



# Réflexion sur l'étymologie

COMMUNICATION D'ANDRÉ GOOSSE  
À LA SÉANCE MENSUELLE DU 13 MARS 1993

L'origine des mots intéresse beaucoup de gens, et cela depuis longtemps, comme le montre l'ancienneté du mot *étymologie*. Mais on a pu mettre des choses diverses sous cette étiquette.

C'est un mot grec formé sur l'adjectif ετυμος, qui signifie « vrai ». L'étymologie a d'abord été la recherche du vrai dans et par les mots. Cela suppose un rapport étroit et nécessaire entre le mot et la chose qu'il représente, rapport facile à admettre pour ceux qui pensaient que le langage avait été inventé par les philosophes.

L'étymologie n'était donc pas une enquête historique : on ne disait pas que tel mot *vient* de tel autre mot, mais qu'il est lié par le sens avec tel autre mot ou tels autres mots, en quelque sorte qu'il *est* cet autre mot ou ces autres mots.

J'emploie le pluriel pour deux raisons. La première est qu'une étymologie ainsi conçue peut très bien coexister avec un autre rapprochement, fondé sur un autre aspect de la réalité désignée. La deuxième est que la méthode consistait souvent à tronçonner le mot qu'on voulait expliquer. L'exemple le plus fameux est sans doute le latin *cadaver*, dans lequel on retrouvait les syllabes initiales de *caro data vermibus*. Celui-ci est du Moyen Âge misogyne : *Eva = extra vadens*, « par ce, si dient li philosophe, que feme va volentiers hors de voie de sapience et de raison ». Celui-ci traite plus hardiment ce qu'on ose à peine appeler la phonétique : *gladius* « le glaive » = *gulam dividens*, littéralement « qui sépare la gueule ». Plus poétique et plus simple : les *alouettes louent* Dieu.

Ce sont là des jeux de mots, si l'on veut, des calembours, mais pris au sérieux, car ils s'harmonisent avec le système platonicien comme avec le symbolisme du Moyen Âge chrétien.

Ils ne disparaissent pas ensuite. Somaize, au dix-septième siècle, prend la précaution de prévoir une évolution « phonétique » : *casquette* est pour *casquotte*, avant de lancer son explication : *capitis salus quotidiana*, le salut quotidien de la tête.

Vous refusez de prendre au sérieux l'ami des précieuses ? Et le grave Joseph de Maistre ? Il fait venir *ancêtre* de *ancien être*, *beffroi* de *bel effroi* et *sortir* de *se-hors-tir*, « se tirer dehors », comme si le français de 1821 s'expliquait tout entier par le français de 1821. On pense au faussaire célèbre qui réussit à vendre une lettre en français plus ou moins du seizième siècle dans laquelle Lazare remerciait Jésus de l'avoir ressuscité.

Et Claudel ? « Nous ne naissons pas seuls. Naître, pour nous, c'est connaître. Toute naissance est une connaissance. » Ou encore : « *Naître* (avec l'initiale négative), c'est-à-dire être ce qui n'est pas. » À côté de ces textes célèbres, il y en a bien d'autres, par exemple dans le *Journal*. À propos du latin *mens* « esprit », Claudel glose : « Me-ens, ce qui est en moi, moi en présence de mon être et le réalisant. » Le nom féminin *tempe* est suivi du mot *étymologie*, lui-même suivi de cette alternative : « *tempus ? templum ?* » Même la physionomie graphique des mots est mise en relation avec leur sens : « Moustique - ti : le dard, la piqûre - m : l'insecte posé suçant - osqe : le rond, le vertige armé ; u u : le sifflement pointu - le zigzag insaisissable avec l'i qui revient sur lui-même. » Je souligne le fait que les u censés exprimer le « sifflement pointu » sont tout à fait absents du mot prononcé.

Sont-ce là jeux gratuits de poète, rêveries, imaginations libérées ? Certains textes feraient croire le contraire : « Entre un objet quelconque, entre un quelconque fait, sentiment ou action, et sa représentation sonore ou graphique, il n'y a pas apparemment de rapport (*quoique personnellement avec Platon je sois persuadé du contraire*<sup>1</sup>). » Certaines de ces *étymologies* sont d'ailleurs prises comme

---

<sup>1</sup> Cité par M.-Fr. Guyard, *Claudiel et l'étymologie*, dans *Cahiers de l'Association internationale des études françaises*, mai 1959, p. 298-299.

de véritables arguments, comme des points de départ pour une réflexion théologique ou mystique. Claudel est aussi un homme du Moyen Âge.

Les poètes ne sont pas seuls. « Chaque fois qu'une étymologie m'intéresse, me retient, m'amuse, écrit Duhamel, les spécialistes entrent en transe [formule un peu surprenante à propos d'austères érudits !] et me démontrent aussitôt que cette étymologie est fantaisiste. » Les usagers ont peine à s'accommoder de ce qu'on appelle l'arbitraire du signe, ainsi que de son corollaire : l'infinie relativité du langage à travers l'espace et à travers le temps. L'étymologie qui plaît à Duhamel comme à bien d'autres est celle qui lie étroitement le mot à la chose désignée, celle qui révèle des origines pittoresques et suggestives. D'où les historiettes imaginées pour expliquer les mots apparemment difficiles. Un exemple cocasse du dix-septième siècle : les Goths mutinés, ayant mis à mort leur chef Alaric, « beuvoient par derision à sa santé en proferant ces mots A Ti Alaric Got » ; voilà, selon Fleury de Bellingen (*L'étymologie ou explication des proverbes français*, 1656), d'où vient à *tire-larigot*. Mais c'est au vingtième siècle qu'André Billy, parlant de l'origine du nom *bouillabaisse*, citait sans rire une *abbesse* appelée *Bouille* ou encore le bandit Gaspard de Besse : « On disait sur le marché : Ne touchez pas à ça : c'est la *bouille* à Besse ! »

Le fossé entre l'étymologie sérieuse et l'étymologie qui séduit est particulièrement visible dans le domaine des noms de lieux, dont on a peine à admettre qu'ils perpétuent le souvenir d'un obscur propriétaire, une configuration passagère des lieux ou d'autres choses banales et terre à terre. L'imagination remédie à cette grisaille. Dans le toponyme de la région liégeoise *Dolembreux*, la deuxième partie est un mot signifiant « marais » ; au début de ce siècle, un abbé proposait de chercher ici le souvenir d'une bataille perdue contre César : *Dolor Eburonum*. Rabelais a tiré des effets plaisants de pareilles étymologies. Le géant Gargantua a payé sa bienvenue aux Parisiens en les arrosant comme vous savez ou comme vous devinez : « Quand furent au plus hault de l'Université suans, toussans, crachans et hors d'halene, commencerent à renier et jurer [...] : “ Carymary, caramara ! par sainte Mamye, nous son baignez par rys ! ” Dont fut depuis la ville nommee *Paris*, laquelle auparavant on appelloit *Leucece*, comme dit Strabo, lib. iiij, c'est à dire en grec, *Blanchette*, pour les blanches cuisses des dames dudict lieu. »

Il convient d'ajouter que, plus d'une fois, dès l'Antiquité, des érudits ont fait des propositions moins superficielles et moins aventurées que celles que j'ai citées, ou plus proches d'une véritable filiation que d'une analogie approximative ou tarabiscotée. Certaines sont encore admises : comme le rapprochement de *cadaver* et du verbe *cadere* « tomber ». La sémantique a son rôle dans l'étymologie, et, armé de bon sens (c'est le cas de le dire), on peut reconstituer des cheminements vraisemblables.

Au seizième siècle — et cela coïncide avec un recul de l'esprit scolastique —, on s'intéresse à la langue pour elle-même et non plus seulement pour atteindre la vérité à travers elle, et les recherches étymologiques vont concerner les rapports entre les langues. Mais c'est aussi le temps où l'on découvre avec enthousiasme l'Antiquité grecque. De même que Jean Lemaire de Belges, Ronsard et d'autres voyaient ou feignaient de voir dans les Français des descendants des Troyens, des érudits contemporains s'efforçaient de rattacher la langue française à la grecque. D'autres, plus religieux, expliquaient tout par l'hébreu, qui se serait corrompu au temps de la tour de Babel et aurait alors donné naissance aux autres langues.

De semblables théories ne cherchent pas appui sur des faits historiques véritables ; elles supposent que chacune des langues a toujours été telle qu'on la pratiquait ou connaissait ; pour les faits particuliers, elles se satisfont d'analogies superficielles et expliquent les différences phonétiques avec désinvolture : pour Budé, le français *foison* vient du grec  $\varphi\omicron\acute{\alpha}$  moyennant le changement de  $p$  en  $\varsigma$ , de  $o$  en  $oi$ , de  $a$  en  $o$  et l'addition d'un  $v$ .

Au dix-huitième siècle sévira une crise de celtomanie. D'autres ne jurent que par le germanique. Un isolé, l'Anversois Goropius Becanus fait venir du flamand toutes les langues du monde. Plus tard, d'autres encore, superficiellement frottés d'une science plus moderne sur les langues indo-européennes, verront partout, même dans nos noms de lieux, des survivances du sanscrit, en faisant fi des intermédiaires.

Les mots de nos dialectes peuvent aussi être auréolés d'une noble origine : le nom d'une danse du Borinage, l'alion, a été rattaché au grec  $\eta\lambda\iota\omicron\varsigma$  et au culte du soleil ; plus surprenant encore, en 1973, un lexique du parler de Comines fait venir du grec  $\chi\lambda\iota\acute{\alpha}\varsigma$  « millier » un mot dont l'étymologie est pourtant parfaitement claire

et aussi prosaïque qu'il est possible : le nom féminin *chuyée* au sens de « grande quantité ».

Ceci est l'occasion d'exprimer un des principes fondamentaux de la recherche étymologique : pour qu'une langue donne un mot à une autre langue, il faut que les usagers de l'une soient en contact avec les usagers de l'autre (ou que des intermédiaires puissent être déterminés), oralement, ou, pour les mots de nature savante, par écrit. Ni Borains ni Cominois n'ont côtoyé les Grecs, et ils n'ont pas été tous des familiers d'Homère ou de Platon.

Entretemps, l'idée que le français était issu du latin apporté par la conquête romaine (on ajoutait même parfois : du latin populaire) avait fait son chemin. Ce n'est pas le lieu de rendre hommage à tous ceux, connus ou méconnus, qui ont contribué à faire admettre ce que nous considérons aujourd'hui comme une évidence. Mais il ne suffisait pas que cela fût accepté pour que l'étymologie eût acquis le statut de science.

Gilles Ménage (1613-1692), maître de M<sup>me</sup> de Sévigné, mais aussi modèle supposé du Vadius des *Femmes savantes*, auteur du premier *Dictionnaire étymologique de la langue française* (1694), d'abord publié sous le titre *Origines de la langue française* (1650), n'a pu empêcher, malgré ses vastes connaissances et ses grandes qualités, que l'étymologie reste, aux yeux des gens sérieux, un jeu arbitraire. Il a même contribué à ce discrédit.

À l'actif de Ménage, outre la place qu'il reconnaît au latin, il faut mettre le fait qu'il annonce la méthode comparative et la méthode historique : d'une part, il prend en considération l'italien et l'espagnol, ainsi que les patois ; d'autre part, il donne au français ancien une place seulement limitée par la connaissance que l'on pouvait en avoir alors. Il n'était pas banal de tenir compte, dans une recherche ambitieuse, des patois et de l'ancien français, dédaignés par beaucoup comme des formes linguistiques inférieures. Grâce à tout cela, les étymologies exactes de Ménage sont plus nombreuses que les autres : on a calculé que soixante-douze pour cent de ses explications ont été conservées par Diez, qui au début du dix-neuvième siècle, inaugure la science étymologique moderne.

Malheureusement, il y a le déchet, moins gênant par l'aspect quantitatif que par la qualitatif. Quand il se trompe, Ménage pousse l'erreur jusqu'au ridicule.

Sa démarche part du sémantique : devant expliquer un mot français, Ménage cherche en latin un mot de sens analogue. Quand celui-ci a en même temps une forme proche, les choses se passent bien (il serait d'ailleurs injuste d'identifier les soixante-douze pour cent à des cas évidents). Mais les différences formelles les plus considérables n'arrêtent pas Ménage, convaincu que le synonyme latin est l'étymon, et il va le conduire de force jusqu'au mot français.

Je prends l'exemple célèbre de *haricot*. Ménage trouve comme synonyme latin *faba* (qui, par parenthèse, a donné *fève*). Il commence par justifier l'altération de l'initiale : « De *foris*, on a dit de mesme *horis*, dont nous avons fait *hors*. » Il est reconnu aujourd'hui encore que *hors* vient de *foris*, mais cette substitution de *h* à *f* est en français un phénomène exceptionnel qui demande une explication particulière (l'intermédiaire de *deforis*) et qu'il n'est pas permis d'appliquer à n'importe quel autre mot.

Cette erreur procède du souci louable de rendre raison d'une évolution qu'on envisage. Ménage sentait qu'il devait y avoir des règles, qu'il appelle « conversions de lettres » (il faudra longtemps avant qu'on donne aux sons le rôle primordial). Mais les lois phonétiques n'ont été découvertes qu'au dix-neuvième siècle, ainsi que ce principe fondamental : elles ne sont valables que dans un lieu et dans un temps déterminés et pour les sons placés dans des conditions identiques ; par exemple, *c* devant *a* devient *ch* à l'initiale du mot (*campus* → *champ*), mais non entre voyelles (*pacare* donne *payer*).

Le reste de l'histoire du mot, chez Ménage, pseudo-suffixations et contractions, est totalement fantaisiste : *fabarius*, *fabaricus*, *fabaricotus*, *habaricotus*, *haricotus*, *haricot*.

Je pourrais citer d'autres exemples et vous faire rire aux dépens de Ménage : comment *rat* vient de *mus* ou *hanneton d'asilus*. Notons que, dans les trois cas, il s'agit de réalités (le haricot, le rat et le hanneton) qui n'étaient pas connues des Latins. Autre exigence de l'étymologie moderne : il faut aussi connaître l'histoire des choses et des concepts.

Ménage est-il bien mort ? Il me semble qu'au vingtième siècle un article de Pierre Fouché et un livre de Pierre Guiraud, *mutatis mutandis*, n'en sont pas si éloignés.

Il s'agit, non d'obscurs amateurs, mais de linguistes connus (notamment par des œuvres ne prêtant pas le flan aux mêmes critiques).

L'article de Pierre Fouché<sup>2</sup> concerne une racine pré-indo-européenne, la racine \**kal-*. Ceci demande quelques éclaircissements préalables, *éclaircissements* étant à prendre avec une certaine réserve, puisque nous allons nous perdre dans la nuit des temps.

Les romanistes ont l'énorme avantage d'étudier l'histoire d'une langue dont l'ancêtre est une langue attestée, le latin. Cela donne aux lois phonétiques allant du connu au connu une solidité particulière. Mais le latin, en s'imposant, a évincé des langues indigènes : l'étrusque, le gaulois, le ligure, etc. Il est normal que les gens parlant ces langues aient gardé dans leur latin un certain nombre de leurs anciens mots : d'où la présence dans le français de mots d'origine gauloise. La conquête romaine n'a pas remplacé non plus tous les noms de lieux. La difficulté est que nous savons peu de choses sur le vocabulaire gaulois : il est vraisemblable, en particulier, que notre toponymie recèle des traces qui nous sont en quelque sorte devenues invisibles ou que nous expliquons en tenant compte seulement du peu de mots gaulois qui nous sont accessibles. C'est un domaine où il faut accueillir avec une certaine réserve les propositions des savants les plus qualifiés.

Les populations gauloises, qui appartenaient à la famille indo-européenne, ont été précédées d'autres populations, non indo-européennes, sur lesquelles les découvertes archéologiques apportent des renseignements divers. Mais celles-ci sont, naturellement, muettes sur la nature et les éléments de la ou des langues pré-gauloises. Il semble bien pourtant que des traces en subsistent jusqu'à nous, dans les noms des grands cours d'eau (*la Meuse*) et des montagnes (*les Alpes*). La comparaison entre des langues diverses, de part et d'autre de la Méditerranée, permet de rattacher à une langue pré-indo-européenne ou méditerranéenne un certain nombre de noms communs. Cette recherche demande des connaissances linguistiques exceptionnelles et une grandissime prudence.

La racine \**kal-* ferait partie de ce fonds très ancien. Pierre Fouché a dressé une liste d'environ sept cents noms de lieux qui conserveraient la trace de cette racine prolifique, étonnamment plastique puisque *r* peut remplacer *l*, *o* peut

---

<sup>2</sup> *À propos de \*kal-. Étude de toponomastique pré-indo-européenne*, dans *Anales del Instituto de lingüística*, 1943, p. 57-93.

remplacer *a*, la consonne initiale peut tomber, la voyelle peut tomber, l'initiale et la voyelle peuvent tomber ensemble, etc., tout cela sans qu'il soit jamais question ni de la prononciation actuelle des noms, ni de la phonétique locale, ni des formes anciennes, ni de la configuration des lieux (et si le *mont Chauve* était réellement chauve ?). Un seul exemple : *Goutelle*, nom d'un modeste ruisseau de nos Ardennes, se voit ainsi doté d'une ancienneté prodigieuse, vu sa faible utilité comme repère, alors que sa forme ne rappelle que vaguement la racine *\*kal-* et alors qu'il est tout à fait satisfaisant de rattacher à la famille de *goutte* le nom d'un ruisseau de faible débit.

Le *Dictionnaire des étymologies obscures* de Pierre Guiraud ne date que de 1982. Il n'a pas rencontré beaucoup de faveur chez les spécialistes. Seuls le *Grand Robert* et le *Robert historique* (le titre exact est *Dictionnaire historique de la langue française*) tout récent le citent avec une grande fréquence. C'est même à peu près le seul étymologiste qui y soit mentionné dans les articles. On pourrait donc croire que, dans le domaine, c'est ce que l'on a fait de plus original et de plus solide. Original, soit, mais solide ?

La situation dans laquelle s'est mis Pierre Guiraud est assez voisine de celle de Ménage. Celui-ci, auteur du premier dictionnaire étymologique du français, devait, et il n'avait pas de vrai prédécesseur, envisager l'étymologie de *tous* les mots français et, par conséquent, imaginer des évolutions là où le rapport avec le latin tel qu'il le connaissait ne sautait pas aux yeux. Guiraud retient les seize cents mots dont le dictionnaire de Bloch-Wartburg ne donne pas d'étymologie ou pas d'étymologie jugée acceptable, et il veut trouver à *tous* une origine, de préférence latine. C'est une position peu confortable, dont Ménage n'est pas sorti sans que sa réputation en ait souffert.

Comparons avec un des grands étymologistes : Walther von Wartburg, auteur du *Französisches etymologisches Wörterbuch* en plus de trente volumes, a commencé par réunir une énorme documentation, à la fois sur le passé de chaque mot français et sur le vocabulaire de tous les parlers de la France, de la Wallonie et de la Suisse romande. Cela lui a permis de voir d'une façon neuve beaucoup de problèmes et, par conséquent, d'émettre des propositions étymologiques

auxquelles n'avaient pas pensé ou n'auraient pas pu penser les érudits qui l'avaient précédé. Malgré cela, trois volumes sont prévus pour les mots d'origine inconnue.

Guiraud n'apporte aucune documentation neuve et ne cherche même pas à prendre connaissance de celle qui existe ; comment en aurait-il le temps ? Je sais bien qu'un rapprochement inédit peut surgir tout à coup de la confrontation de données connues. Mais imaginer de résoudre systématiquement, dans le silence de son bureau, sans aucune enquête ni vérification, *tous* les cas ayant jusqu'à présent laissé des générations de chercheurs dans la perplexité, quelle ambition !

Les réflexions que je viens de faire amènent à mettre en doute le sérieux d'une telle entreprise, mais il ne serait pas de bonne méthode de considérer pour cela qu'aucune des propositions n'est à retenir. Il conviendrait de les vérifier. C'est un énorme travail, que l'auteur ne facilite guère, car les *démonstrations* sont souvent rapides, sommaires, et l'hypothétique peu distinct de ce qui est réellement attesté.

Je me permettrai de développer un seul cas, pour lequel même les données réunies dans le *Robert historique* justifient des réserves expresses : le mot vulgaire *fion*.

Données objectives : le mot est attesté depuis le dix-huitième siècle, en 1744 au sens de « coup », en 1792 au sens de « dernière touche à un ouvrage », en 1793 au sens d'« injure » en Suisse ; enfin, en 1880 pour désigner le derrière ; dans ce dernier sens, il y a une variante *fignon*, je l'ajoute, de 1898. L'hypothèse de Guiraud est un latin vulgaire supposé *\*fionem*, résultant de l'altération, par dissimilation des deux *n*, d'un latin vulgaire supposé *\*finionem*, dérivé de *finis*, qui a donné régulièrement *fin*.

Ce n'est qu'avec la plus grande précaution que l'on doit faire remonter au fonds primitif latin un mot attesté pour la première fois en français plus d'un millénaire après la disparition du latin parlé, a fortiori si ce mot latin n'est pas attesté. Or, non seulement ce mot latin est supposé, mais il serait altéré d'une forme également supposée. A-t-on d'autres exemples d'une telle dissimilation ? La dérivation supposée *\*finio* trouve-t-elle des confirmations dans d'autres mots ? Questions sans réponses. Quel rapport y a-t-il entre le sens premier du français, « coup », et les sens du latin *finis* et de sa famille ? Pas de réponse. Toutefois le *Robert historique* a cette précision : l'hypothèse expliquerait le sens « derrière, postérieur », c'est-à-dire ce qui est à l'extrémité du corps (?) ; mais ce sens est le

dernier attesté en français : cent quarante ans après le premier. Dernier paradoxe, la variante *fignon*, apparue plus tard encore, dériverait de la forme latine *\*finio*. Autrement dit, la forme et le sens primitifs ne seraient représentés que par la dernière forme et par le dernier sens attestés en français.

La discussion est un peu technique, veuillez me le pardonner. Elle me semblait nécessaire pour étayer le rapprochement entre Guiraud et Ménage. Sans doute n'observe-t-on pas chez notre contemporain l'ignorance des lois phonétiques que l'on a constatée chez le second. Ce n'est ni un mérite pour l'un ni un démérite pour l'autre, puisqu'on les a découvertes au dix-neuvième siècle. Pour le reste, cette étymologie pourrait être signée Ménage.

Mes commentaires ont surtout porté sur la préhistoire de l'étymologie, sur la période où les recherches étaient viciées par trois défauts : un parti pris sur le but même de l'étymologie ; l'indifférence ou l'ignorance en ce qui regarde les faits historiques ; l'ignorance des lois de l'évolution phonétique. J'ai aussi montré que ces défauts n'avaient pas disparu, même au vingtième siècle. Il a donc été question surtout de l'étymologie telle qu'elle ne devrait pas être. Il me restait à la voir telle qu'elle doit être, à donner une vue plus positive de l'étymologie, de ses méthodes et de ses résultats. Si je ne vous ai pas trop ennuyés, ce pourrait être l'objet d'une autre communication.

Copyright © 1993 Académie royale de langue et de littérature françaises de Belgique. Tous droits réservés.

**Pour citer cette communication :**

André Goosse, *Réflexion sur l'étymologie* [en ligne], Bruxelles, Académie royale de langue et de littérature françaises de Belgique, 1993. Disponible sur : < [www.arlfb.be](http://www.arlfb.be) >