

Le fonds primitif

Le celtique

Dès le Vème s. AC, des tribus celtes venues des régions de l'Est, correspondant à la Bohême et à la Bavière actuelles, s'installent sur ce qui va devenir la Gaule (Gaule vient de *Galli*, terme latin qui désignait les Celtes). Les Gaulois parlent de nombreux dialectes qui dérivent tous du celtique. C'est une langue qui dispose d'un vocabulaire assez riche, notamment dans les domaines qui occupent le quotidien des Gaulois. Bon nombre de mots sont parvenus jusqu'à nous¹. Ce sont des termes ayant trait à l'agriculture (sillon, charrue, ruche...), la description des terrains (lande², boue, galet, breuil³, bruyère⁴...), les spécialités maisons (cervoise, brasserie, tonneau...), des noms de plantes et d'animaux (alouette, chêne, arpent⁵, bouc, bec, sapin, mouton...). En tout, environ septante termes sont d'origine gauloise dans le français actuel. Mais les « gauloiseries » sont surtout présentes dans les toponymes qui sont la source d'information la plus vivante à la fois sur la langue et sur la vie des Gaulois. Grâce à la toponymie, qui est l'étude des noms de lieux, on peut préciser les quatre types de lieux que connaissaient les Gaulois : les lieux défensifs, les lieux de marché, les lieux sacrés et les situations particulières.

Les lieux défensifs : on les reconnaît aisément grâce aux suffixes *-dunum*, *-durum* ou *-rato*, « forteresse, place forte », mais à condition de se souvenir que l'évolution phonétique a généralement effacé une partie des syllabes d'origine.

Exemple : *Lugdunum* → ... ; *Carbantorato* → ...

Les lieux de marché : ils peuvent être identifiés grâce au suffixe *-magus*, « marché »

Exemple : *Noviomagus* → ...

Les lieux sacrés : ils peuvent se reconnaître grâce à la forme *nemeto*, « sanctuaire ».

Exemple : *Nemetodurum* → ...

Les situations particulières : ils sont repérables grâce à des termes comme *lano*, « plaine », ou *bona*, « port ».

Exemple : *Mediolanum* → ...

Le latin

80% du vocabulaire français vient du latin.

Dès 57 AC, Jules César conquiert la Gaule. La conquête de la Gaule s'accompagne, dans les faits, de l'expansion de la langue latine, et ce grâce aux marchands, aux soldats, les artisans et... les prostituées. Ce n'est donc pas un latin littéraire, le latin de Cicéron, qui se répand mais bien celui des rues. Par ailleurs, il faut être conscient que, lorsque les Romains arrivent en Gaule, il y a près de cent peuples gaulois qui ne parlent pas tous la même variété de gaulois. Et que, par conséquent, chacun a pu apporter la touche de son « accent » dans sa prononciation

¹ Cfr feuilles en annexe.

² Étendue de terre où ne croissent que certaines plantes sauvages.

³ Vieilli : petit bois, champ. Ce mot a donné beaucoup de noms de lieu (Briel, Le Breuil...)

⁴ Petit arbrisseau des landes

⁵ Ancienne mesure agraire qui valait, selon les régions, de 20 à 50 ares (are : mesure agraire de superficie de cent mètres carrés).

du latin. Sous l'effet de ces contacts entre populations aux usages linguistiques différents, mais aussi en raison de nouveaux besoins de communautés qui se transformaient, le latin parlé en Gaule s'est modifié en se diversifiant et en donnant naissance à des parlers dont les différences étaient sensibles selon les régions.

Le germanique

L'occupation romaine de la Gaule avait pendant plus de cinq cents ans, apporté aux habitants de ce pays une nouvelle langue, en même temps qu'une autre civilisation : le pays avait été organisé, l'économie s'était développée, des routes avaient été construites et tout aurait été pour le mieux si ne s'était pas produit le grand bouleversement des invasions. Entre la chute de l'Empire romain (476) et l'apparition de ce que l'on a appelé le premier « monument » de la langue française, le texte des *Serments de Strasbourg* (842), ce sont près de quatre siècles d'obscurantisme pour les populations, qui sont pour nous quatre siècles d'obscurité, faute de documents. C'est pourtant une période décisive pour l'histoire du français, car c'est celle où s'est développé le processus de différenciation et dialectalisation.

Malgré la rareté des documents de cette époque, nous allons essayer de comprendre comment ont pu se produire les changements linguistiques dont nous constatons l'aboutissement au IX^{ème} siècle.

Tout d'abord, il convient de mettre la *Provincia* (correspondant aux actuelles régions de Provence, du Languedoc, du Dauphiné et de la Savoie). En effet, dans ces zones, la romanisation avait été rapide et durable. Cette région qui avait été une des zones où les Gaulois avaient eu le moins d'influence sur les Ligures qui occupaient le pays avant eux, est aussi une qui subira le moins les influences germaniques. C'est pourquoi, les dialectes provençaux sont restés proches du latin.

Les envahisseurs qui entrent en ligne de compte pour l'histoire générale de la langue aux alentours du V^{ème} siècle sont les populations germaniques : les Francs, les Burgondes et les Wisigoths. Parmi eux, les Francs ont joué le rôle le plus important car leur implantation en Gaule s'est prolongée pendant des siècles. En effet, bien avant le V^{ème} siècle, les Francs étaient présents sur le sol gaulois. Au V^{ème} siècle, Clovis a étendu le royaume franc dans toute la moitié nord de la France actuelle (à l'exception de la Bretagne).

Les Wisigoths s'étaient fait attribuer le sud de la Gaule mais en 507, Clovis étendra sa domination sur ce territoire.

Les Burgondes, eux, se sont installés dans la région de l'actuelle Savoie. Ils ont ensuite occupé un territoire correspondant approximativement aux régions administratives de Bourgogne, de Franche-Comté et de Rhône-Alpes.

Sur le plan linguistique, cette occupation du territoire par des populations d'origines diverses semble être en partie responsable des grandes divisions dialectales qui partagent encore aujourd'hui la France :

- dialectes d'**oïl** au nord du territoire (car on affirmait en disant *oïl*)
- dialectes d'**oc** au sud (où « oui » se disait *oc*)
- dialectes **francoprovençaux** dans une zone intermédiaire qui recouvre le bassin moyen du Rhône et s'ouvre au sud de Vichy, en englobant les régions de Lyon, de Genève et de Grenoble.

Cette différenciation dialectale en deux zones principales et une zone intermédiaire était probablement accomplie à la fin des invasions, mais pas la fragmentation en variétés dialectales plus petites.

La zone des parlers d'oïl correspond approximativement au territoire occupé par les Francs (dont l'influence linguistique sur la langue française a été considérable).

La zone méridionale, très tôt et profondément romanisée, n'a pratiquement pas été influencée par l'installation, d'ailleurs très brève, des Wisigoths. La zone intermédiaire, qui constitue le domaine francoprovençal, partage certains traits avec les parlers d'oïl et d'autres traits avec les parlers d'oc.

L'influence germanique est sensible dans le vocabulaire, où plus ou moins quatre cents mots sont d'origine germanique (et la plupart parmi eux, d'origine franque). Cela n'a rien d'étonnant si l'on se rappelle que la grande invasion franque du Vème siècle avait été précédée d'une longue cohabitation. Ainsi, on peut estimer que au moins la moitié des mots français recensés comme étant d'origine germanique est antérieure à l'invasion

L'influence germanique est aussi visible sur la prononciation. Le *h* qu'on appelle « aspiré » en français est traduit, en français moderne, par une absence de liaison (*les hanches* et non *les-z-hanches*) ou par une absence d'élision (*le hêtre* et non *l'hêtre*). Mais pourquoi une différence de traitement des *h* initiaux ? Il est en effet correct de dire *l'homme*... Ce sont les Francs qui en sont la cause. Il y eut une période où des mots comme *hêtre* et *hanche*, introduits en Gaule par des populations germaniques, se prononçaient avec un vrai *h*, qui se traduisait par une forte expiration d'air (alors qu'en latin, le *h* n'était plus prononcé).

Une autre trace se retrouve dans les toponymes. Suivez l'explication, en latin, l'adjectif pouvait se placer avant ou après le nom déterminé, selon sa fonction, épithète ou attribut. En ancien français, des formes telles que *des rouges manteaux*, étaient courantes. En français moderne, c'est la position après le nom qui est privilégiée. Or, dans les langues germaniques, c'est le déterminant qui précède toujours le déterminé : en français, on dit *la langue française* ; mais en allemand, *die französische Sprache* et, en anglais, *the French language*. Ainsi l'influence germanique peut encore se vérifier dans les noms de lieux : c'est la position de l'adjectif devant le nom qui a prévalu dans les régions d'implantations franque, tandis que dans les autres régions, l'adjectif se plaçait après le nom.

Exemple : au Nord, on retrouve *Neuville*, *Neuvelle*, *Neuveville* alors qu'au Sud, les *Villeneuve* sont plus fréquents.

Le morcellement linguistique du Moyen-Age

L'influence des envahisseurs germaniques peut expliquer une partie des différences régionales que l'on constate au Moyen-Age. Elle ne peut les expliquer toutes. Comment concevoir, en effet, qu'à partir du latin venu de Rome, on en soit arrivé quelques siècles plus tard à un morcellement linguistique tel que chaque région avait son dialecte ? Pour essayer de s'expliquer comment s'est produite cette différenciation en dialectes divers, il faut se rappeler les conditions de vie sous le régime féodal. Fondée sur les relations de vassal à suzerain, la vie s'organisait sur la terre du seigneur, autour du château. Le seigneur était donc amené à avoir de multiples rapports avec les paysans, dont il devait pouvoir se faire comprendre, mais il ne fréquentait guère ses pairs, sinon à de rares occasions. C'est donc à l'intérieur des limites d'un fief qu'ont dû naître et se développer des divergences, d'abord minimales mais qui ont pu s'accroître lorsque les conditions géographiques augmentaient l'isolement de deux domaines.

Emprunts au latin et au grec

Une seconde couche du lexique comprend les emprunts directs au latin depuis les débuts de la langue écrite (*Serments de Strasbourg*, 842) jusqu'au XVIème siècle par les savants et les lettrés ; ils permettent la constitution d'un vocabulaire abstrait

rendu nécessaire par l'évolution de la civilisation et des techniques. Déjà au VIII^{ème} siècle, Charlemagne avait fait appel à Alcuin et lui avait demandé de redonner un peu de splendeur et de pureté à un latin déjà très évolué. La « renaissance carolingienne » avait eu pour résultat de réintroduire de nombreuses formes du latin classique dans une langue qui ne lui ressemblait pas beaucoup. À cette époque, la langue française en formation a été perturbée dans son évolution naturelle : ce moment est capital dans le développement de la langue, qui devient dès lors doublement latine : une fois directement par évolution naturelle et une seconde fois, de façon artificielle, par injection ultérieure de formes savantes non altérées. Ainsi, par exemple, le mot *frêle*⁶ vient du mot latin *fragilem* : il a évolué normalement. *Fragile*, lui, a été emprunté tel que au latin tardivement. Dès lors, on comprend pourquoi le même mot latin se retrouve en français sous deux formes différentes : on les appelle les *doublés*. Il est intéressant de remarquer que les deux termes n'ont jamais le même sens.

Dès le XVI^{ème} siècle et plus massivement au XVII^{ème} et XVIII^{ème} siècles, on a formé des mots nouveaux en utilisant des éléments toujours empruntés directement au latin mais surtout parallèlement au **grec** pour faire face au besoin de nommer les inventions et les concepts nouveaux dans le domaine de la médecine, de la philosophie, des sciences humaines.

Emprunts aux autres langues

Les emprunts aux autres langues, témoignages des relations de notre pays avec d'autres peuples, sont multiples. Ils représentent un peu moins de 13% du vocabulaire. Sur 4200 mots d'origine étrangère, nos plus grosses dettes vont à :

- l'anglais, 1054 mots : *barman, football, look, match, sketch...*
- l'italien, 707 mots : *aquarelle, balcon, concerto, lavande, confetti, solfège...*
- l'arabe⁷, 215 mots : *amiral, algèbre, chiffre, luth, zénith, zéro...*
- l'allemand, 164 mots : *blocus, choucroute, ersatz, képi, valse, vasistas...*
- l'espagnol, 159 mots : *banane, casque, chocolat, guérilla, maïs, sieste...*
- le néerlandais, 153 mots : *matelot, maquereau, layette, gaufre...*

⁶ Dont l'aspect ténu donne une impression de fragilité.

⁷ Le français a très tôt emprunté à l'arabe ; le mot *amiral* (émir de la mer) se retrouve déjà dans la *Chanson de Roland* (XI^{ème} siècle).