

Linguistique variationnelle historique et traitement de textes : présentation d'une méthode d'analyse

Harald Völker (Université de Göttingen)

1. L'histoire du français standard et la recherche scriptologique

La grammaire du français standard, telle que nous la connaissons aujourd'hui, est le résultat d'un processus historique de standardisation linguistique. Pour le français, ce processus de normalisation est bien connu et bien documenté à partir du XVI^e et surtout du XVII^e siècle. Pour la période antérieure, les faits se présentent de façon plus opaque. Certes, nous connaissons les prises de positions contemporaines, souvent citées, par rapport à la qualité et à la réputation des idiomes de la France moyen-âgeuse: Roger Bacon, Conon de Béthune, Aimon de Varennes et d'autres.¹ Ces manifestations soulèvent pourtant un problème méthodologique, puisque ces opinions sont en réalité des métaévidences proto-scientifiques au delà de la terminologie géo-linguistique nette et bien définie dont nous disposons aujourd'hui. Quand Roger Bacon fait par exemple la distinction entre l'idiome de la Picardie, celui de la Normandie, celui de la Bourgogne, celui de Paris et celui qu'il appelle *Gallicum*, il nous laisse dans l'incertitude quant à l'étendue exacte de ces signifiants, surtout du dernier.²

¹ Cités et commentés par exemple dans Winkelmann 1990, 336-338, Lodge 1993, 95-104, et Dieckmann 2000.

² Voir Lodge 1993, 97.

C'est en particulier avec Pfister 1993 qu'on a commencé à intégrer les résultats et les méthodes de la scriptologie dans la recherche des origines du français standard. Ceci a été le point de départ pour les travaux effectués dans le cadre du projet de recherche « Westmitteldeutsche und ostfranzösische Urkunden- und Literatursprachen im 13. und 14. Jahrhundert » ('Ancien français de l'est et moyen haut allemand de l'ouest dans les chartes et dans la littérature des XIII^e et XIV^e siècles') qui a été entrepris aux universités de Trier (Trèves) et de Göttingen dans le cadre de l'Unité de Recherches Interdisciplinaires 235 « Zwischen Maas und Rhein. Beziehungen, Begegnungen und Konflikte in einem europäischen Kernraum » ('Entre Meuse et Rhin: rapports, rencontres et conflits dans une région centrale de l'Europe'). Ce projet de recherche interdisciplinaire a été subventionné entre 1995 et 2001 par la Deutsche Forschungsgemeinschaft, le fonds national de la recherche en Allemagne. Plusieurs publications issues de ces travaux sont parues à ce jour: La monographie Holtus/Overbeck/Völker 2003b « Luxemburgische Skriptastudien » ('Études scriptologiques sur le Luxembourg'), la thèse Völker 2003 « Skripta und Variation. Untersuchungen zur Negation und zur Flexion des Substantivs in den altfranzösischen Urkunden Gräfin Ermesindes und Graf Heinrichs V. von Luxemburg, 1237-1281 » ('Scripta et variation. Recherches sur la négation et la déclinaison du substantif dans l'ancien français des chartes de la comtesse Ermesinde et du comte Henri V de Luxembourg, 1237-1281'), qui met en relief les aspects méthodologiques, et la thèse Overbeck 2003b « Literarische Skripta in Ostfrankreich » ('Scripta littéraire dans l'est de la France'), qui applique la démarche méthodologique de Trèves au domaine littéraire; en outre, un volume recueillant des articles du début des travaux (Gärtner/Holtus 1995), trois actes de colloques tenus à Trèves (Gärtner/Holtus 1997, Gärtner/Holtus/Rapp/Völker 2001 et Gärtner/Holtus sous presse) et un certain nombre d'articles parus dans des journaux, dans des festschriften et dans des actes de colloque.³

Le but de cet article est dans un premier temps de présenter la démarche méthodologique entreprise dans les travaux de ce projet. Dans un deuxième temps, nous tenterons de démontrer l'importance de certains de nos résultats pour la recherche de l'origine du français standard.

2. La méthodologie et sa mise en œuvre

2.1. La linguistique variationnelle

Même si la linguistique variationnelle ne peut passer pour établie et reconnue que depuis les années quatre-vingt, sa genèse date de trente ans plus tôt. Le premier à

³ Par exemple Holtus/Körner/Völker 2001, Holtus/Overbeck/Völker 2003a, Holtus/Völker 1999, Holtus/Völker 2004, Overbeck 2003a, Völker 2000, Völker 2004 et Völker 2005. Une bibliographie des publications issues du projet de Trèves est disponible sous <RM.net> (<http://gepc189.uni-trier.de/cgi-bin/iRMnet/RMnetIndex.tcl?&nav=su&cnt=supr&cid=32&tpid=D.7A&proj=SFB235&hea=&for=&mode=flitdb>).

distinguer des « coexistences structurales régulières » au sein des langues est le Norvégien Leiv Flydal (1952, 255), qui a en même temps créé la notion d'« architecture de langue » (ib. 257). Peu de temps après Flydal, Uriel Weinreich 1954 (cf. 389s.) a introduit les notions de « variety » et de « diasystem ». Eugenio Coseriu (1970, en particulier 32-37) a élargi et précisé les instruments terminologiques en formant les premiers adjectifs « dia- », comme diatopique, diastratique et diaphasique. Ludwig Söll 1974 et Bodo Müller 1975 ont été les premiers à utiliser la nouvelle terminologie dans le cadre de la recherche pratique, à savoir dans leurs descriptions du français contemporain.⁴

La linguistique variationnelle part du principe que les langues naturelles ne sont pas des systèmes autonomes, mais qu'elles dépendent au contraire de facteurs extralinguistiques. Cette branche de la linguistique est donc basée sur les mêmes fondements que la pragmatique linguistique et la sociolinguistique. Cependant, la linguistique variationnelle est beaucoup plus liée aux faits et aux structures intralinguistiques que ces deux dernières disciplines, parce que son but est non seulement de décrire la variation de la *parole*, mais aussi de rechercher les régularités normatives caractéristiques des différentes normes diasystématiques. Les « variétés » ou bien « subsystemes » se distinguent en effet à la fois par leurs traits linguistiques caractéristiques et par des facteurs extralinguistiques aisément détectables (comme par exemple la provenance géographique du locuteur ou bien son statut social) dont le rapport avec la variation linguistique est évident. La linguistique variationnelle offre donc la possibilité de systématiser les différentes catégories variationnelles (telles que sociale, dialectale, pragmatique etc.) en fonction de leur impact sur la structure interne du langage.

La réalité linguistique – voilà le credo de la linguistique variationnelle – ne consiste pas en un système idéal et indépendant des contextes et de son utilisation. La réalité linguistique est plutôt déterminée par la coexistence et l'interaction dynamiques des différentes variétés d'une langue. Dans cette perspective, le standard est défini comme une variété qui ne dépend pas ou presque pas de facteurs extralinguistiques, mais il n'est, lui aussi, qu'une variété parmi d'autres.⁵

Pour des raisons évidentes, la linguistique variationnelle devrait avoir beaucoup d'importance pour le travail en diachronie linguistique : dans la plupart des langues, le standard ne joue un rôle prééminent au sein de l'architecture linguistique que depuis relativement peu de temps. À l'époque de l'ancien français, les dialectes et les *scriptae* régionales avaient une position beaucoup plus forte qu'aujourd'hui; il n'existait guère de standard linguistique. Mais le problème des variétés linguistiques dépasse la géographie: pour ne pas restreindre l'interprétation des variantes et de leur importance statistique au nom d'*a priori*, nous avons décidé de ne pas nous contenter de l'analyse de la simple variation diatopique, mais de tenir compte en outre de trois autres catégories de variation dont les données

⁴ Voir Holtus/Radtke 1983 et iid. 1986-1990 pour la portée des questions traitées par la linguistique variationnelle. Comme introduction à la thématique, on se servira de Holtus 1978, et 1992, Prüßmann-Zemper 1990 ainsi que de Thelen 1999, 7-28.

⁵ Voir Müller 1975, 5 et 34s., ainsi que Prüßmann-Zemper 1990, 830s.

extralinguistiques pouvaient être déterminées sans efforts démesurés, à savoir la variation diachronique, diastratique et idiolectale.⁶ Les raisons pour lesquelles il s'agit là d'une innovation, c'est-à-dire les raisons pour lesquelles la linguistique variationnelle n'a pas encore pu engendrer des méthodes de recherche répandues en diachronie, sont multiples et, jusqu'à un certain point, compréhensibles :

- Beaucoup d'archives ont été dévastées par des guerres et par d'autres catastrophes si bien que l'on rencontre toutes sortes de difficultés lorsque l'on veut constituer des corpus concis et assez importants. Guyotjeannin/Pycke/Tock (1993, 355) soulignent que les archives souvent n'ont pu garder des séries de documents que des « blocs erratiques ».
- S'ajoute à cela que les éditions de textes, même de textes d'archives, sont souvent soumises à des normalisations. Ceci s'explique très bien par le but de la plupart des éditions: rendre le contenu des textes moyen-âgeux accessible au lecteur moderne. Mais pour nous autres linguistes, c'est justement cela qui pose problème: toute intervention, toute normalisation nuit au linguiste variationniste, puisque la variante, la faute, la spécificité est la matière première sans laquelle il ne peut travailler.⁷
- Enfin, les données extralinguistiques, qui, dans le travail en synchronie contemporaine, se livrent d'elles-mêmes au chercheur, doivent être, en diachronie, l'objet d'une enquête historique minutieuse. Et, dans bien des cas, il arrive que ces données ne soient tout simplement plus accessibles.

2.2. La constitution du corpus

Le corpus que nous avons examiné à Trèves englobe les 180 chartes originales écrites en ancien français reçues ou expédiées entre 1237 et 1281 par la comtesse Ermesinde de Luxembourg et son fils, le comte Henri Blondel. Si nous avons choisi les comtes du Luxembourg, c'est pour les raisons suivantes:

- Le comté de Luxembourg avoisine les régions mosellanes examinées par nos collègues germanistes.
- Le comté de Luxembourg a été étudié de manière approfondie, surtout en ce qui concerne le XIII^e siècle, par des collègues historiens luxembourgeois et allemands, travaillant eux aussi dans l'Unité de Recherches Interdisciplinaires de Trèves.
- Dans les années 40 et 50 du dernier siècle, les chartes luxembourgeoises ont été inventoriées de façon exemplaire par l'historien luxembourgeois Camille Wampach.
- L'édition des chartes publiée par Wampach ne suffit pourtant pas aux besoins linguistiques.

⁶ Voir Völker 2003, 88-102, pour la réflexion théorique menant à cet élargissement de perspective.

⁷ Pour le différend sur les principes d'édition faisant suite à la *New Philology*, voir Holtus/Völker 1999 et Overbeck 2003a.

La matière première de toute analyse variationnelle, nous venons de le dire, est la variante. Il est donc évident que notre projet de recherche s'inspire, entre autres, de l'*Éloge de la variante* et de la *New Philology*. Cerquiglini 1989 et la *New Philology* ont ouvert la voie à un débat nécessaire, qui nous a permis, à nous autres linguistes variationnistes, étymologues, lexicologues etc. de nous communiquer mutuellement nos besoins.

Avant de nous rendre dans des Archives luxembourgeoises, belges, françaises et allemandes pour copier les 180 chartes choisies, nous avons établi des règles pour la transcription des chartes. Au cours de notre première lecture des chartes, nous avons complété ces principes qui suivent d'ailleurs plus ou moins celles des *Documents linguistiques de la France*. Pourtant, pour suffire à nos besoins linguistiques, nous nous sommes imposé, dans certains cas, des règles encore plus strictes, à savoir...

- ...maintenir toutes les variantes du texte même lorsqu'il semble s'agir de fautes évidentes
- ...maintenir toutes les variantes graphiques dont l'utilisation peut indiquer l'appartenance à une variété régionale, sociale ou bien idiolectale (comme par exemple lettres minuscules vs. lettres majuscules)
- ...mettre les abréviations écrites en toutes lettres entre parenthèses pour pouvoir les exclure de certaines analyses (p.ex. morphologique, graphématique) mais pour pouvoir quand même en tenir compte dans les analyses lexicologiques.

Voici quelques détails de nos principes d'édition:⁸

1. Abréviations

Les abréviations sont transcrites en toutes lettres; cette intervention est marquée par des parenthèses. Pour l'interprétation des abréviations, nous nous sommes reportés aux graphies développées dans d'autres parties de la même charte ou dans d'autres chartes du même scribe; en cas d'hésitation, nous proposons une variante dans l'apparat. Étant donné le caractère aléatoire des transcriptions d'abréviation, celles-ci sont à exclure de l'analyse scriptologique des chartes.

2. Formes graphiques des lettres

Quand, lors de notre première lecture des chartes, nous avons remarqué que la forme graphique d'une lettre pourrait être pertinente pour la définition d'une variété, nous avons décidé de conserver les différentes formes graphiques de cette lettre. Cela signifie que:

- nous maintenons la distinction originale entre *u* et *v*.
- nous maintenons la distinction originale entre *i* et *j*.

⁸ Version abrégée de Holtus/Overbeck/Völker 2003a, 116s. La liste entière des principes d'édition peut être consultée dans Holtus/Overbeck/Völker 2003b, 6-14.

- nous faisons la différence entre *s* rond et *ſ* long.
- en outre, les majuscules et les minuscules sont différenciées suivant leur apparition dans les originaux.

3. Séparation des mots

Nous avons renoncé à introduire des séparations modernes de mots. La séparation des mots nous paraît être un trait spécifique de la scripta individuelle d'un scribe, voire d'un style de chancellerie; la conserver peut aider à identifier des mains de scribes et des chancelleries.

4. Ponctuation

Nous maintenons tous les signes de ponctuation présents dans les chartes. L'analyse de la pertinence variationnelle et surtout syntaxique de la ponctuation médiévale n'a jusqu'ici pas encore pu être commencée, puisque presque toutes les éditions de textes médiévaux, y compris les *Documents linguistiques de la France*, renoncent à respecter la ponctuation originale.⁹ Le signe le plus fréquent est le point volant (·) qui segmente la phrase et encadre des chiffres romains. Si le point se trouve sur la ligne de base, nous le représentons comme dans l'original (·). Un autre signe est la virgule, qui consiste dans la plupart des cas en un trait oblique montant de gauche à droite (/). Moins fréquents, le *punctus elevatus* (point-virgule renversé) et les diverses combinaisons de points sont, pour des raisons de typographie, commentés dans des notes.

5. Fautes

Les fautes évidentes ne sont en aucun cas corrigées dans la transcription, car elles représentent souvent des particularités dialectales ou individuelles des scribes. Un *sic* suivi d'un essai d'explication en note signale qu'il s'agit d'une faute évidente.

2.3. La recontextualisation catégorisée

Comme nous l'avons vu, nous cherchons en linguistique variationnelle à découvrir les corrélations entre variantes intralinguistiques et cadre extralinguistique. Les analyses nécessaires peuvent très bien se faire à l'aide d'un ordinateur, à condition que l'ordinateur puisse digérer les données extra-linguistiques. Pour cela il faut lui mettre ces données à disposition de manière catégorisée. Tout en sachant que la liste suivante peut et devrait être élargie dans l'avenir, nous avons décidé de commencer par quatre catégories qui, en linguistique variationnelle, sont connues sous les noms de « diachronique », « diatopique », « diastratique » et « idiolectale ».

⁹ Voir Monfrin 1974, LXIV: « La ponctuation que l'on rencontre dans les actes originaux est tout à fait sommaire et ne semble pas obéir à des règles précises ». Une analyse modèle de la ponctuation originale d'une chartre nous est livrée par Boutier 2001.

Dans l'investigation de ces données extralinguistiques, dont nous avons besoin pour catégoriser les chartes, nous avons pu profiter à maints égards de la coopération interdisciplinaire au sein de l'Unité de Recherche de Trèves.

1. La datation des chartes s'est faite sans grands problèmes, puisque toutes nos chartes indiquent en clause finale, ce qui ne surprend pas, leur datation exacte.¹⁰
2. La « localisation » des chartes: Voilà un problème méthodologique qui mérite un intérêt approfondi. Dans la tradition scriptologique, les chartes passent en général pour facilement localisables. Mais il n'en est pas ainsi si l'on cherche à éviter les cercles vicieux qui résultent de la localisation intralinguistique (basée sur des « traits dialectaux »). Si, de plus, on ne se contente pas de la localisation de la production de la charte et que l'on cherche à localiser l'acte communicatif représenté par le texte, de nouveaux facteurs entrent en ligne de compte. Comme les chartes de petite portée communicative présentent certainement une plus grande précision diatopique que les chartes de grande portée communicative, nous avons par exemple tenu compte de la portée administrative des chartes.¹¹
3. Pour le statut social des personnes impliquées dans l'affaire juridique ou politique documentée par la charte, nous avons établi trois catégories ; 1^{er} niveau : petite noblesse, bas clergé et bourgeoisie (groupe A) ; 2^e niveau : haute noblesse et haut clergé (groupe B) ; 3^e niveau : royal/impérial (groupe C). La recherche de ces données a été plutôt facile, puisque les acteurs politiques au sein et autour du comté de Luxembourg au XIII^e siècle sont bien documentés grâce aux travaux de nos collègues historiens.¹²
4. Pour établir des séries de mains de scribes, nous nous sommes servi d'une liste de critères paléographiques calquée sur le modèle de la paléographie néerlandaise, pays à la pointe de la méthodologie pour l'identification de scribes. 24 séries de scribe ont pu être identifiées à l'aide de cette méthode, une série à neuf chartes, une à huit, une à sept, trois à cinq, deux à quatre, quatre à trois et douze à deux chartes.¹³

Ces données ont été entrées dans un fichier dit « Segmentdatei » (‘fichier segmenté’), une spécificité du logiciel TUSTEP qui permet d'ajouter au texte-même un metatexte (ou bien hypertexte)¹⁴ sous forme de base de données relationnelle, reliée directement au texte, mais reliée en même temps aux autres facteurs extralinguistiques.

¹⁰ Pour plus de détails, voir Völker 2003, 128-130.

¹¹ Pour plus de détails, voir Völker 2003, 134-142, et Holtus/Völker 2004.

¹² Pour plus de détails, voir Völker 2003, 150s.

¹³ Pour plus de détails, voir Holtus/Overbeck/Völker 2003b, 17-43, et Völker 2003, 157-162.

¹⁴ Voir Völker 2005.

2.4. Choix et balisage des unités d'analyse

Les variétés d'une langue – voilà pourquoi elles peuvent être regroupées sous le toit d'une et seule langue – partagent un fonds d'éléments et de structures linguistiques. Pour une analyse qui vise à distinguer les différentes variétés d'une langue, il faut donc choisir des éléments intralinguistiques qui sont susceptibles de varier et non ceux qui laissent prévoir une grande stabilité à travers les variétés. Nous nous sommes décidé entre autres pour des unités linguistiques qui étaient soumises à des restructurations au XIII^e siècle. Comme le changement linguistique ne se produit pas d'un seul coup, mais, d'une variété à l'autre, en différentes étapes et à différentes vitesses, il s'agit là, pour reprendre le mot célèbre de Paul Imbs (dont il s'est servi, lui, dans un autre contexte) d'éléments malléables dont l'apparence est souvent reliée aux influences extralinguistiques.¹⁵ Voilà pourquoi nous avons choisi, entre autres, la négation et le système bicasuel comme objets de recherche. Au sein d'un fichier TUSTEP, les variantes des éléments choisis ont été balisées et complétées de toute information nécessaire pour l'analyse automatique.¹⁶

2.5. Analyses multifactorielles

Procédons au programme d'analyse. Comme d'habitude en informatique, le programme que nous avons écrit¹⁷ a l'air d'être compliqué, mais il ne l'est point quand on en connaît l'idée et le principe. L'idée est simple et provient du célèbre Louis Remacle, père de la notion de scripta: « Si on allait y voir? »¹⁸

Le principe consiste à relier les variantes des éléments balisés aux facteurs extra-linguistiques pour savoir si l'on trouve dans l'ensemble des chartes des corrélations nettes et évidentes entre une variante déterminée et un facteur extra-linguistique. Dans une telle démarche, l'avantage des chartes est évident: les chartes nous livrent, de par leur « Sitz im Leben » (leur rôle dans la vie réelle), plus d'informations fiables sur les conditions politiques et socioculturelles, dans lesquelles l'acte communicatif s'est joué. Sur cette base, il a été possible de voir quels facteurs extralinguistiques sont liés à quelle variante, quelle innovation ou bien à quelle forme hypercorrecte.¹⁹

¹⁵ Voir Imbs 1961, 138 : « Tous ceux qui sont familiarisés avec l'ancien français savent d'expérience à quel point les mots d'ancien français sont proches de la parole, c'est-à-dire de l'incarnation dans des phrases particulières, des conditions concrètes de leur emploi [...]. De là les difficultés des définitions de type logique par simple indication du genre et de la différence spécifique. Le mot médiéval est un mot malléable et souple, et c'est ici, dans le vocabulaire médiéval, que la description phénoménologique me paraît tout particulièrement à sa place ».

¹⁶ Pour les détails du balisage voir Völker 2003, 103-128 et 169-182, ainsi qu'Overbeck/Völker sous presse.

¹⁷ Le programme est présenté de manière plus détaillée dans Völker 2004 et Overbeck/Völker sous presse.

¹⁸ Remacle 1948, 14.

¹⁹ Sur le rôle des analogies et en particulier des formes hypercorrectes dans l'interprétation des variantes, voir Holtus/Körner/Völker 2001.

3. L'origine du standard à la lumière des résultats du projet de recherche

Une nouvelle dimension de notre démarche a consisté à élargir la perspective diasystématique. Tandis que la plupart des travaux scriptologiques considèrent la variation linguistique uniquement comme une variation diatopique, nous avons ajouté aux analyses diatopiques des analyses diachroniques et diastratiques et nous nous sommes intéressés en plus au rôle des scribes. Ainsi, nous avons pu constater que les chartes échangées entre le roi de France et le comte de Luxembourg se distinguent sensiblement des autres chartes du corpus. Tandis que, par exemple, dans la majorité des chartes du corpus le système bicasuel se voit toujours bien (voire très bien) respecté, dans les documents royaux²⁰ se manifeste déjà le début d'un système unicasuel.

<i>Système bicasuel</i> ²¹				
	Cas régime, occurrences...		Cas sujet, occurrences...	
	...conformes	...non conformes	...conformes	...non conformes
total	5999 (99,40%)	37 (0,60%)	1445 (83,80%)	280 (16,20%)
A ²²	1881 (98,70%)	24 (1,30%)	512 (99,00%)	5 (1,00%)
B ²³	3490 (99,70%)	11 (0,30%)	882 (99,00%)	9 (1,00%)
C ²⁴	628 (99,70%)	2 (0,30%)	51 (16,10%)	266 (83,90%)

Tab. 1: Le système bicasuel

Il en va de même pour la négation avec particule supplémentaire. La forme qui l'a emporté en français moderne est plus fréquente dans les chartes du groupe C que dans les autres chartes:

²⁰ On peut considérer comme assuré que ces documents ont été rédigés sous la responsabilité de la chancellerie royale ; voir Völker 2000.

²¹ Voir Völker 2003, 188s., et, pour la définition de „conforme/non conforme“, ib. 169-181.

²² Groupe niveau petite noblesse, bas clergé et bourgeoisie.

²³ Groupe niveau haute noblesse et haut clergé.

²⁴ Groupe niveau royal/impérial.

<i>Négation en...</i> ²⁵				
	...ne mie	...ne pas	...ne point	...ne nient (adv.)
total	37 (50,00%)	24 (32,40%)	12 (16,20%)	1 (1,40%)
A	7 (77,80%)	1 (11,10%)	1 (11,10%)	0 (0,00%)
B	22 (75,90%)	3 (10,30%)	3 (10,30%)	1 (3,50%)
C	8 (22,20%)	20 (55,60%)	8 (22,20%)	0 (0,00%)

Tab. 2: La négation

À la lumière de notre corpus, ce n'est donc pas la périphérie géographique qui a abandonné une règle grammaticale trop compliquée,²⁶ mais bien le centre social du pouvoir politique qui a poussé à la simplification. Ces résultats sont d'autant plus importants que les chartes en question proviennent de différentes mains de scribes. Il peut donc être exclu, qu'il s'agisse d'un trait linguistique purement individuel.²⁷

En outre, il est important de ne pas confondre la position sociale de la chancellerie royale avec la situation géographique de Paris ou de l'Île-de-France. Ce qui est sûr, c'est que la royauté, et par conséquent la scripta des chartes en question, est un fait social (diastatique, peut-être aussi diatechnique); ce qui est loin d'être sûr, voire peu probable au XIII^e siècle, c'est que la chancellerie royale soit un fait de géographie stable et de diatopie fixe. Ni le côté producteur (les scribes de provenance diverse, les affaires éparses traitées dans les documents) ni le côté récepteur (souvent des souverains provenant des quatre coins du royaume) ne peuvent être considérés comme enracinés dans une localité déterminée. Il en découle que les documents produits par des chancelleries royales ou impériales ne peuvent en aucun cas servir de témoins diatopiques.²⁸

En ce qui concerne la standardisation du français, ces résultats mettent à nouveau en question l'hypothèse de l'origine francienne du français standard. Bernard Cerquiglini, dans sa « Naissance du français » fait la critique du francien:

« C'est dès lors très précisément la question de l'origine du français littéraire, commun, national, du « bon » français, qui se trouve posée. Et résolue, car depuis longtemps une réponse scientifique a été apportée, réponse convenue, couramment admise, et que l'on peut lire dans toutes les Histoires de la langue: le français national provient très directement du dialecte de l'Île-de-France, et résulte de sa suprématie » (Cerquiglini 21993, 114).

Il soulève ensuite 4 objections contre le francien qui portent sur la perspective téléologique des historiens de langue du XIX^e siècle, sur le statut diasystématique

²⁵ Voir Völker 2003, 153.

²⁶ Comme l'a prétendu Titz 1926, 80s., – et il n'est pas le seul – pour le système bicasuel.

²⁷ Comme notre corpus n'implique pas de documents de l'ouest, nos résultats ne contredisent pas ceux de Dees 1978 qui situent l'origine de l'abandon du système bicasuel à l'ouest; voir aussi Dees 1980.

²⁸ Voir aussi Völker 2000, 164, et id. 2003, 136, avec d'autres exemples.

de la variété dite « parisienne » ainsi que sur le développement de la monarchie française, et il conclut :

« Le français national, notre français, ne provient donc pas d'un terroir, mais de la littérature. De cette *scripta* essentiellement poétique, quasi nationale dans les *Serments* (qui doivent annoncer un État), interrégionale d'oïl dans les textes littéraires qui suivent, et qu'élaborent les clercs, d'expérience en expérience, jusqu'à ce qu'elle se fige en ancien français commun. Le français résulte de ce travail séculaire d'écriture, de cette édification cléricale » (Cerquiglini 21993, 118s.).

Bernard Cerquiglini n'est pas le seul à faire la critique du francien. Comme lui, Anthony Lodge met en question le rôle prépondérant de la diatopie.²⁹ Le problème méthodologique nous semble que ni Cerquiglini ni Lodge – comme beaucoup d'autres qui se sont exprimés sur l'origine du standard français – ne se basent sur une méthodologie empirique.

Nos résultats, eux produits par une approche empirique, vont cependant, au moins en ce qui concerne la critique du francien, dans le sens de Cerquiglini; et ils sont en outre en accord avec les résultats de la dialectologie moderne:

« Quoiqu'il en soit ces études montrent que le dialecte central différait nettement du français officiel du XVI^e s. Malgré cela on s'ingénie encore à le considérer comme sorti d'un dialecte propre à l'Île-de-France. (...). Ceci s'explique toujours par le mythe de la pureté du français et par l'évidence que les français centraux sont proches parents les uns des autres » (Fondet 1980, 676).³⁰

En ce qui concerne le caractère extralinguistique de la ou des variétés directrices, on aurait pourtant tendance à contredire Cerquiglini 21993, puisque selon nos résultats on ne devrait pas écarter la chancellerie royale comme facteur important pour le développement d'un proto-standard (pour éviter le terme anachronique d'ancien français commun). Il reste, certes, à préciser le rôle exact de la chancellerie royale dans l'ensemble des autres facteurs qui entrent en ligne de compte : le facteur diatopique et le francien (favorisé des manuels par tradition ; voir récemment Burdy 2004, 340), la koinésation littéraire (p. ex. Cerquiglini 21993), la koinésation orale à Paris (p. ex. Lodge 1998 et 2004), les scriptoria monastiques (p. ex. Stanovaïa 2003)... Il nous semble d'ailleurs probable que sur l'axe diachronique ces facteurs se soient entremêlés et succédés.

C'est dans ce but que nous plaçons pour une augmentation des efforts visant à saisir numériquement les chartes écrites en ancien français : pour compléter le panorama diatopique, nous avons besoin de chartes provenant de toutes les régions francophones;³¹ pour compléter le panorama diachronique nous devons aborder le XIV^e et le XV^e siècle; pour connaître mieux le rôle des scriptoria et des chancelleries, nous avons besoin de plus de séries scribales... Si nous ajoutons à tout cela d'autres catégories « dia- » (p. ex. la diatextualité pour tenir compte des spécificités du langage littéraire), la linguistique de corpus et la linguistique variationnelle historique nous permettront peut-être un jour de dresser une image

²⁹ Voir Lodge 1998 et 2004.

³⁰ Voir aussi Wüest 1985, qui va dans le même sens.

³¹ Les résultats de Dees 1980, cartes 186 et 206-208, selon lesquels le déclin du système bicasuel commence à l'ouest, pourraient par exemple être réanalysés sous une perspective variationniste.

du développement de l'ensemble des variétés de la langue française et des processus de standardisation. Le but de cette démarche n'est certainement pas aisé à atteindre, mais il est réalisable: dépeindre la diachronie des architectures du français et ainsi connaître mieux les origines de la langue standard.³²

Bibliographie

- Boutier, Marie-Guy, *Études sur des chartes luxembourgeoises*, in: Gärtner/Holtus/Rapp/Völker 2001, 419-447.
- Burdy, Philipp, Rez. zu Holtus/Overbeck/Völker 2003b, in: *Vox Romanica* 63 (2004), 338-342.
- Cerquiglini, Bernard, *Éloge de la variante. Histoire critique de la philologie*, Paris, Seuil, 1989.
- Cerquiglini, Bernard, *La naissance du français*, Paris, PUF, 1993.
- Coseriu, Eugenio, *Einführung in die strukturelle Betrachtung des Wortschatzes*, Tübingen, [Narr], 1970.
- Dees, Anthonij, *A toz cels qui cez lettres verront. La déchéance de la déclinaison et l'ordre des mots*, in: Stuij, René E. V. (ed.), *Langue et littérature françaises du moyen âge*, Assen/Amsterdam, Van Gorcum, 1978, 2-11.
- Dees, Anthonij, *Atlas des formes et des constructions des chartes françaises du 13^e siècle*, avec le concours de Pieter Th. van Reenen et de Johan de Vries, Tübingen, Niemeyer, 1980.
- Dieckmann, Sandra, *Langue de fransois: die andere Sprache? Zur Selbst- und Fremdwahrnehmung im Floriment von Aimon de Varennes*, in: Dahmen, Wolfgang, et al. (edd.), *Schreiben in einer anderen Sprache. Zur Internationalität romanischer Sprachen und Literaturen. Romanistisches Kolloquium XIII*, Tübingen, Narr, 2000, 21-34.
- Flydal, Leiv, *Remarques sur certains rapports entre le style et l'état de langue*, in: *Norsk Tidsskrift for Sprogvidenskap* 16 (1952), 241-258.
- Fondet, Claire, *Dialectologie de l'Essonne et de ses environs immédiats*, 2 vol., Lille/Paris, Université de Lille/Champion, 1980.
- Gärtner, Kurt/Holtus, Günter (edd.), *Beiträge zum Sprachkontakt und zu den Urkundensprachen zwischen Maas und Rhein*, Trier, THF, 1995.
- Gärtner, Kurt/Holtus, Günter (edd.), *Urkundensprachen im germanisch-romanischen Grenzgebiet. Beiträge zum Kolloquium am 5./6. Oktober 1995 in Trier, Mainz, Zabern*, 1997.
- Gärtner, Kurt/Holtus, Günter (edd.), *Überlieferungs- und Aneignungsprozesse auf dem Gebiet der westmitteldeutschen und ostfranzösischen Urkunden- und Literatursprachen*, Trier, Kliomedien, sous presse.
- Gärtner, Kurt/Holtus, Günter/Rapp, Andrea/Völker, Harald (edd.), *Skripta, Schreiblandschaften und Standardisierungstendenzen. Urkundensprachen im Grenzbereich von Germania und Romania im 13. und 14. Jahrhundert. Beiträge zum Kolloquium vom 16. bis 18. September 1998 in Trier*, Trier, Kliomedien, 2001.
- Guyotjeannin, Olivier/Pycke, Jacques/Tock, Benoît-Michel, *Diplomatique médiévale*, [Turnhout], Brepols, 1993.
- Holtus, Günter, *Zu einigen Beschreibungsversuchen der Varietäten und Strukturen der französischen Gegenwartssprache*, in: *Französisch heute* 9 (1978), 161-169.
- Holtus, Günter, *En torno a una lingüística variacional de la lengua asturiana*, in: *Lletres asturianas* 43 (1992), 21-36.
- Holtus, Günter/Körner, Anja/Völker, Harald, *«Endogene» und «exogene» Analogien. Hyperkorrektismen und andere Analogienbildungen in den altfranzösischen Urkunden der Grafen von Luxemburg (1237-1281)*, in: Henrard, Nadine/Moreno, Paola/Thiry-Stassin, Martine (edd.),

³² L'auteur tient à remercier Sandra Félix et Günter Holtus (Göttingen) ainsi que Cécile Veneman (Berlin) pour la relecture critique du texte et pour leurs indications précieuses.

- Convergences médiévales. Épopée, lyrique, roman. Mélanges offerts à Madeleine Tyssens*, Bruxelles, De Boeck, 2001, 257–270.
- Holtus, Günter/Overbeck, Anja/Völker, Harald, *Ce qu'un texte médiéval peut nous faire savoir. Édition et analyse d'un corpus de chartes luxembourgeoises*, in: Sánchez Miret, Fernando (ed.), *Actas del XXIII Congreso Internacional de Lingüística y Filología Románica. Salamanca, 24-30 septiembre 2001*, vol. IV: *Sección 5: Edición y crítica textual – Sección 6: Retórica, poética y teoría literaria*, Tübingen, Niemeyer, 2003, 111-122 (=2003a).
- Holtus, Günter/Overbeck, Anja/Völker, Harald, *Luxemburgische Skriptastudien. Edition und Untersuchung der altfranzösischen Urkunden Gräfin Ermesindes (1226–1247) und Graf Heinrichs V. (1247–1281) von Luxemburg*, Tübingen, Niemeyer, 2003 (=2003b).
- Holtus, Günter/Radtke, Edgar (edd.), *Varietätenlinguistik des Italienischen*, Tübingen, Narr, 1983.
- Holtus, Günter/Radtke, Edgar (edd.), *Sprachlicher Substandard*, 3 vol., Tübingen, Narr, 1986-1990.
- Holtus, Günter/Völker, Harald, *Das Ineinandergreifen von Diachronie und Synchronie bei der Edition altfranzösischer Urkundentexte*, in: Eggers, Eckhard/Becker, Joachim/Udolph, Jürgen/Weber, Dieter (edd.), *Florilegium Linguisticum. Festschrift für Wolfgang P. Schmid zum 70. Geburtstag*, Frankfurt a.M. et al., Lang, 1999, 199-223.
- Holtus, Günter/Völker, Harald, *La localizzazione di documenti medievali scritti in volgare*, in: Świątkowska, Marcela/Sosnowski, Roman/Piechnik, Iwona (edd.), *Maestro e Amico. Miscellanea in onore di Stanisław Widlak / Mistrz i Przyjaciel. Studia dedykowane Stanisławowi Widlakowi*, Kraków, Wydawnictwo Uniwersytetu Jagiellońskiego, 2004, 165-170.
- Imbs, Paul, *La place du vocabulaire ancien dans un thesaurus de la langue française*, in: *Lexicologie et lexicographie françaises et romanes. Orientations et exigences actuelles (Strasbourg, 12-16 novembre 1957)*, Paris, CNRS, 1961, 133-147.
- Lodge, R. Anthony, *French: From Dialect to Standard*, London/New York, Routledge, 1993.
- Lodge, Anthony, *Vers une histoire du dialecte urbain de Paris*, in: *Revue de Linguistique Romane* 62 (1998), 95-128.
- Lodge, R. Anthony, *A Sociolinguistic History of Parisian French*, Cambridge, Cambridge University Press, 2004.
- Monfrin, Jacques, *Introduction*, in: Gigot, Jean-Gabriel, *Documents linguistiques de la France (série française). Chartes en langue française antérieures à 1271 conservées dans le département de Haute-Marne*, Paris, CNRS, 1974, XI-LXXX.
- Müller, Bodo, *Das Französische der Gegenwart. Varietäten, Strukturen, Tendenzen*, Heidelberg, Winter, 1975.
- Overbeck, Anja, *Zur Edition altfranzösischer Texte. Editionspraktische Überlegungen im Anschluß an die 'New Philology'*, in: *Editio* 17 (2003), 67-88 (=2003a).
- Overbeck, Anja, *Literarische Skripta in Ostfrankreich. Edition und sprachliche Analyse einer französischen Handschrift des Reiseberichts von Marco Polo (Stockholm, Kungliga Biblioteket, Cod. Holm. M 304)*, Trier, Kliomedia, 2003 (=2003b).
- Overbeck, Anja/Völker, Harald, *TUSTEP et chartes médiévales. Remarques informatiques*, in: *Actes du XXIV^e Congrès International de Linguistique et de Philologie Romanes (Aberystwyth 2004)*, Tübingen, Niemeyer, sous presse.
- Pfister, Max, *Scripta et koinè en ancien français aux XII^e et XIII^e siècles?*, in: Knecht, Pierre/ Marzys, Zygmunt/ Destraz, Dominique (edd.), *Écriture, langues communes et normes. Formation spontanée de koinès et standardisation dans la Galloromania et son voisinage. Actes du colloque tenu à l'université de Neuchâtel du 21 au 23 septembre 1988*, Neuchâtel/Genève, Faculté des Lettres/Droz, 1993, 17-41.
- Prüßmann-Zemper, Helga, *Französisch: Varietätenlinguistik des Französischen*, in: Holtus, Günter/Metzeltin, Michael/Schmitt, Christian (edd.), *Lexikon der Romanistischen Linguistik, vol. V,1: Die einzelnen romanischen Sprachen und Sprachgebiete von der Renaissance bis zur Gegenwart: Französisch*, Tübingen, Niemeyer, 1990, 830-843.

- Remacle, Louis, *Le problème de l'ancien wallon*, Liège, Faculté de Philosophie et Lettres, 1948.
- Söll, Ludwig, *Gesprochenes und geschriebenes Französisch*, Berlin, Schmidt, 1974.
- Stanovaia, Lydia A., *La standardisation en ancien français*, in: Goyens, Michèle/Verbeke, Werner (edd.), *The Dawn of the Written Vernacular in Western Europe*, Universitaire Pers Leuven, Leuven, 2003, 241-272.
- Thelen, Udo, *Sprachliche Variation und ihre Beschreibung. Zur Markierungspraxis in der französischen Sprachlehre und Grammatikographie zwischen Maas und Rhein vom 16. bis zum 18. Jahrhundert*, Tübingen, Niemeyer, 1999.
- Titz, Karel, *La substitution des cas dans les pronoms français*, Brno/Paris, Piša/Champion, 1926.
- Völker, Harald, *Chartes luxembourgeoises du 13^e siècle. Scripta régionale, locale ou « individuelle »?*, in: Englebert, Annick/Pierrard, Michel/Rosier, Laurence/Van Raemdonck, Dan (edd.), *Actes du XXII^e Congrès International de Linguistique et de Philologie Romanes. Bruxelles, 23-29 juillet 1998*, vol. 5: « *Les manuscrits ne brûlent pas* ». *Travaux de la section « Philologie, codicologie, éditions de textes »*, Tübingen, Niemeyer, 2000, 159-166.
- Völker, Harald, *Skripta und Variation. Untersuchungen zur Negation und zur Substantivflexion in altfranzösischen Urkunden der Grafschaft Luxemburg (1237–1281)*, Tübingen, Niemeyer, 2003.
- Völker, Harald, *Zwischen Textedition und historischer Varietätenlinguistik. Tustep-unterstützte hypertextuelle Strukturen zur Analyse rekontextualisierter Texte*, in: Reeg, Gottfried/Schubert, Martin J. (edd.), *Edieren in der elektronischen Ära*, Berlin, Weidler, 2004, 127-144.
- Völker, Harald, *Hypertextstrukturen in historischen Textkorpora*, in: Pusch, Claus D./Kabatek, Johannes/Raible, Wolfgang (edd.), *Romanistische Korpuslinguistik II. Korpora und diachrone Sprachwissenschaft / Romance Corpus Linguistics II. Corpora and Diachronic Linguistics*, Tübingen, Narr, 2005, 45-55.
- Wampach, Camille, *Urkunden- und Quellenbuch zur Geschichte der allluxemburgischen Territorien bis zur burgundischen Zeit*, 10 vol., Luxemburg, Paulus, 1935-1955.
- Weinreich, Uriel, *Is a Structural Dialectology Possible?*, in: *Word* 10 (1954), 388-400.
- Winkelmann, Otto, *Französisch: Sprachnormierung und Standardsprache*, in: Holtus, Günter/Metzeltin, Michael/Schmitt, Christian (edd.), *Lexikon der Romanistischen Linguistik*, vol. V,1: *Die einzelnen romanischen Sprachen und Sprachgebiete von der Renaissance bis zur Gegenwart: Französisch*, Tübingen, Niemeyer, 1990, 334-353.
- Wüest, Jakob, *Le «patois de Paris» et l'histoire du français*, in: *Vox Romanica* 44 (1985), 234-258.