

Linguistique comparée des langues modernes

L4LFZ042, jeudi 16h30-18h30 (Centre universitaire Malesherbes, 108 blvd Malesherbes)

Professeur : André THIBAUT

Semaine 4 : Les langues dans le temps, l'espace et la société (II) : dialectologie galloromane, quelques concepts ; l'élaboration d'une « langue historique ». Concept de *scripta* médiévale ; les « dialectes » de l'ancien français.

0. Introduction

Après nous être interrogés la semaine dernière sur les concepts de langue, de dialecte et de patois, nous allons observer de plus près cette semaine les deux faces d'une même médaille : d'une part, la réalité dialectale de la Galloromania ; d'autre part, la lente apparition et élaboration du français comme langue standardisée.

1. Quelques concepts

Avant de présenter les grandes aires dialectales de la Galloromania, il faut d'abord expliquer brièvement ce que sont la **dialectologie** et la **géographie linguistique**. En effet, les zones dialectales dont nous allons parler n'ont pu être délimitées qu'à l'aide de ces sciences que l'on peut définir respectivement comme l'étude des dialectes, et de la différenciation linguistique dans l'espace, c'est-à-dire des différences que l'on observe entre les parlers d'une région donnée.

1.1. La variation linguistique

Ces sciences trouvent leur raison d'être dans le fait que la langue ne peut être décrite d'une façon satisfaisante que si l'on tient compte de sa variabilité, de sa tendance à la variation ; la langue est, en effet, un objet essentiellement hétérogène – et plus encore dans les cas où elle vit « en liberté » (par opposition à la situation des langues standardisées, plus ou moins encadrées dans le respect d'une norme). Cette variation s'observe dans le temps et dans l'espace, les deux phénomènes étant d'ailleurs reliés. Plus deux états de langue sont éloignés dans le temps, plus ils risquent d'être différents ; de même, plus deux variétés d'une même langue sont éloignées dans l'espace, plus elles différeront l'une de l'autre – mais seulement si cet éloignement est ancien, le facteur temps s'avérant beaucoup plus déterminant que le facteur espace.

1.2. Méthodologie

Comment les linguistes s'y prennent-ils pour délimiter les zones dialectales, et caractériser les dialectes ? La seule façon d'y arriver consiste à étudier les parlers d'un très grand nombre de localités, d'y relever un très grand nombre de phénomènes linguistiques (phonétiques, lexicaux, etc.), et de tracer sur une carte les frontières qui séparent les phénomènes relevés. Lorsqu'on a tracé plusieurs de ces frontières, on constate normalement que certaines d'entre elles se rejoignent et forment des séparations plus importantes entre les aires. C'est la prise en considération de ces frontières, que l'on réunit dans des atlas linguistiques, qui permet de tracer une carte des principales aires dialectales. Si l'on regarde par exemple la carte « Limites linguistiques », sur laquelle ont été reportées de nombreuses *isoglosses* (c'est le terme technique employé, v. glossaire ci-dessous), on constate que certaines zones ressortent. C'est sur la base

de cartes d'atlas de ce genre que l'on peut proposer des cartes idéalisées, comme celle tirée du *Lexikon der romanistischen Linguistik*, vol. V, p. 580, où les aires dialectales ont été clairement identifiées (vous pouvez télécharger cette carte sur le site, si ce n'est déjà fait).

Les linguistes du galloroman n'ont pas toujours été d'accord sur l'existence d'aires dialectales objectivement identifiables dans la Galloromania. Plusieurs prétendaient, il y a environ un siècle, que les patois se modifiaient graduellement d'un endroit à l'autre, peu à peu, et que les différents regroupements que l'on étiquetait « gascon », « limousin », « champenois », etc., n'étaient que des vues de l'esprit. Or, la juxtaposition d'un très grand nombre d'isoglosses sur la carte de la Galloromania montre à l'évidence qu'en certains endroits les différences entre parlars sont très fortes et permettent donc de conclure à l'existence de frontières réelles entre des parlars qui, pour être apparentés, n'en sont pas moins identifiables comme des unités dont l'identité est démontrable à partir de critères objectifs. Cela dit, la « distance » entre parlars peut varier énormément selon les cas.

1.3. Terminologie

- *dialectologie* : en synchronie, « étude scientifique des dialectes ; discipline qui s'est donnée pour tâche de décrire comparativement les différents dialectes dans lesquels une langue s'est diversifiée dans l'espace » ; s'emploie aussi pour la description de parlars pris isolément, sans référence aux parlars voisins ; en diachronie, « étude des phénomènes de différenciation dialectale par lesquels une langue relativement homogène à une époque donnée subit au cours de son histoire une fragmentation qui aboutit à des dialectes ou même à des langues distinctes ».
- *géographie linguistique* (aussi appelée *géolinguistique*) : « présentation systématique des données dialectales sous la forme d'atlas linguistiques, dans le but de présenter un domaine linguistique dans toute sa variété, et de localiser les uns par rapport aux autres les différents dialectes d'une langue ».
- *isoglosse* (phonétique, lexicale, etc.) : « frontière qui sépare deux aires présentant pour un trait linguistique donné des réalisations différentes ». Par exemple, l'une des lignes qui traversent la carte « Limites linguistiques » de l'est à l'ouest représente probablement la frontière entre les régions, au nord, où la terminaison latine des verbes du premier groupe -ARE est devenue -er, des régions du sud où elle est restée -ar.
- *faisceau d'isoglosses* : « ensemble d'isoglosses superposées ou proches ». Le plus important de la Galloromania est celui qui sépare le poitevin du gascon (v. encore une fois la carte « Limites linguistiques »).
- *aires dialectales* : « aires délimitées par des faisceaux d'isoglosses ». C'est le cas, par exemple, de l'aire du wallon (au nord-est de la Galloromania).
- *foyer d'expansion* : « centre à partir duquel les innovations linguistiques se diffusent ». Dans la Galloromania, le principal foyer d'expansion est Paris, mais les grandes métropoles régionales ont aussi pu jouer ce rôle, en particulier dans la diffusion du français dans leurs aires d'influence respectives ; dans la Romania de l'Antiquité, c'était Rome.
- *aires latérales* : « aires situées de part et d'autre d'un foyer d'expansion ». À l'échelle de la francophonie (et non de la Galloromania), on peut considérer la Suisse romande, la Belgique et le Québec comme des aires latérales par rapport à la France. On y remarque, par exemple, l'existence d'archaïsmes comme la désignation des repas (*déjeuner, dîner, souper* au lieu de *petit-déjeuner, déjeuner, dîner*).
- *aires résiduelles* : « poches de survivance de phénomènes linguistiques archaïques ». Il est facile d'imaginer qu'une localité isolée, en milieu rural, ne reçoive qu'en retard des inno-

vations exportées par les foyers d'expansion et conserve plus longtemps qu'ailleurs des réalisations archaïques. Par exemple, la survivance du [h] aspiré dans la prononciation de certains mots (*en haut, la hotte*) a été relevée dans les Vosges ; le maintien de la prononciation apicale (« roulée ») du *r* s'observe encore chez les gens âgés en Bourgogne.

- *le Croissant* : « zone intermédiaire de transition entre le nord-occitan et la langue d'oïl, caractérisée par un étalement progressif des isoglosses (ces dernières se rejoignent à l'ouest pour former une pointe, et se divisent en deux faisceaux à l'est pour ceinturer le francoprovençal) » (v. la carte tirée du *Lexikon*).

2. Les zones dialectales du galloroman

Nous allons maintenant détailler les zones dialectales du galloroman, en nous reportant à la carte du *Lexikon*.

2.1. Les dialectes d'oïl

Ce sont les plus éloignés du latin (leur phonétique est souvent dite « révolutionnaire »). Cela s'explique par leur plus grand éloignement géographique du centre de la Romania, ainsi que par l'influence du superstrat franc, beaucoup plus forte dans le nord de la France. Nous avons déjà dit que l'appellation « d'oïl » fait référence au mot *oïl* qui est la forme du mot *oui* en ancien français (et qui vient du latin HOC ILLE, qui veut dire à peu près « c'est ça »). Cette appellation s'oppose à *langue d'oc* (v. ci-dessous). La langue d'oïl a connu de nombreux « dialectes littéraires » au moyen âge, qu'il est convenu d'appeler aujourd'hui des « scriptas » ou « scriptae » (v. ci-dessous).

Voici les plus importants dialectes d'oïl :

2.1.1. Le picard, l'ancien dialecte littéraire de Picardie, qui survit aujourd'hui à l'état de patois (dont l'une des variétés est appelée *chti* ou *chtimi*). Quelques caractéristiques linguistiques de ce dialecte :

- article féminin *le* ou *li* pour *la*
- maintien de la distinction des nasales *an* et *en*, confondues en français dès le 11^e s.
- absence de palatalisation des groupes lat. KA et GA, ex. *canter, gambe* pour *chanter, jambe*, que l'on attribue à la force du superstrat germanique
- lat. CE et CI aboutissent à *ch*, ex. *cherf, chité* pour *cerf, cité*, de même que les groupes [kj] et [tj], ex. *canchon* pour *chanson*
- maintien du *w*-germanique tel quel alors qu'il est devenu [g] en français, ex. *warder, waster, waignier* pour *garder, gâter, gagner*, etc.)

2.1.2. Le wallon, le dialecte de Wallonie (en Belgique). Quelques caractéristiques :

- article féminin *le* ou *li* pour *la*
- maintien de la distinction des nasales *an* et *en*
- maintien du *w*-germanique tel quel (trois traits que le wallon partage avec le picard, que nous venons de voir)

- le [e] accentué issu de lat. A tonique libre (comme par ex. dans la terminaison de l'infinitif des verbes du premier groupe) se diphtongue en [ej], ex. infinitifs en *-eir*, participes passés en *-ei*, etc.).

2.1.3. *Le normand* (et anciennement l'anglo-normand). Quelques phénomènes repérables dans les textes :

- *k-* pour *qu-*
- *u* pour *o*
- diphtongue nasale *aun*, ex. *aunte* « tante », cf. anglais moderne *aunt*
- Le *w-* initial d'origine germanique, resté [w] en picard et en wallon mais devenu [g] en français (v. ci-dessus) est devenu [v] en normand (ex. **werwolf* > fr. *garou* mais normand *varou*).
- Le *h-* initial dans les mots d'origine germanique s'est maintenu (cf. par ex. *hache*, du francique HAPPIA).

2.1.4. *L'anglo-normand* est une variété de normand exportée en Angleterre au 11^e siècle, qui allait influencer énormément la langue anglaise.

L'existence de l'anglo-normand est due à la conquête de l'Angleterre par Guillaume le Conquérant, en 1066. Il faut d'abord expliquer qu'après les Wisigoths, les Burgondes et les Francs, une dernière vague d'envahisseurs germaniques est venue s'installer en France, aux 9^e et au 10^e siècles : il s'agit des **Normands**. Leur nom, étymologiquement, veut dire *hommes du Nord*, et en effet les envahisseurs normands venaient de Scandinavie ; ce ne sont ni plus ni moins que ceux que l'on connaît plus généralement sous le nom de Vikings. Au 10^e siècle, donc, les Normands s'installent dans la région de l'embouchure de la Seine, qui justement fut baptisée d'après eux (la *Normandie*). Ce sont des chefs de guerre, comme les Francs ; ils dominent la société galloromaine de Normandie, mais ils sont minoritaires et très vite ils s'assimilent linguistiquement aux populations sur place ; en quelques générations, ils perdent leur langue germanique et se mêlent aux populations locales. L'ancien dialecte d'oïl qu'ils parlaient, et qui est très proche des autres variétés d'ancien français, est connu sous le nom de **normand**. Seulement voilà, non contents d'avoir conquis la Normandie, les Normands continuent de guerroyer, et envahissent l'Angleterre en **1066**, sous le règne de **Guillaume le Conquérant**. Ils y amènent avec eux leur langue, le normand, qui est donc pratiquement une variété d'ancien français. À partir du moment où ce dialecte littéraire a été transplanté en Angleterre, il a commencé à vivre de sa vie propre : on l'appelle alors l'**anglo-normand**.

Comme les Normands constituaient l'élite là-bas, pendant plusieurs siècles, au moyen âge, en Angleterre, le français a eu un statut très enviable de langue de culture, et de nombreux textes, littéraires ou juridiques, y ont été écrits dans cette langue. Comme la plus grande partie de la population parlait anglais et non anglo-normand, c'est quand même l'anglais qui a fini par s'imposer, et au cours des siècles on a fini par abandonner la pratique de l'anglo-normand en Angleterre, même parmi les élites ; cela dit, l'anglo-normand survit à travers l'anglais, qui de toutes les langues germaniques est de loin celle dont le vocabulaire a été le plus influencé par le français (ainsi que par le latin des savants, grande langue de culture au moyen âge et qui a transmis énormément de mots à l'anglais dans des traductions et des adaptations de textes écrits d'abord en latin médiéval).

De manière générale, on peut dire que bien souvent, lorsque l'on compare l'anglais et l'allemand, on constate qu'en cas de divergence le mot anglais vient du français, à travers l'anglo-normand, alors que le mot allemand est d'origine germanique. Prenons des mots aussi simples que *table* ou *chaise*. En allemand, cela ne ressemble pas du tout au français : *der Tisch, der Stuhl*. En anglais, par contre, vous avez *the table*, qui a connu bien sûr une évolution phonétique, mais qui dans sa graphie est très archaïque car il a gardé exactement la même que celle du français. Quant à *the chair*, on remarque quand même une différence : le mot se termine par un *-r* et non par *-se* comme dans le mot français *chaise*. C'est que le mot anglo-normand était légèrement différent du mot du français central. Cela dit, la forme avec *-r* s'est aussi perpétuée en français, mais dans un emploi différent : c'est cette forme qu'on retrouve dans le mot *chaire*, qui désigne la tribune d'un professeur ou d'un ecclésiastique, et par extension métonymique le poste de professeur dans une université. Cette variante phonétique désignait elle aussi un siège à l'origine, mais a connu en français une évolution sémantique particulière. En anglais, la seule forme transmise par l'anglo-normand, *chair*, a les deux sens (*a chair to sit on* ; *a chair of linguistics at the University*).

2.1.5. Le français : ce qui est devenu la langue standard de toute la France et qui a pour base historique une sorte de koinè écrite reposant semble-t-il sur les usages les moins excentriques de la France d'oïl au moyen âge se distingue au sein des autres langues romanes (je ne parle plus ici des seuls parlers d'oïl) par :

- la palatalisation des occlusives vélaires initiales devant A (donc, *chanter* et non *cantar*)
- le passage de [u] à [y] (cf. lat. LUNA > fr. *lune*)
- le passage de A tonique libre à [e], ce qui fait que les infinitifs en *-ar(e)* du latin, de l'italien, de l'espagnol ou du portugais sont en *-er* en français
- la diphtongaison du E long tonique latin en *ei*, puis *oi* [wa], ce qui nous donne donc *toile* au lieu de *tela*, *poire* au lieu de *pera*, etc.
- la nasalisation totale des voyelles entravées par d'anciennes consonnes nasales (phénomène assez rare dans les autres langues romanes, à l'exception notable du portugais)
- la chute des occlusives intervocaliques (*muer* pour esp. port. *mudar*, ital. *mutare* ; *jouer* pour esp. *jugar*, port. *jogar*, ital. *giocare*)
- le caractère oxyton de son accentuation (c'est-à-dire que l'accent tombe toujours sur la dernière syllabe – orale – du mot ; donc, esp. port. *fantástico* mais fr. *fantastique*)

2.1.6. D'autres dialectes d'oïl qui ne survivent aujourd'hui qu'à l'état de patois :

- les parlers de l'Ouest (gallo, angevin)
- les parlers du Sud-Ouest (poitevin, saintongeais)
- les parlers du Centre (berrichon, bourbonnais)
- les parlers de l'Est (champenois, lorrain, bourguignon, franc-comtois)

2.2. Le francoprovençal

Localisation : centre-est de la France, Suisse romande, Val d'Aoste.

Caractéristiques :

- A tonique libre lat. reste *a* (comme en occitan) : CANTARE > (t)sāntá ;
- A final atone latin reste *a* (comme en occitan) : TELA > teila ;
- -U final latin atone (passé à -o en latin vulgaire) se maintient après un groupe consonantique comme voyelle de soutien, au lieu de passer à -e comme en français : *miraclo, proprio, autro, sieglo* (au lieu de fr. *miracle, propre, autre, siècle*, etc.).

Ces deux derniers facteurs contribuent à faire du francoprovençal une langue paroxytonique (c'est-à-dire, dont la plupart des mots sont accentués sur l'avant-dernière syllabe), à la différence du français qui est exclusivement oxytonique (à l'exclusion des accents méridionaux qui maintiennent la prononciation du -e muet final).

Le francoprovençal se caractérise par son extrême morcellement dialectal. Il n'a guère connu d'illustration littéraire au moyen âge, contrairement aux dialectes d'oïl et d'oc.

Le nom de *francoprovençal* a été donné à cette subdivision dialectale de la Galloromania en raison de ses caractéristiques linguistiques qui font qu'elle se rattache tantôt à la langue d'oïl, tantôt à la langue d'oc.

Aujourd'hui, il n'y a plus que dans les régions montagneuses du Valais (Suisse), du Val d'Aoste (Italie) et de la Haute-Savoie (France) que l'on peut encore trouver des locuteurs de patois francoprovençaux.

2.3. La langue d'oc

Localisation : le sud de la France, une partie du Piémont (Italie) ; petit débordement en Espagne (le Val d'Aran, dans les Pyrénées). La langue d'oc est la plus conservatrice des langues de la Galloromania (c'est-à-dire celle qui s'est le moins éloignée du latin) ; on dit que c'est la langue-soeur du catalan, parce qu'elle en est très proche. Ce caractère conservateur s'explique par la moindre distance du centre de la Romania, le haut degré de romanisation à l'époque de l'Empire, et la moindre influence des superstrats germaniques. On rappellera que le nom de *langue d'oc* vient de la façon de dire *oui* dans cette langue, c'est-à-dire *oc*, du latin HOC (qui voulait dire « cela, c'est cela »). On la connaît aussi sous le nom d'*occitan*. Au moyen âge, l'occitan (que l'on connaît populairement sous le nom de *provençal*, même s'il n'était pas limité à la Provence) était une grande langue de culture ; c'était, plus précisément, la langue des troubadours, et à ce titre elle était alors encore plus prestigieuse que les dialectes littéraires de la France d'oïl. Aujourd'hui, l'occitan ne survit qu'à travers de nombreuses variétés plus ou moins rurales et très dialectalisées (c'est-à-dire différenciées régionalement), qui ne se maintiennent qu'avec le statut social d'un patois.

Variétés dialectales :

2.3.1. *Le nord-occitan*. Comprend le limousin, l'auvergnat et le provençal alpin. Caractéristiques : palatalisation de CA et GA latins (resp. > [j] et [ʒ]) comme en langue d'oïl, mais contrairement au reste du domaine occitan.

2.3.2. *L'occitan moyen*. Regroupe les parlers les plus conservateurs de la Galloromania. Comprend le languedocien et le provençal.

2.3.3. *Le gascon*. Se distingue par la chute de -N- et -L- intervocaliques latins, phénomène dû à un substrat pyrénéen et que le gascon partage avec le portugais, et par le passage de F- latin à *h-*, trait que le gascon partage avec le castillan. FARINA devient donc *haría*. À vrai dire, les spécialistes considèrent aujourd'hui que le gascon représente un ensemble à part entière, qui s'est individualisé déjà au 6^e siècle après J.-C. Bien que vous risquiez de trouver encore aujourd'hui le gascon comme un sous-ensemble de l'occitan dans la plupart des manuels, il conviendrait de le considérer comme un sous-ensemble du galloroman à part entière.

3. Distinction patois / français régional

J'attire encore une fois votre attention sur le fait que les patois ne doivent pas être confondus avec le français régional. La raison pour laquelle on a parfois tendance à ne pas bien distinguer les deux réside dans le fait que les patois sont de plus en plus moribonds, qu'ils disparaissent peu à peu, et qu'ils sont bien sûr remplacés par le français, dans ses différentes variétés régionales. Il y a donc chez les locuteurs bilingues beaucoup d'alternance entre les deux (patois et français), et beaucoup d'interférences, mais il n'en reste pas moins que ce sont deux réalités bien différentes, du point de vue de leur parcours historique.

4. L'élaboration du français, langue « standardisée », « langue-toit »

Nous avons terminé le cours la semaine dernière avec les concepts de *langue écart*, *langue par élaboration* et *langue-toit*. Nous avons caractérisé le français comme une langue par écart, car il ne peut être considéré comme le dialecte d'aucune autre langue, étant bien sûr suffisamment différent de ses plus proches parents standardisés (espagnol, italien, etc.) ; nous l'avons également considéré comme une langue par élaboration, puisqu'il a subi au fil des siècles nombre d'interventions externes sur sa forme même, tendant à lui conférer toutes les caractéristiques d'une langue standardisée, contrairement aux patois :

- 1) tendance à l'homogénéisation dialectale, à la « dédialectalisation » (par opposition à la nature normalement très dialectalisée des langues « en liberté ») ;
- 2) apparition de nombreux technoclectes rendant la langue apte à l'expression de la pensée dans les champs les plus divers de l'activité humaine, des domaines les plus concrets aux plus abstraits ;
- 3) enrichissement de l'inventaire des procédés syntaxiques (prépositions, conjonctions, etc.) et morphologiques (dérivation, composition) permettant de structurer, d'enrichir et de préciser l'expression de la pensée.

Enfin, nous lui avons reconnu le statut de « langue-toit », puisque ce processus d'élaboration a débouché sur une cohabitation du français avec les autres dialectes galloromans ainsi que les « langues régionales », avec répartition fonctionnelle asymétrique : fonctions prestigieuses monopolisées par le français, fonctions peu prestigieuses assumées par les autres (mais aussi bien sûr par le français, lorsqu'il devient la seule langue en présence, les patois s'étant peu à peu estompés).

Mais ce français « par élaboration » dont l'existence nous paraît si évidente et qui nous semble avoir toujours existé est né comment ? Que sait-on de ce que fut le français dans les pre-

miers siècles de son histoire ? Quelle en est la base historique, en particulier des points de vue diastatique (quelle classe sociale) et diatopique (quelle région) ? Quel est le rapport entre le français écrit des textes les plus anciens, et la langue que les gens parlaient à l'époque ? Aurait-on raison de croire que les textes anciens nous présentent un portrait fidèle de l'usage oral, et de sa diversité dans l'espace ? Ces questions ont été au centre des préoccupations des linguistes diachroniciens qui se sont occupés du problème de la genèse du français « écrit » (on l'appelle parfois « littéraire », mais ce n'est pas très heureux car de nombreuses sources de grande importance pour la problématique de l'origine du français « langue par élaboration » ne sont justement pas littéraires, mais simplement utilitaires, comme les chartes), en particulier depuis une trentaine d'années.

Lorsque l'on se penche sur les textes de l'ancienne langue rédigés dans ce que l'on appellera très grossièrement « l'ancien français » (je me restreins donc ici à la France d'oïl ; il ne sera pas question aujourd'hui du francoprovençal ni de l'occitan), on constate qu'ils présentent des caractéristiques régionales ; des traits graphiques ou lexicaux dont la présence est plus marquée dans des textes originaires de telle ou telle région. Cela a amené les philologues à parler des « dialectes littéraires » de l'ancien français. On dira d'un texte marqué par un grand nombre de traits reconnus comme typiques de telle ou telle région qu'il a été écrit en dialecte « picard », « normand », « wallon », « lorrain », « champenois », « anglo-normand », « poitevin », « franc-comtois », etc. En d'autres mots, on ne peut pas parler pour la période de l'ancien français (surtout aux 12^e et 13^e siècles) d'une langue standardisée dont la variation sur l'axe diatopique serait réduite à très peu de choses : dans cette première étape de son histoire, le français écrit se présente à nous dans des textes marqués par une certaine différenciation dialectale.

La grande question que l'on s'est longtemps posée est la suivante : ces variétés régionales de l'ancienne langue, telles qu'elles nous apparaissent à l'écrit, sont-elles le témoin fidèle de la façon dont les gens parlaient à l'époque dans les différentes régions concernées ? Il est *a priori* très difficile de répondre à cette question : on ne sait absolument rien, de première main et de source sûre, de la langue orale de l'époque. On connaît très bien, en revanche, les patois modernes, qui avant de disparaître ont été l'objet d'investigations intensives et approfondies, au 19^e et au 20^e siècles. De nombreuses monographies dialectales, ainsi que des monuments de la recherche dialectologique galloromane comme l'*Atlas linguistique de la France* de Gilliéron¹, ou l'ALW² de l'école dialectologique belge, ou encore la somme que représente le *Französisches Etymologisches Wörterbuch* (FEW)³ de Walther von Wartburg, nous renseignent de façon exhaustive sur la nature des patois.

Or, que remarque-t-on lorsque l'on compare la langue des textes d'ancien français dialectalement marqués à celle des patois modernes ? En fait, *grosso modo*, il semble que les textes en question ne soient jamais *intégralement* rédigés dans un dialecte quelconque, ou plus précisément dans une langue qui aurait pu devenir ce que sont les patois modernes, mais seulement qu'ils laissent apparaître çà et là des traits qui sont effectivement représentés dans les patois

¹ GILLIÉRON, Jules; EDMONT, Edmond. *Atlas linguistique de la France*, Paris, Champion, 1902-1910, 35 fasc. en 17 vol.

² *Atlas linguistique de la Wallonie*. Tableau géographique des parlers de la Belgique romane d'après l'enquête de Jean Haust et des enquêtes complémentaires. Tomes parus : 1 (1953), 2 (1969), 3 (1955), 4,1 (1976), 5 (1991), 8,3 (1994), 9,1 (1987), 15,2 (1997), 6,1 (2006), 17 (2011).

³ WARTBURG, Walther von. *Französisches Etymologisches Wörterbuch. Eine Darstellung des galloromanischen Sprachschatzes*, Bonn, Klopp, 1928 ; Leipzig-Berlin, Teubner, 1934 et 1940 ; Basel, Helbing & Lichtenhahn, 1946–1952 ; Basel, Zbinden, 1955–2002 [25 volumes].

modernes. On trouve même souvent des formes diatopiquement marquées qui cohabitent dans le même texte avec des formes d'usage plus répandues. En outre, la fréquence de ces traits diatopiquement différenciés diminue avec le temps, comme si les scribes avaient voulu que leur langue écrite tende vers une plus grande neutralité du point de vue de la variation dialectale.

C'est la raison pour laquelle il faut être bien conscient du fait que la langue médiévale telle qu'elle nous apparaît dans les textes ne représente fort probablement pas la langue parlée à l'époque dans telle ou telle région, mais plutôt une certaine façon *d'écrire* la langue vulgaire (je dis *langue vulgaire* par opposition au latin) marquée dès les débuts par une tentative d'uniformisation. Les linguistes parlent, pour désigner cette réalité qui relève de la langue écrite dans une époque primitive de l'élaboration du français comme langue standardisée, de **scripta** (le romaniste wallon Louis Remacle est le père de ce terme, qui est devenu d'emploi courant aujourd'hui parmi les philologues francisants, en grande partie grâce aux travaux du romaniste suisse Carl Theodor Gossen). Il faut distinguer conceptuellement et terminologiquement ce que furent les dialectes parlés à l'époque, et dont on sait bien peu de choses (à part ce que l'on peut reconstituer à partir des parlers modernes), de ce qu'étaient les variétés écrites de cet ancien français qui sont parvenues jusqu'à nous à travers les textes. Le concept de **scripta** nous permet d'exprimer cette distinction. Il est généralement admis que les *scriptas* régionales ne reflètent pas l'ancien usage oral, mais que les traits qui les distinguent y trouvent leur source.

Parallèlement au problème de la distinction entre dialectes et *scriptas*⁴ se pose le problème de l'origine de ce français écrit déjà précocement unifié. Si les textes de l'époque présentent déjà une certaine uniformité, en particulier lorsqu'on les compare à ce que sont devenus les patois par la suite, à quoi cela est-il dû ? Plusieurs théories ont été défendues. On a beaucoup évoqué autrefois le prestige de la langue de Paris, langue de la cour, langue du roi ; on a aussi dit que ce qui devait être la langue parlée en Île-de-France était vite devenu digne d'imitation de la part des scribes de tout le domaine d'oïl, qui se seraient appliqués à la reproduire tant bien que mal, laissant seulement transparaître çà et là leurs particularités régionales. On utilisait autrefois le terme de **francien** pour désigner ce que l'on croyait avoir été le dialecte *oral* de l'Île-de-France. Ce concept a été largement critiqué, ne serait-ce que parce qu'il est très difficile de savoir ce que les locuteurs parlaient vraiment à l'époque, et ne peut encore être utilisé aujourd'hui que pour désigner la *scripta* d'un texte écrit dans une langue où ne semble transparaître aucun régionalisme, aucune caractéristique diatopique ; ce n'est donc plus un concept diatopique, mais simplement philologique (c'est-à-dire relié à la philologie, à l'étude de la langue des textes).

D'autres auteurs (en particulier Anthonij Dees) ont défendu une thèse différente, selon laquelle l'influence du centre, de Paris, n'était pas plus importante que celle d'autres régions. Ils en veulent pour preuve le fait que certains traits semblent s'être diffusés dans le temps et dans l'espace en suivant des trajectoires nord-sud ou est-ouest, par exemple, qui ne font que balayer Paris, alors que l'hypothèse traditionnelle aurait plutôt fait attendre une irradiation des traits innovateurs à partir du centre.

« Les cartes confirment que la perte de la flexion casuelle se propage de l'Ouest vers le Centre, alors que la tendance à placer le complément d'objet direct après le verbe se développe depuis le Sud-Est du pays d'oïl. Ces considérations permettent de poser deux hypothèses. L'une est que le

⁴ Certains font parfois le pluriel en *scriptae*, s'alignant ainsi sur la morphologie du latin, mais nous préférons traiter ce mot comme un emprunt intégré aux structures morphologiques du français en faisant systématiquement son pluriel en *-s* : *scriptas*.

fonds français commun des chartes du XIII^e siècle fut le fait d'une langue commune bien vivante dans les régions où les documents étaient écrits et ne devaient rien à un quelconque envahissement parisien. L'autre est que des traits du français commun du XIV^e siècle, que l'administration royale contribua indéniablement à répandre, ont pu parfois tirer leur origine de dialectes fort éloignés de Paris. » S. LUSIGNAN, dans CHAURAND (dir.) 1999:102.

En fait, si les textes de la région de l'Île-de-France semblent moins dialectalement marqués que les autres, c'est peut-être dans un premier temps simplement parce que cette région est centrale par rapport aux autres⁵, à la croisée de toutes les tendances et innovations, et non parce que le parler de Paris aurait eu dès le début un prestige déterminant. Cela dit, dans une étape plus avancée (14^e siècle) du processus de dérégionalisation des variétés écrites de langue vulgaire en domaine d'oïl, le modèle de langue écrite diffusé par les chancelleries parisiennes a dû jouir d'un prestige croissant.

Quoi qu'il en soit, il faut bien prendre conscience du fait que la langue française telle que nous la connaissons aujourd'hui remonte à une forme de langue d'abord écrite, donc jusqu'à un certain point artificielle, sorte de *koinè* (langue commune) dont aurait été éliminées très tôt les caractéristiques régionales les plus excentriques, et dont les rapports avec les variétés orales de l'époque restent obscurs. Il est clair que les parlers d'oïl de l'époque, bien que certainement moins différenciés les uns des autres qu'ils n'allaient le devenir quelques siècles plus tard, se présentaient sûrement sous une forme dialectalisée, ce qui est le propre de toute langue non encore standardisée, « en liberté ». Au sein de toute cette variété dialectale, comment le processus de sélection qui allait mener à une forme écrite plus ou moins uniformisée s'est-il déroulé ? Du point de vue diatopique, il semble que ce qui était commun au plus grand nombre s'est naturellement imposé ; du point de vue diastratique, il est vraisemblable que l'usage des sociolectes des locuteurs les plus instruits et les plus influents ait servi de modèle à la langue écrite.

On entend souvent dire que le français est, à la base, la langue de Paris et de l'Île-de-France. Ce n'est pas si simple. Parallèlement à la langue écrite, des registres de langue parlée distincts de la langue écrite en de nombreux aspects ont très longtemps – sinon toujours – existé ; il ne faut pas confondre langue de Paris et langue française « standardisée ». La langue française standard telle qu'elle s'est élaborée au cours des siècles a été très influencée par les sociolectes les plus élevés de la société française septentrionale, qui ne coïncident pas toujours, loin de là, avec le français parisien ou de l'Île-de-France des classes les plus humbles. C'est pourquoi on a pu parler d'un « patois » de Paris pour désigner cette réalité ; d'autres (dans le monde anglo-saxon, cf. Anthony Lodge) ont évoqué un « dialecte urbain ». Un des exemples les plus connus de cette dichotomie entre ancien français populaire parisien et français standard concerne la prononciation du trigramme <eau>, anciennement triphthonguée au moyen âge, qui est devenue [jo] (écrit <iau>) en parisien et francilien populaire (*troupiou*, *siau*, *biau*), alors que le français normé, influencé par les usages de la cour, optait pour [o] (en gardant la même graphie : *troupeau*, *seau*, *beau*). Il ne faut jamais oublier que la problématique de la genèse du français ne relève pas que de la variation diatopique, mais aussi de la variation diastratique.

L'étude de l'ancienne langue à travers les textes qui sont parvenus jusqu'à nous est un terrain parsemé d'embûches. Il importe d'en prendre conscience pour comprendre tous les désac-

⁵ « Ainsi, ce serait dans l'organisation de l'espace linguistique d'oïl elle-même que serait inscrite la place particulière d'un dialecte estiné à devenir directeur. C'est cette organisation qui fait, p.ex., qu'un dialecte situé au centre d'un domaine linguistique est moins divergent des autres, toutes choses égales par ailleurs, que ceux-ci ne le sont entre eux. » Yan Greub / Jean-Pierre Chambon, « Histoire des dialectes dans la Romania : Gallo-romania », HSK 3 (*Romanische Sprachgeschichte / Histoire linguistique de la Romania*), 2008, p. 2509.

cords qui existent sur la façon dont le français « langue par élaboration » est né. Voici les principaux facteurs qui rendent l'exploitation des textes anciens si malaisée :

- 1) Sources littéraires vs. sources « utilitaires » (chartes, documents notariaux). On a eu tendance au début à éditer de préférence des sources littéraires, leur trouvant évidemment plus d'intérêt intellectuel que les textes ennuyeux et stéréotypés des archives notariales. Malheureusement pour les problèmes de localisation des traits régionaux, les textes littéraires sont très difficiles à localiser dans l'espace et dans le temps, alors que les textes d'archives se caractérisent justement par le contraire : les chartes sont normalement datées et comportent souvent des indices textuels qui permettent d'évaluer quel fut leur lieu de rédaction. L'étude des scriptas régionales (je dis bien des scriptas et non des dialectes tels qu'ils devaient exister à l'oral) bénéficie évidemment davantage des données tirées des chartes que de celles tirées des textes littéraires. Cela dit, il faut signaler avant de clore ce point que la localisation géographique des chartes n'est pas toujours facile à déterminer avec précision, le scribe pouvant provenir d'une autre région que celle où la charte a été rédigée, ou la charte mentionnant deux localités assez distantes l'une de l'autre, ou ne mentionnant aucun toponyme qui aide au repérage, etc.
- 2) Originaux et copies : il faut savoir qu'au moyen âge, en l'absence de toute technologie permettant la reproductibilité à l'identique du texte écrit, les copies manuscrites étaient la seule façon de contribuer à la diffusion d'un ouvrage (dans le cas des textes littéraires), ou à sa conservation (dans le cas des chartes). Or, un texte recopié est susceptible d'être marqué par de nombreuses altérations : les cas où il nous est possible de comparer textes originaux et copies nous montrent que les scribes se permettaient à peu près n'importe quoi (changer la graphie, changer les mots, enlever, en ajouter, etc.). La valeur de témoignage linguistique d'une copie est évidemment très incertaine, et les données qu'on peut en retirer ne nous sont guère utiles dans une étude de géographie linguistique historique.
- 3) Circularité de l'identification des marques régionales : certains auteurs ont abusé d'un raisonnement circulaire consistant à dire qu'un texte est, par ex., picard, parce qu'on y trouve quelques traits généralement tenus pour picards, et que par conséquent tous les traits particuliers qui s'y trouvent sont picards parce qu'ils se trouvent dans un texte picard. Ce genre de démarche est dangereux car il peut entraîner le philologue dans une spirale d'imprécision croissante, de déductions reposant sur de bien maigres certitudes.
- 4) Non prise en compte de la distinction « langue parlée » vs. « langue écrite » d'une part (comme genre textuel), doublée d'une non prise en compte de la distinction « graphie » vs. « prononciation ».
Certains auteurs se sont imaginés, assez naïvement, que les textes écrits en langue vulgaire au moyen âge rendaient une image quasiment « polaroid » de la langue parlée à l'époque. Or, le fait d'écrire est en soi une activité artificielle, un rituel, quelque chose d'appris caractérisé par tout un ensemble de comportements acquis auprès de professeurs ainsi que de textes modèles ; les modèles en question pouvaient très bien avoir été rédigés en latin, voire en grec ancien, avec toutes les influences que cela a pu avoir sur la nature de ce français écrit naissant. Du point de vue du genre textuel, ni les chartes notariales, avec leur style répétitif et stéréotypé, ni la grande littérature épique, avec ses ambitions esthétiques, ni les plus anciens textes scientifiques, lourdement influencés par le latin, ne peuvent être considérés comme des témoins fiables de ce que fut la langue parlée.
Quant au problème de l'interprétation phonétique des graphies, il est de taille. La toute

première chose dont il faut prendre conscience est que l'alphabet romain hérité du latin qui était à la disposition des scribes médiévaux pour donner à la langue vulgaire une forme écrite n'était tout simplement pas conçu pour l'ancien français. De nombreux sons qui n'existaient pas en latin avaient fait leur apparition dans l'ancienne langue d'oïl, pour lesquels les scribes devaient inventer différentes combinaisons autrefois inconnues (consonnes palatales, plus ou moins transcrites <gn> et <ill> ; consonnes affriquées, transcrites par des digrammes, comme <ch> [tʃ], ou à l'aide de la cédille, comme <ç> [ts] ; voyelles, diphtongues et triphthongues, plus ou moins transcrites avec les moyens du bord : <oi>, <eau>, <ei>, etc.).

Or, comment savoir hors de tout doute que tel ou tel graphème ou combinaison de graphèmes correspond bien à tel ou tel son ? La phonétique historique a fait beaucoup de progrès dans ce travail de reconstruction, en tenant compte d'une multitude de facteurs (le point de départ, à savoir l'étymon latin, et le point d'arrivée, à savoir la forme telle qu'on la prononce aujourd'hui, en français et dans les patois ; l'étude des rimes ; l'étude de la métrique ; le traitement phonétique des emprunts ; les quelques témoignages métalinguistiques d'époque). Cela dit, il n'est jamais exclu que dans certaines circonstances on ne puisse arriver à y voir clair. Considérons deux cas de figure :

D'abord, il peut très bien arriver qu'une graphie à première vue divergente ne soit qu'une autre manière, dans une autre région, de transcrire le même son. On constate par exemple que dans les textes écrits en anglo-normand, le graphème <k> est très fréquemment employé là où on aurait simplement <c> ou <qu> dans d'autres scriptas. Il n'y a probablement pas de conséquences à tirer de cette graphie en ce qui concerne la phonétique de l'anglo-normand. Ce n'est qu'une habitude graphique, qui aide à repérer l'origine du scribe et du manuscrit mais qui ne nous fournit pas d'information sur la prononciation.

Inversement, une même graphie peut très bien avoir été prononcée de plusieurs façons selon l'origine du scribe. Ceci est un fait d'expérience que chacun peut observer dans son quotidien : que ce soit en français, en espagnol, en portugais ou en allemand, etc., les différences d'accent, même très fortes, ne se reflètent pas dans la graphie, qui est autant que possible unifiée (sauf cas exceptionnels où l'auteur veut absolument rendre dans la graphie une prononciation particulière, mais cela ne se fait guère que dans des genres textuels très précis comme le théâtre ou les passages narratifs des romans). Il n'est pas exclu que les textes écrits dans cette *koinè* médiévale qui s'est retrouvée assez tôt dépouillée de traits dialectaux fortement excentriques aient été lus à voix haute dans le respect de la prononciation locale, de la même façon qu'aujourd'hui un Américain et un Anglais liront le même texte chacun avec leurs accents respectifs (qui peuvent être très éloignés l'un de l'autre, comme vous le savez). [Il faut savoir que les textes littéraires relevant du style épique étaient souvent conçus pour être récités en public ; les lecteurs devaient certainement être influencés d'une part par leur usage personnel spontané, d'autre part par la nécessité d'être compris par leur auditoire, ce qui théoriquement pouvait les amener à prendre des libertés avec la lettre du texte.]

- 5) Parallèlement à ce problème des graphies, il faut signaler l'importance des écoles de scribes dans la transmission des scriptas régionales. Jusqu'à un certain point, les scriptas reflètent peut-être davantage des habitudes graphiques (significatives ou non du point de vue phonétique, ce qui n'est pas toujours facile à déterminer) que de véritables différences dialectales. Les scribes n'apprenaient évidemment pas à écrire tout seuls dans leur coin, mais bien sûr au sein d'écoles où des habitudes graphiques s'élaboraient et se transmettaient. Il n'y avait pas encore d'orthographe fixée à l'époque, comme c'est le cas aujourd'hui ; aucune norme ne s'était encore imposée, et chaque

école diffusait ses propres habitudes graphiques. Il faut donc être très prudent pour ne pas prendre une particularité graphique pour une véritable particularité dialectale.

- 6) Enfin, un dernier problème que j'aimerais évoquer : la distance entre les patois tels qu'on les connaissait encore au 20^e siècle, et ce que dut être la langue parlée dans les différentes régions du domaine d'oïl au moyen âge. Certains auteurs ont défendu l'idée que la langue écrite au moyen âge nous apparaît comme déjà passablement unifiée tout simplement parce que plus on remonte dans le temps, moins les patois devaient être différenciés les uns des autres (la différenciation étant une chose qui augmente avec le temps, en l'absence de toute norme centralisatrice). Cette vision des choses ne résiste toutefois pas aux faits. On constate par exemple que dans un même texte coexistent des formes non-marquées régionalement avec leur équivalent régional. On trouve aussi des textes beaucoup plus marqués dialectalement que d'autres. Cela montre bien que la variation dialectale existait, mais que le scribe pouvait choisir entre le terme régional et le terme plus général ; en d'autres mots, qu'il pouvait produire un texte partiellement ou entièrement *dérégionalisé*.

Ce qui a mis la puce à l'oreille de certains chercheurs, en particulier ceux issus de l'école de dialectologie wallonne (Jules FELLER, puis Louis REMACLE), est le phénomène de l'**hypercorrection**. Ce phénomène consiste à produire des formes fautives en appliquant à tort des règles de correspondance entre deux registres, deux parlers, voire deux langues. Pour donner un exemple contemporain, on constate qu'il y a des gens qui, voulant trop bien faire, disent « *je m'en suis rendu **contre*** » au lieu de « *rendu **compte*** ». Cela est dû au fait que, bien souvent, on ne prononce pas le *-re* final des mots se terminant par CONSONNE + <-re> : *quat'e*, *cont'e*, etc. Toutefois, on s'efforce de le restituer dans la diction soignée. Or, si l'on ne fait pas attention, on risque de mettre des <-re> là où il n'en faut pas : c'est un exemple de ce que l'on appelle « l'hypercorrection ». La citation ci-dessous illustre ce phénomène pour la période médiévale :

« On trouve dans une charte de 1271 publiée par M. Wilmotte la forme insolite *octable* pour *octave*. Comment cette francisation malhabile s'est-elle produite ? L'auteur ignorait le français *octave* ; il connaissait un *octâve* wallon, sans se douter que c'était un mot d'emprunt [au français central]. Or, d'autres mots wallons, comme *stâve*, *tâve*, *fâve*, *ovrâve*, *mariâve*, *vindâve*, correspondent à des mots en *-able* en français. Il francise donc *octâve* en *octable* ! » Jules Feller, cité dans Völker 2003 : 36.

De telles formes démontrent à l'évidence que l'auteur s'efforçait d'écrire dans une langue qui ne pouvait pas être le reflet fidèle de sa langue maternelle ; il n'essayait pas d'écrire en wallon, mais bien dans cette espèce de koinè scripturale de l'époque qu'on peut appeler *a posteriori* et d'une façon un peu anachronique le « français ».

5. Langues standardisées et variation diatopique

Quoi qu'il en soit de la genèse du français « langue par élaboration », on constate qu'à partir de l'époque du moyen français (dès le 14^e siècle) celui-ci s'est imposé peu à peu, gagnant des territoires de plus en plus étendus, se diffusant des classes sociales les plus prestigieuses aux plus humbles, et des noyaux urbains aux zones semi-rurales, puis rurales. Je ne voudrais toutefois pas vous laisser sur l'impression que la substitution du français aux autres dialectes ou patois galloromans, ou simplement la diffusion du français sur l'espace galloroman (et au-delà : Alsace, Bretagne, etc.), équivaut à la disparition de toute variation diatopique. Les langues par élaboration essayent de tendre, effectivement, vers une certaine uniformité dans l'espace ; cela dit, bien sûr, cette recherche de l'uniformité s'apparente un peu au mythe de

Sisyphé, à quelque chose d'inaccessible. Plus une langue se diffuse, plus justement elle risque de retomber dans la dialectalisation, en raison d'une foule de facteurs qui favorisent la différenciation diatopique : archaïsmes, innovations, influence des langues en contact. Le poids de la norme et les exigences de l'intercompréhension tendent à faire en sorte que la variation diatopique d'une langue standardisée reste à l'intérieur de certaines limites ; cela dit, cette variation est consubstantielle à la nature même de la langue, même d'une langue « par élaboration » (malgré le fait que la variation d'une langue « en liberté » soit évidemment beaucoup plus grande que celle d'une langue standardisée).

Dans certaines configurations internationales, une langue « par élaboration » peut même développer plusieurs « normes nationales », avec intercompréhension mais avec des processus de standardisation jusqu'à un certain point autonomes les uns des autres. C'est le cas, à des degrés divers, de toutes les langues européennes transplantées sur le continent américain (français au Canada, anglais au Canada et aux Etats-Unis, espagnol dans les différents pays d'Amérique latine et portugais au Brésil). C'est également le cas de ce que l'on appelle en Suisse alémanique le *Schweizer Hochdeutsch*, qui est une variété nationale d'allemand standardisé, distincte des dialectes alémaniques de Suisse mais aussi de l'allemand standard d'Allemagne, dont elle se différencie par de nombreux détails, à tous les niveaux (phonétique, orthographique, lexical, pragmatique, etc.).

6. Références bibliographiques

CHAURAND, Jacques (éd.), *Nouvelle histoire de la langue française*, Paris : Seuil, 1999 (2^e partie : S. LUSIGNAN, « Langue française et société : du XIII^e au XV^e siècle », en particulier pp. 91-124).

GREUB, Yan, *Les mots régionaux dans les farces françaises : Étude lexicologique sur le Recueil Tissier (1450-1550)*, Strasbourg : Société de linguistique romane, 2003 (lire l'introduction, pp. 1-22).

GREUB, Yan / CHAMBON, Jean-Pierre. « Histoire des dialectes dans la Romania : Galloromania », dans *Romanische Sprachgeschichte / Histoire linguistique de la Romania*, Berlin / New York : Walter de Gruyter, 2009, 2499-2520 (§ 214).

LODGE, Anthony, « Vers une histoire du dialecte urbain de Paris », dans *Revue de linguistique romane* 62 (1998), pp. 95-128.

VÖLKER, Harald, *Skripta und Variation : Untersuchungen zur Negation und zur Substantivflexion in altfranzösischen Urkunden der Grafschaft Luxemburg (1237-1281)*, Tübingen : Niemeyer, 2003 (lire chap. 1 et 2).