguistique Française - CMLF 2018*, 23p. <https://doi.org/10.1051/shsconf/20184602003>
23. Glikman J. & S. Schneider (2018) « Constructions parenthétiques, marques d'oralité et type de textes en diachronie du français », *Nouvelles voies d'accès au changement linguistique*, W. Ayres Bennet, A. Carlier, J. Glikman, T. M. Rainsford, G. Siouffi et C. Skupien Dekens (dir.), Classiques Garnier, p. 317-334.

24. Goldman, J.-P., Y. Scherrer, J. Glikman, M. Avanzi, C. Benzitoun, P. Boula de Mareüil (2018) « Crowdsourcing Regional Variables and Automatic Geolocalisation of Speakers of European French. » *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Nicoletta Calzolari, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Koiti Hasida, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, H el ene Mazo, Asuncion Moreno, Jan Odijk, Stelios Piperidis and Takenobu Tokunaga ( eds), *European Language Resources Association (ELRA)*, Miyasaki, Japon, p. 3336-3342. <http://www.lrec-conf.org/proceedings/lrec2018/summaries/517.html>
25. Oberl e B., C. Schnedecker, E. Baumer, D. Capin, J. Glikman, C. Guo, Th. Revol, A. Todirascu, J. Tushkova (2018) « Les cha enes de r ef erence dans les textes encyclop ediques du 12e au 21e si ecle :  tude longitudinale. », *Travaux de Linguistique*, 2018/2 n o 77, p. 67 – 141. DOI : 10.3917/tl.077.0067. URL : <https://www.cairn.info/revue-travaux-de-linguistique-2018-2-page-67.htm>.
26. Glikman J. (2019) « Les locutions conjonctives malgr e que et  a cause que : normes et usages en diachronie », *Des parlers d’oil  a la francophonie. Contact, variation et changement linguistiques*. A. Dufter, K. Gr ubl, T. Scharinger (eds), Beihefte zur Zeitschrift f ur romanische Philologie 440, De Gruyter, p. 75-96. <https://www.degruyter.com/view/book/9783110541816/10.1515/9783110541816-004.xml>
27. Mazziotta N. et J. Glikman (2019) « Oral repr esent e et narration en ancien fran ais. Sp ecificit es syntaxiques dans trois textes de genres distincts », *Linx*, 78 | 2019, <https://journals.openedition.org/linx/3151>.
28. Combettes, B. & J. Glikman (2020) «  volution de la phrase complexe », *Grande Grammaire Historique du Fran ais*, C. Marchello-Nizia, B. Combettes, S. Pr evost and T. Scheer (dir.), Berlin, De Gruyter, p. 1338-1342.
29. Glikman, J. (2020a) « Les subordonn es finales », *Grande Grammaire Historique du Fran ais*, C. Marchello-Nizia, B. Combettes, S. Pr evost and T. Scheer (dir.), Berlin, De Gruyter, p. 1360-1368.
30. Glikman, J. (2020b) « Les subordonn es r egies », *Grande Grammaire Historique du Fran ais*, C. Marchello-Nizia, B. Combettes, S. Pr evost and T. Scheer (dir.), Berlin, De Gruyter, p. 1342-1359.
31. Glikman J. & B. Bouard (2021) « La locution conjonctive pour que dans l’histoire du fran ais entre usage et construction de la norme : un exemple d’ echec des recommandations des remarqueurs », Aquino-Weber, Doroth e, Cotelli Kureth, Sara et Skupien Dekens, Carine ( ed.), *La norme du fran ais et sa diffusion dans l’histoire*, Paris, Honor e Champion, p. 143-161.
32. Glikman J. & T. Verjans (2021) « Apr es moult rebondissements : sur les emplois modernes de moult », *De la diachronie  a la synchronie et vice versa. M elanges offerts  a Annie Bertin*. Glikman J., H. Bat Zeev Shyldkrot, S. Lehmann, F. Sitri & T. Verjans (dir.), Chamb ery, Presses Universitaires Savoie Mont Blanc, p. 375-404.
33. Capin D., J. Glikman, C. Schnedecker & A. Todirascu (2021) « Le r ole des cha enes de r ef erence dans la structuration textuelle :  tude diachronique de l’ancien fran ais au

- français moderne », *Langages* n°224, p. 87-107. DOI : 10.3917/lang.224.0087. URL : <https://www.cairn.info/revue-langages-2021-4-page-87.htm>.
34. Glikman J., C. Fauth (2022) « Un nouvel accès à la parole spontanée : les vocaux » *34e Journées d'Études sur la Parole*, JEP2022, Noirmoutier, 13-17 juin 2022.
35. Glikman J. & A. Patard (2022) « Être pour aller en français d'Europe », *CMLF 2022, SHS Web Conf.* 138 02001 (2022). DOI: 10.1051/shsconf/202213802001
36. Verjans T. & J. Glikman (2022) « Un ou une espèce de changement ? », *Synchronie et diachronie : l'enjeu du sens*, Champion, Paris, collection Bibliothèque de Grammaire et de Linguistique N°67, sous la responsabilité d'Annie Bertin, Thierry Ponchon et Olivier Soutet, p. 423-434.

Annexes : introductions et tables des matières des volumes et numéros de revue édités

- i. Glikman J., Mansour L., Weiser S. (éds) (2007) *Actes du 4e quatrième colloque des doctorants et jeunes chercheurs de Paris X Nanterre (Coldoc 07)* « Le vocabulaire scientifique et technique en Sciences du Langage », actes en ligne sur le site du laboratoire Modyco.
- ii. Glikman J. & M. Avanzi (dir.) (2012 [2009]) *Entre rection et incidence : des constructions verbales atypiques ? Etudes sur je crois, je pense et autres parenthétiques*, *Linx* 61-2009. <https://journals.openedition.org/linx/1322>
- iii. Schneider S., Glikman J. & M. Avanzi (éds) (2015) *Parenthetical Verbs*, *Linguistische Arbeiten* 557, De Gruyter. <https://doi.org/10.1515/9783110376142>
- iv. Schnedecker C., J. Glikman & F. Landragin (dir.) (2017) Les chaînes de référence en corpus, *Langue française*, 2017/3 (N° 195). <https://www.cairn.info/revue-langue-francaise-2017-3.htm>
- v. Ayres Bennet, W., A. Carlier, J. Glikman, T. M. Rainsford, G. Siouffi & C. Skupien Dekens (dir.) (2018) *Nouvelles voies d'accès au changement linguistique*, Classiques Garnier. <https://classiques-garnier.com/nouvelles-voies-d-acces-au-changement-linguistique.html>
- vi. Capin D, J. Glikman, V. Obry & T. Revol (éds) (2019) *Le Français en diachronie : Moyen français, Segmentation des énoncés, Linguistique textuelle*, Eliphi, TraLiRo, Morphologie, syntaxe, grammaticographie, Strasbourg.
- vii. Glikman J. & C. Gérard (dir.) (2019) La linguistique des genres, en actes et en questions, *Linx*, 78 | 2019, <https://journals.openedition.org/linx/2892>.
- viii. Glikman J., G. Parussa & R. Waltereit (éds) (2019) Les marqueurs du discours en diachronie du français : nouvelles perspectives. *Studia Linguistica Romanica*, n° 2, 2019, <https://studialinguisticaromanica.org/index.php/slr/issue/view/2>.
- ix. Glikman J. & T. Verjans (dir.) (2021) Regards linguistiques sur l'édition des textes médiévaux, *Diachroniques* n°8, Paris, Sorbonne.
- x. Glikman J., H. Bat Zeev Shyldkrot, S. Lehmann, F. Sitri & T. Verjans (dir.) (2021) *De la diachronie à la synchronie et vice versa. Mélanges offerts à Annie Bertin*. Chambéry, Presses Universitaires Savoie Mont Blanc.