

Premier chapitre

Origine et naissance du français

Comment, du latin parlé, émergea une langue nouvelle : le français

L'histoire des langues est à bien des égards une affaire de familles. Pour mieux en prendre conscience, il n'est pas inutile d'opérer un petit retour en arrière jusqu'au siècle des Lumières. Au XVIII^e siècle en effet, l'étude de la langue sacrée de l'Inde, le sanscrit, permet aux linguistes de mettre au jour des ressemblances troublantes entre cette langue très ancienne, contemporaine du latin et du grec, et les langues européennes. Dès 1787, l'Anglais William Jones affirme que ces ressemblances ne peuvent pas être le fait du hasard ni le résultat de contacts entre les populations. Il en déduit que le sanscrit, le latin, le grec, mais aussi les langues celtiques et germaniques, sont des sœurs, ou plutôt des cousines ayant un ancêtre commun. Ce dernier, qui a donné naissance aux langues indiennes et aux langues européennes, reçoit logiquement le nom d'« indo-européen ». Bien que nous n'ayons, hélas, aucune trace de cette langue « mère », les intuitions des premiers linguistes ont été amplement confirmées, au siècle suivant, par une nouvelle science alors en plein essor : la grammaire comparée¹.

Les langues indo-européennes forment un ensemble immense ; c'est pourquoi on a coutume de distinguer en leur sein des « sous-familles » (balte, slave, celtique, etc.). La famille « celtique », dont fait partie le gaulois, ne nous est pas indifférente et nous verrons d'emblée que le français moderne

1. On considère souvent que le fondateur de cette nouvelle science est l'Allemand Franz Bopp, auteur d'une *Grammaire comparée des langues indo-européennes* (1833-1849 pour la première édition).

recèle quelques vestiges de ce lointain parler gaulois. Mais pour nous, la famille la plus importante est sans conteste celle des langues italiennes, à laquelle appartient le latin ; celui-ci est en effet l'ancêtre d'une nouvelle famille, celle des langues romanes, qui regroupe l'italien, l'espagnol, le portugais, le roumain et, *last but not least*, le français.

L'Histoire est très longue, et ce livre ne tient pas à l'être ; c'est pourquoi, dans ce premier chapitre, nous ne remonterons pas jusqu'aux Indo-Européens, peuple lointain (dans le temps et dans l'espace) dont nous ne savons du reste pas grand-chose. Nous prendrons en revanche le temps d'évoquer l'héritage de nos « ancêtres » gaulois et germaniques ; nous observerons aussi la façon dont le latin parlé évolua peu à peu pour donner naissance à une langue nouvelle et autonome : la nôtre.

1/ Nos ancêtres les Gaulois ?

Les manuels du XIX^e siècle, pour des raisons idéologiques et politiques qu'il serait un peu long de justifier ici, se plaisaient à faire de nous, Français, les descendants directs des Gaulois. Il est vrai que ces derniers, volontiers représentés comme de vaillants guerriers chevelus, étaient aussi de bons cultivateurs et de grands amateurs de bière ; ainsi nous ont-ils légué le tonneau, la charrue, mais aussi les braies – sorte de pantalon guère utilisé aujourd'hui mais dont il nous reste au moins le diminutif, la fameuse « braguette ». Nous devons aussi aux Gaulois, mais fort indirectement, notre emblème national ; en latin, Gaulois se dit en effet *Galli* (habitants de la *Gallia*) ; or, au XVI^e siècle, les humanistes rapprochèrent ce nom de celui du coq : *gallus* en latin. Tout naturellement, le coq « gaulois » devint donc l'animal emblématique des Français, descendants des Gaulois, et il le reste à l'époque moderne : il figure notamment sur les maillots de l'équipe de France de football.

Arrivés autour du V^e siècle en Gaule¹, les Celtes parlaient – évidemment ! – des langues celtiques très diverses dont le tort principal – pour nous en tout cas – était de ne pas être écrites². Parmi les explications invoquées, on retiendra que les druides, qui faisaient alors office de prêtres,

1. Les Celtes, qui se rattachent aux populations indo-européennes, arrivèrent en Europe occidentale en deux vagues successives. Poussés par les Germains, ils s'installèrent dans des régions précédemment occupées par des populations non indo-européennes dont nous ne savons pas grand-chose ; de leur langue ne subsistent que de rares noms de lieux et ceux de nombreux cours d'eau tels que la Loire, la Seine, le Rhône, la Garonne, l'Aube, la Saône ou encore la Duranc.

2. Les rares traces écrites que nous possédions sont des inscriptions votives ou des devises sur des pièces de monnaie ; on citera aussi le calendrier de Coligny, une petite commune de l'Ain ; découvert à la fin du XIX^e siècle et aujourd'hui conservé au Musée gallo-romain de Fourvière, ce calendrier luni-solaire présente 5 années de 12 mois (de 29 ou 30 jours) et semble dater du II^e siècle après J.-C.

de juges et de maîtres d'école, considéraient le savoir comme trop pur, trop précieux pour pouvoir être couché sur des matériaux d'origine animale ou gravés dans la pierre. Désireux de le préserver sans le souiller, ils choisirent donc de le transmettre sous une forme exclusivement orale et de le conserver dans la seule mémoire humaine. Selon Jules César, les futurs druides passaient même parfois vingt ans à étudier et retenir par cœur des milliers de formules sacrées ; on imagine aisément les inconvénients d'une telle pratique : il ne nous reste quasiment rien de la langue gauloise et du savoir des druides.

Quelques (rares) traces du gaulois en français moderne

Les Gaulois nous ont légué environ 150 mots d'usage courant (*alouette*, *briser*, *jante*, *charrue*...) dont nous reparlerons dans le chapitre consacré au lexique. Ils ont aussi influencé la prononciation du latin vulgaire : plusieurs philologues pensent que le passage, en langue d'oïl, du *u* latin, prononcé « *ou* », au *u* français, serait dû à l'influence gauloise¹ ; certaines palatalisations et le maintien du *s* final s'expliqueraient de la même façon. Les Gaulois seraient aussi les responsables de notre étonnant « quatre-vingts » ; leur façon de compter, typiquement celtique, utilisait en effet le nombre 20 pour base ; il s'agit là d'une numération « vicésimale » qui resta courante sous l'Ancien Régime : l'hospice créé par Louis IX vers 1260, qui regroupe 300 lits (15×20), porte naturellement le nom de « Quinze-Vingts » ; au xvii^e siècle encore, il n'est pas rare d'entendre « six-vingts » plutôt que « cent vingt ». Bien que l'usage ait généralisé l'emploi de « cent vingt » en français moderne, *quatre-vingts* n'a pas été supplanté (en France en tout cas) par la forme *octante*.

Le déclin du Gaulois et sa disparition ne s'expliquent pas seulement par des pratiques culturelles spécifiques. Lorsque les Romains conduits par César envahirent la Gaule, au 1^{er} siècle avant J.-C., celle-ci se romanisa de manière progressive et profonde. Pendant près de 500 ans, la fameuse période gallo-romaine, le gaulois et le latin parlé coexistèrent ; au vi^e siècle encore, le témoignage de Grégoire de Tours atteste la survivance de la langue gauloise. Mais il se passa finalement sur notre territoire ce qui s'était déjà passé en Italie : le latin, qui n'était à l'origine que la modeste langue d'un petit peuple installé à Rome et dans la région du Latium (c'est-

1. Nous verrons cependant que cette hypothèse ne fait pas l'unanimité parmi les philologues : la plupart d'entre eux estiment que la palatalisation du *u* latin en *u* « français » (noté [y] dans l'API) est due aux Germains.

à-dire la vallée inférieure du Tibre et les régions avoisinantes), s'imposa au détriment du gaulois : comme dans toutes les provinces de l'Europe occidentale, la domination politique romaine se traduisit par une domination linguistique¹.

Le facteur politique n'est certes pas négligeable ; il faut néanmoins compléter et nuancer cette approche belliqueuse en rappelant que les Romains, où que ce fût, ne tentèrent pas d'imposer leur langue par la force. En fait, les Gaulois (comme les Ibères en Espagne) prirent très vite conscience que le latin, langue de civilisation et langue commerciale, constituait pour eux un instrument de promotion sociale non négligeable² ; le latin devint aussi rapidement, ne l'oublions pas, la langue de la religion chrétienne³ ; enfin, les langues celtiques, dont fait partie le gaulois, étaient celles qui, en Europe, se rapprochaient le plus des langues italiques, dont relève le latin : cette parenté favorisa sans doute le bilinguisme des populations indigènes⁴.

Le breton descendrait-il du gaulois ?

Auteur de nombreux ouvrages sur la langue bretonne, le chanoine François Falc'hun (1901-1991), linguiste et universitaire breton, soutint dans les années 1950 que le gaulois, contrairement à ce que l'on imaginait, n'avait pas complètement disparu. Selon lui en effet, le dialecte vannetais contemporain était une survivance gauloise peu influencée par l'apport breton ; quant aux autres dialectes bretons, ils étaient un gaulois « simplement plus marqué par la langue des immigrants d'origine insulaire⁵ ». Cette thèse déplut fortement aux nationalistes bretons ; elle impliquait en effet que les Bretons étaient encore plus « français » que les Français puisque leur région était l'ultime lieu de conservation de l'ancienne langue gauloise. Mais la thèse de Falc'hun fut surtout remise en cause pour des raisons scientifiques : à la fin des années 1960, le chercheur anglais Kenneth Jackson démontra que le breton était le descendant du brittonique, langue parlée dans la province romaine de *Britannia*. Plus précisément, Jackson rappela que le déclin de l'Empire romain d'Occident avait poussé des

-
1. En Europe occidentale seulement ; car en Orient, le latin ne parvint pas à supplanter le grec.
 2. Dès le 1^{er} siècle de notre ère, on sait que les nobles gaulois envoyaient leurs enfants dans les écoles romaines pour qu'ils s'y familiarisent avec la langue et la culture de l'occupant.
 3. Il semble que les chrétiens, au 1^{er} siècle de notre ère, employaient déjà la langue latine pour communiquer entre eux ; par ailleurs, une traduction latine de la Bible due à saint Jérôme, la célèbre *Vulgate*, fut rédigée dès le 4^e siècle.
 4. Attention cependant car ce dernier point est aujourd'hui discuté : selon Pierre-Yves Lambert, spécialiste de la langue gauloise, les traits communs à l'italique et au celtique ne sont en fait pas si nombreux et l'unité du groupe italo-celtique mérite d'être discutée voire mise en cause (voir Pierre-Yves Lambert, *La Langue gauloise*, Paris, Errance, 2003, p. 13).
 5. François Falc'hun, *Perspectives nouvelles sur l'histoire de la langue bretonne*, Paris, Union Générale d'Éditions, 1981, p. 530.

communautés entières de Britto-Romains à émigrer dans une partie de l'Armorique ; ces émigrants avaient apporté avec eux leur culture, leur organisation mais aussi leur langue, qui devait dès lors être considérée comme l'ancêtre « officielle » du breton moderne.

NB : pour la petite histoire, on signalera que la thèse de Falc'hun, délaissée par les scientifiques, a du moins survécu dans des ouvrages connus de tous ; c'est d'elle en effet que s'inspirèrent Goscinny et Uderzo quand, dans les années 1960, ils situèrent en Armorique le fameux village gaulois d'Astérix, dernier résistant à l'envahisseur romain.

2/ Du latin classique au français : le rôle décisif du « latin vulgaire »

Le gaulois et – *a fortiori* – les langues pré-celtiques n'ont guère laissé de traces dans notre langue actuelle. En fait, à l'instar des autres langues romanes, le français est pour l'essentiel le descendant du latin ; encore faut-il savoir de quel latin on parle : il ne faut pas confondre le latin écrit, classique et cicéronien, dont quelques privilégiés continuent d'apprendre la richesse et les difficultés sur les bancs du collège, et le latin parlé par le peuple – pas cicéronien du tout celui-là –, couramment dénommé latin vulgaire (du latin *vulgus* désignant précisément le « peuple¹ »). De fait, c'est bien ce latin parlé et instable qui constitue notre ancêtre direct ; mal connu hélas, puisqu'il est par essence une langue orale, le latin vulgaire subsiste à l'état de « traces » dans quelques graffitis ou lettres.

Pour savoir à quoi ressemblait cette langue, les linguistes doivent « traquer », dans les textes tardifs, les solécismes commis par leurs auteurs ; car ces « écarts » par rapport à la norme classique sont souvent révélateurs des usages populaires. Ils peuvent aussi s'appuyer sur quelques textes de grammaire normative datant de la période du latin tardif, dans lesquels sont condamnés certains usages de la langue orale. La plupart du temps néanmoins, il leur faut observer les diverses langues romanes et les comparer au latin classique afin de *reconstituer* les formes intermédiaires employées par les locuteurs. Prenons l'exemple des noms féminins pluriels de la 1^{re} déclinaison : en latin classique, au nominatif, « roses » se dit *rosae*. Pourtant, en français médiéval, on trouve la forme *roses* au cas équivalent (appelé alors « cas sujet ») ; or, il est impossible de faire dériver *roses* de *rosae*. Les philologues en ont déduit que la forme *rosae*, en latin vulgaire, avait été remplacée par **rosas* ; comme celle-ci n'est pas attestée à l'écrit,

1. L'expression « latin vulgaire » date des années 1870. Elle s'inspire d'une expression cicéronienne, *sermo vulgaris*, qui désigne la langue « populaire »

on la fait précéder d'un astérisque, mais son existence ne fait aucun doute puisque seule une forme terminée en *-as* peut expliquer la désinence *-es* de nos *roses* françaises.

Le latin, un ancêtre prestigieux... et parfois encombrant

La langue française, au Moyen Âge, ne fait pas l'objet d'une réflexion théorique poussée. Dès le XII^e siècle cependant, les érudits ont bien conscience de son origine latine. Une telle filiation est certes prestigieuse ; mais elle est, d'un point de vue politique, un peu gênante : on craint que la subordination du français au latin ne signifie la dépendance du royaume de France à l'Empire germanique (successeur de l'Empire romain) ou à l'Italie. C'est pourquoi on réaffirme fort à propos, sous le règne de Philippe Auguste (1165-1223), l'origine troyenne des Gaulois et des Francs¹ ; cette nouvelle filiation, plus prestigieuse encore, est aussi moins encombrante. Au XVI^e siècle, certains philologues prétendent même que ce sont les Gaulois qui ont fondé la ville de Troie : nos ancêtres auraient ensuite diffusé leur savoir dans toute la Grèce avant que celui-ci ne revienne en France !

Ces hypothèses, aussi farfelues soient-elles, ont été relayées par nombre d'érudits : le grand savant Nicole Oresme (1322-1382²), dès le XIV^e siècle, soutient que le français descend du grec. On affuble alors le nom des *Francs* d'un étymon grec³ ; quant à la capitale du royaume, Paris, elle porte bien sûr le nom grec que les envahisseurs francs lui ont attribué⁴. Les humanistes de la Renaissance ne sont pas en reste : Guillaume Budé, au XVI^e siècle, fait dériver du grec des termes aussi usuels que *pantoufle*, *dîner* ou *agrapher*⁵. La vérité est bousculée, mais quel prestige !

-
1. La croyance en une origine troyenne des Francs n'est pas neuve : elle est attestée depuis le VII^e siècle. Mais c'est bien au XIII^e siècle que les philologues la remettent à l'honneur, pour des raisons évidemment plus politiques que linguistiques !
 2. Nicole Oresme, né à Bayeux en 1325, fut un des plus grands savants de son temps : il publia divers traités scientifiques en latin (sur la cosmologie, les mathématiques, la physique), traduisit et commenta en français plusieurs ouvrages d'Aristote (*l'Éthique*, *la Politique*, *l'Économique*, *Du ciel et du monde*), et, par ses propres traités scientifiques en langue française (je pense en particulier au *Livre de divinations* et au *Traité de la sphère*), il contribua à donner au français le statut de langue scientifique et philosophique capable de rivaliser avec le latin.
 3. Selon Aimoin de Fleury, *franc* signifie « vaillant » en langue attique (voir Frédéric Duval, *Mille ans de langue française, histoire d'une passion*, Paris, Perrin, 2007, p. 212).
 4. Plus généralement, Guillaume le Breton, auteur de la *Philippide* (poème écrit vers 1220 en l'honneur de Philippe Auguste), soutient que les Francs, quand ils ont envahi la Gaule, ont attribué des noms grecs aux villes qu'ils fondaient. Paris n'est que le plus célèbre d'entre eux.
 5. À la fin du Moyen Âge et au début du XVI^e siècle, les érudits diffusent l'hypothèse d'une transmission de la langue grecque par l'intermédiaire des druides gaulois. Là encore, le cheminement est complexe et historiquement (très) discutable, mais le sentiment national y trouve son compte.

Le latin vulgaire est né lorsque le latin est devenu une langue de culture, aux alentours du III^e siècle avant J.-C. ; à partir de ce moment en effet, une rupture s'est opérée entre la version écrite de la langue latine, plus stable et soumise à des normes, et sa version orale, plus prompte à évoluer et sans doute moins rigoureuse. Contrairement au latin écrit, le latin vulgaire eut donc la « chance » de ne pas être bridé par des règles syntaxiques ou orthographiques ; c'est pourquoi le premier évolua très peu et très lentement (il existe peu de différences entre le latin classique du I^{er} siècle avant J.-C. et le latin « impérial » des siècles suivants¹) tandis que le second put facilement s'adapter aux besoins linguistiques de son temps : cette souplesse joua un rôle majeur dans son expansion et sa vitalité.

NB : le latin vulgaire étant la version parlée de la langue latine, il a toujours coexisté avec le latin écrit et sa dénomination n'est pas liée à une époque précise ; il ne faut donc pas le confondre avec le latin tardif, qui est un latin écrit datant des III^e-V^e siècles après J.-C.

Le latin vulgaire, une langue plus simple et plus expressive

Assez logiquement, la langue latine parlée se distingue de la langue écrite par une simplicité et une expressivité accrues.

- Sur le plan **morphologique**, de nombreux verbes classiques déponents (à la conjugaison fort compliquée) sont remplacés à l'oral par des verbes plus simples, tel *moriri* par **morire* ; les infinitifs *esse* et *velle* cèdent la place à deux formes plus « normales » (si l'on peut dire...) : **essere* et **volere*². Du côté des substantifs, les imparisyllabiques de la 3^e déclinaison tendent à disparaître et les noms du type *flos, florem, floris...* se déclinent désormais **floris, florem, floris...* Les substantifs de la 5^e déclinaison – peu nombreux donc atypiques – sont en général remplacés par des doublets en *-a* se déclinant sur le modèle de *rosa* : *materia* (*matière*) élimine *materies* tandis que *luxuria* et *glacia* éliminent respectivement *luxuries* et *glacies*³.

1. À la suite d'Henriette Walter, qui s'interroge dans son ouvrage consacré à *L'Aventure des langues en Occident* sur cette stabilité étonnante du latin écrit, nous affirmerons que les Romains, nation puissante mais regroupant des populations très diverses, avaient « besoin d'un corps administratif solide, régi par des lois précises, au moyen d'une langue sans équivoque » (H. Walter, *L'Aventure des langues en Occident*, Paris, Robert Laffont, 1994, p. 116).

2. Ce sont ces formes non attestées qui sont à l'origine de nos infinitifs *être* (puis *être*) et *vouloir* (devenu *vouloir*) ; *esse* et *velle* n'ont eu aucune descendance.

3. On notera néanmoins que les termes les plus usuels, tel *res* qui désigne la « chose », résistèrent bien mieux aux réfections morphologiques ; cela est facile à comprendre : une forme fréquemment employée est davantage ancrée dans la mémoire des locuteurs et ne peut aisément céder la place à une autre.

Enfin, toutes déclinaisons confondues, les diminutifs, considérés plus expressifs, se multiplient au détriment des formes simples (*auricula* remplace *auris* tandis que *geniculum* se substitue à *geni*).

- Sur le plan **syntactique**, le latin vulgaire généralise l'emploi des prépositions ; déjà présentes dans le système linguistique indo-européen, elles se développent dès le latin classique ; leur nombre et leur emploi restent cependant limités tant que les 6 cas de la langue latine (nominatif, vocatif, accusatif, génitif, datif et ablatif) sont bien respectés et clairement différenciés ; le génitif par exemple suffit à marquer l'appartenance. Mais la langue populaire est moins rigoureuse quant à l'emploi des cas (dans les graffiti de Pompéi apparaissent des confusions entre accusatif et ablatif)¹ ; par ailleurs, à l'oral, on sait que les gens du peuple prononçaient de façon très affaiblie, voire inaudible, les consonnes finales ; or, comment différencier le nominatif *rosa* de l'accusatif *rosam* si le *m* final n'est plus articulé ? Face à cette confusion casuelle, les prépositions deviennent un outil linguistique indispensable ; précisant et clarifiant les rapports syntaxiques unissant les différents termes de la phrase, elles s'imposent dans toutes les langues romanes².

3/ Le français : « la plus germanique des langues romanes »

Le français est une langue romane d'origine latine, c'est un fait acquis. Mais, selon Henriette Walter, « si l'on veut caractériser d'une phrase la langue française, on peut dire que c'est la plus germanique des langues romanes »³. Le nom même de notre langue, le « français », vient d'ailleurs de celui du peuple franc, qui pénétra en Gaule aux premiers siècles de notre ère ; dès le III^e siècle en effet, des mercenaires francs à la solde des Romains vinrent s'installer dans l'extrême nord de la Gaule. Cela dit, l'influence franque se fit surtout sentir à partir du V^e siècle, lors des fameuses « invasions barbares » qui virent déferler sur notre territoire les Alamans (en Alsace), les Francs (au nord, à l'est et dans le bassin parisien), les Burgondes (dans un territoire correspondant à l'actuelle Bourgogne) et les Wisigoths

1. Dès l'époque impériale (II^e-III^e siècles après J.-C.), on suppose que la langue populaire tendit à réduire la déclinaison à deux cas, le nominatif et l'accusatif (c'est-à-dire le cas sujet et le cas régime) ; à l'écrit en revanche, la flexion du latin classique put se maintenir. Nous reviendrons sur ce point lorsque nous aborderons la morphologie du groupe nominal (chapitre VI).

2. L'emploi des prépositions ne fait pas tout ; une autre évolution syntaxique contribua à clarifier la syntaxe en latin vulgaire : l'ordre des mots, jusqu'alors fluctuant, tendit alors à se fixer. En l'absence de différenciation casuelle claire, on ne pouvait plus se contenter, en effet, d'affirmer que *Bella Eulalia juvene agricola amat* (qui est sujet, qui est objet ?) ; on prit donc l'habitude de placer en tête le sujet, suivi du verbe puis du COD : on pouvait dès lors être certain que c'était bien la belle Eulalie qui était amoureuse du jeune paysan.

3. Henriette Walter, *L'Aventure des langues en Occident*, op. cit., p. 225.