

Quand la linguistique française ne saurait que se faire romane : du neuf dans le traitement étymologique du lexique héréditaire

Éva Buchi

ATILF (CNRS & Nancy-Université)
eva.buchi@atilf.fr

Jean-Paul Chauveau

ATILF (CNRS & Nancy-Université)
jean-paul.chauveau@atilf.fr

Xavier Gouvert

ATILF (CNRS & Nancy-Université)
xavier.gouvert@atilf.fr

Yan Greub

ATILF (CNRS & Nancy-Université)
yan.greub@atilf.fr

1 Introduction

Si l'historiographie linguistique doit citer un changement majeur dans la conception et la pratique de l'étymologie romane au 20^e siècle, c'est sans doute le remplacement de « l'étymologie-origine » par « l'étymologie-histoire ». Ce changement n'est pas neuf, puisqu'il se fait sous l'impulsion de Gilliéron, dès le début du siècle, puis est formalisé définitivement, après von Wartburg, par Baldinger (1959). L'étymologie moderne ne recherche donc plus exclusivement l'origine d'un lexème, mais s'efforce d'écrire l'histoire de sa présence à l'intérieur du lexique et de décrire les changements qui l'ont affecté. Cette conception n'est absolument pas contestée, ou du moins pas publiquement, et il n'y pas de raison de penser qu'elle n'a pas été complètement admise par l'ensemble de la discipline « linguistique historique romane » ou par ses sous-ensembles « linguistique historique française » etc.

Cependant, nous savons tous que l'étymologie-origine se pratique couramment, par raccourci de l'étymologie : il suffit d'ouvrir le premier dictionnaire venu pour s'en apercevoir. Cette existence, qui ne cherche pas à se justifier théoriquement, a plusieurs causes et plusieurs formes. Dans le cas du Petit Robert (Rey-Debove & Rey 2009), par exemple, il s'agit de gains de place et de temps, dans un dictionnaire qui ne prétend pas s'intéresser spécifiquement à l'étymologie et qui la réduit donc radicalement à une généalogie (date de naissance + géniteur). Par ailleurs, l'usage de l'étymologie-origine, il faut le préciser, n'est pas seulement le fait d'amateurs restés à l'écart du progrès de la science et des bibliothèques qui le reflètent : le dictionnaire de Baumgartner et Ménard (1996), par exemple, en témoigne.

En fait, on est en face d'un paradoxe : le champ de l'étymologie-origine a été envahi par des formes dégradées, après avoir été abandonné par la science officielle. Celle-ci n'a pas la place de s'y resituer, et pas d'intérêt à le faire non plus.

Et, deuxième problème, corollaire du premier, la recherche scientifique, qui traite l'origine des unités lexicales dans ses descriptions étymologiques, a exclu ce point de sa réflexion théorique ; elle traite donc

le concept d'origine comme un donné, plutôt que comme tous les concepts scientifiques, qui sont susceptibles de révision, et qui dépendent dans leur définition des modèles généraux dans lesquels ils s'insèrent. Le concept d'origine en est donc resté à un degré d'élaboration insuffisant.

Ce n'est que par une extension indue que la recherche étymologique a éliminé, comme on vient de le dire, la question de l'origine de sa réflexion théorique : parce qu'elle a justement éliminé la technique de l'étymologie-origine de ses pratiques, elle a cru pouvoir abandonner aussi la réflexion portant sur l'origine comme partie de la technique de l'étymologie-histoire. Comme la question théorique de l'étymon a été éliminée, l'étymon est la partie la plus mal traitée dans tout article étymologique. Nous nous proposons ici de réintroduire critiquement le concept d'étymon dans le champ de l'étymologie.

Le concept d'étymon a des valeurs essentiellement distinctes selon les cas : dans le cas d'un emprunt, l'étymon se situe dans un autre système linguistique (la langue prêteuse), au moment qui précède immédiatement l'entrée dans la langue emprunteuse, et l'unité linguistique-étymon est distincte, elle aussi, du lexème à étymologiser. Dans le cas d'une formation interne, l'étymon et le lexème à étymologiser appartiennent au même système linguistique, mais ils sont différents. En revanche, la situation d'un lexème héréditaire est tout autre, puisqu'il n'y a ni création d'une unité lexicale, ni transfert d'une langue à une autre, et qu'on ne peut donc pas situer l'origine par l'entrée du lexème dans la langue ; la différence entre les étymons du lexique héréditaire et ceux des deux autres classes étymologiques passe entre la continuité et la discontinuité.

L'auditeur pense déjà qu'il va être très difficile de faire intervenir une quelconque notion d'origine dans un développement continu, et les auteurs partagent ses craintes. L'étymologie-origine, lorsqu'elle traite le matériel héréditaire, a choisi de répondre à cette difficulté par un coup de force : l'invention du latin, langue paradoxale, à la fois ancêtre des langues romanes et essentiellement séparée d'elles, c'est-à-dire pourvue des deux qualités permettant aux langues romanes d'aller y chercher leurs étymons. Cette langue est une aporie : si elle est le point de départ des langues romanes, il n'existe aucun événement qui permette d'établir une démarcation suffisante pour trouver une origine, et si elle n'est pas le point de départ des langues romanes, on n'a pas à aller y chercher l'origine du lexique héréditaire. Mais cette aporie, qu'on ne peut percevoir que si l'on reconnaît le caractère continu de l'évolution linguistique, est par conséquent invisible à l'étymologie-origine, qui procède par saut : elle saute du français, langue bien identifiée dès avant toute recherche étymologique, au latin, langue bien identifiée dès avant toute recherche étymologique. Comme sa méthode ne tient aucun compte de la continuité linguistique, il n'y a aucun problème à sauter du français au latin comme on saute à toute langue à laquelle le français a emprunté : rien ne vient remettre en question la notion d'étymon latin d'un lexème héréditaire français.

L'étymologie-histoire a des ennuis plus sérieux : elle croit éminemment à la continuité de l'évolution linguistique, puisque sa méthode repose sur cette notion, mais elle a hérité son concept d'étymon de l'étymologie-origine, et le place tant bien que mal au bout de l'histoire de l'unité lexicale. Pour maintenir l'illusion, et ne pas avoir à trouver une autre borne au segment d'histoire qu'elle décrit, l'étymologie-histoire a besoin d'une interruption de la continuité : elle la crée en refusant d'utiliser une méthode qui lui permettrait de remonter directement au-delà des premières attestations dans les langues romanes, la reconstruction linguistique. On conçoit dès lors l'importance du trou noir qui précède les premières attestations romanes : une solution de continuité entre le français (ou toute autre langue romane) archaïque et le latin est nécessaire pour conserver à ce dernier son statut d'origine des langues romanes, et pour avoir un endroit où placer, ou chercher, les étymons¹. De là les fausses questions sur le passage du latin aux langues romanes, et sa date. Le truc est un peu gros, et pour le faire paraître moins gros on s'efforcera de rapprocher le latin des premières attestations romanes, en faisant figurer dans son développement historique un latin tardif.

Mais à un esprit non prévenu, cette solution apparaît comme un simple expédient, malgré son succès séculaire ; la question de l'étymon reste ouverte. Nous tenterons d'y répondre à travers l'analyse de l'existant (2), avant de proposer un traitement renouvelé de l'étymologie du lexique héréditaire (3). La présente communication dépend des discussions méthodologiques qui ont préparé et qui entourent la rédaction du DÉRom (*Dictionnaire Étymologique Roman*, cf. Buchi & Schweickard 2008 ; 2009 ;

Florescu 2009) ; elles ont été menées avec J.-P. Chambon², ou sous son impulsion (*cf.* spécialement Chambon 2007 et à paraître).

2 Analyse des pratiques

Notre analyse de la pratique actuelle de l'étymologisation du lexique héréditaire français s'appuiera sur le témoignage de trois lexèmes : *charme* (nom d'arbre), *faire* et *foin*. Nous en examinerons le traitement par les principaux dictionnaires étymologiques consacrés au seul français : Gamillscheg (1928¹–1969²), Bloch & Wartburg (1932¹–1968⁵), Dauzat, Dubois & Mitterand (1938¹–2001), Picoche (1971¹–2002), TLF (1971–1994), le Robert historique (1992¹–1998²), enfin Baumgartner & Ménard (1996). Pour ce qui est du FEW, il traite, malgré son titre trompeur, des quatre langues *galloromanes* (*cf.* son sous-titre) : il n'y a donc pas lieu de le confondre avec nos sept dictionnaires-témoins, qui se limitent au seul français.

L'objet de notre analyse est constitué par les indications proposées par les différents dictionnaires pour énoncer non pas l'étymologie (au sens d'« étymologie-histoire ») de *charme*, *faire* et *foin*, mais seulement ce qui concerne leur étymon, à l'exclusion donc de ce que certains d'entre eux peuvent par ailleurs contenir en matière d'indications historiques (datations, évolution sémantique et morphosyntaxique etc.). Notre analyse se subdivise en cinq parties : étiquette glottonymique (2.1), signifiant (2.2), signifié (2.3), combinatoire restreinte (2.4) et insertion dans le cadre roman (2.5).

2.1 Étiquette glottonymique

En étymologie française et romane, il existe un consensus partagé par l'ensemble de la profession pour écarter l'*etimologia remota* au profit de l'*etimologia proxima* (*cf.* par exemple Pfister & Lupis 2001 : 37)³. Ce parti pris exclut le proto-indo-européen comme langue d'appartenance des étymons du lexique héréditaire du français et des autres idiomes romans : on s'arrête à l'embranchement roman de l'arbre phylogénétique indo-européen, qui a une consistance linguistique (en tant qu'ancêtre commun des langues romanes, par opposition, par exemple, à l'ancêtre commun des langues germaniques), mais aussi sociolinguistique (par sa diffusion à l'intérieur de l'Empire romain).

La phylogenèse romane délimite toutefois aussi des embranchements internes, et on peut se poser la question de savoir si l'*etimologia proxima* ne commanderait pas d'arrêter la reconstruction du lexique héréditaire français à un niveau plus bas, par exemple à l'étape du proto-oïlique (accessible à travers la comparaison avec le wallon, le picard, le champenois etc.) ou à l'étape du « protoroman continental occidental » (accessible à travers la comparaison avec l'occitan, l'italien, l'espagnol etc.). Or la pratique étymologique montre – et nous espérons l'illustrer par les exemples développés par la suite – que le cadre roman est le plus rentable. Dès lors, la langue susceptible de fournir des étymons au lexique héréditaire français sera l'ancêtre commun des parlers romans. Il reste toutefois à savoir comment dénommer cette langue. Le traitement du verbe *faire* par les sept dictionnaires de notre corpus fournira le point de départ de notre argumentation.

Si l'ensemble des auteurs interrogés rattachent l'étymon du verbe français *faire* au latin, on constate trois réalisations légèrement divergentes dans le détail : tandis que Gamillscheg (1928¹–1969²), Bloch & Wartburg (1932¹–1968⁵), Dauzat, Dubois & Mitterand (1938¹–2001), le Robert historique et Baumgartner & Ménard (1996) attribuent l'étymon à un latin non spécifié, J. Picoche (1971¹–2002) le rattache au latin vulgaire, et E. Papin *in* TLF (1980), au latin classique. Le choix retenu majoritairement présente l'avantage de ne pas prodiguer une information notoirement erronée : les étymons des lexèmes héréditaires des langues romanes, et donc du français, appartiennent en effet au latin global (*cf.* de Dardel 2009)⁴ : dès lors, *latin* n'est pas faux. La formule n'est toutefois pas particulièrement heureuse, car elle manque de précision, *latin* renvoyant à des états de langue successifs depuis le latin archaïque, à travers le latin classique, le latin tardif, le latin médiéval et le latin de la Renaissance jusqu'au latin scientifique de l'époque moderne et contemporaine.

Il ne sera pas nécessaire de s'étendre longuement sur l'étiquette retenue par le TLF (« lat. class. »), dont le caractère fautif – dû à un raccourci d'écriture – saute aux yeux : en aucun cas un lexème héréditaire d'une langue romane ne saurait remonter au latin classique, une langue où « cheval » se disait *equus* et où la marque de l'accusatif singulier était *-m* : à l'exception notable de W. Mańczak (ainsi Mańczak 1991), qui soutient envers et contre tous que les langues romanes proviennent en droite ligne du latin de Cicéron, l'ensemble de la communauté scientifique part du principe qu'elles remontent à la variété orale du latin de l'Antiquité, ou plus précisément, pour utiliser la terminologie de Koch & Oesterreicher (cf. en dernier lieu Koch & Oesterreicher 2008), au latin de l'immédiat communicatif. Les romanistes ont coutume de faire référence à cette variété du latin global à travers le terme *latin vulgaire*⁵.

Si un tel glottonyme paraît en l'occurrence plus pertinent que *latin* ou *latin classique*, il ne faut pas se cacher les sérieux problèmes d'ordre conceptuel que pose l'utilisation de ce terme technique et, en fin de compte, son inadéquation à l'objet qui nous occupe. En effet, bien que les auteurs de manuels commencent en général par attribuer à *latin vulgaire* un sens purement diastratique et diaphasique, disons de « latin parlé par le peuple », il est patent qu'ils procèdent à une substitution de concept au moment d'en produire la description interne.

On trouve ainsi, jusque dans les ouvrages les plus récents consacrés à la diachronie romane et française⁶, une description du phonétisme « latin vulgaire » caractérisant ce dernier par un système à sept voyelles (* / a e i o u /), système issu du « bouleversement quantitatif » (all. *Quantitätenkollaps*) et de la « conjonction vocalique » bien connus des francisants. On y apprend par exemple que les prototypes en « latin vulgaire » de fr. *toile* et *poil*, ou de fr. *fleur* et *gueule*, possédaient la même voyelle tonique (respectivement */e/ et */o/) ; hypothèse que la simple observation des séries de correspondances suffit pourtant à infirmer :

roum. *tearǎ* = it. *tela* = sard. *tela* = fr. *toile* = occit. cat. esp. *tela* ;
roum. *pǎl* = it. *pelo* = sard. *pilu* = fr. *poil* = occit. cat. *pel* = esp. *pelo* ;

roum. *floare* = it. *fiore* = sard. *flore* = fr. *fleur* = occit. cat. esp. *flor* ;
roum. *gurǎ* = it. *gola* = sard. *bula* = fr. *gueule* = occit. cat. esp. *gola*.

Il y a en fait quatre séries distinctes, non deux, et il convient donc de poser neuf voyelles au total, non sept, pour rendre compte des formes romanes. De toute évidence, le « latin vulgaire » à sept voyelles – celui avec lequel opèrent la plupart des étymologistes francisants – porte bien mal son nom. Tout au plus s'agit-il d'un « latin vulgaire continental occidental », quelque chose comme l'ancêtre de l'italien, du français, de l'occitan, de l'espagnol etc. – et non du sarde, du sicilien ou du roumain. Ne fût-ce qu'en raison de cette ambivalence, *latin vulgaire* ne convient donc pas comme glottonyme de l'ancêtre commun de tous les parlers romans.

Si donc l'on veut toucher à l'ancêtre commun des langues romanes dans leur ensemble, il faut remonter plus haut, au moment d'avant la séparation en plusieurs sous-groupes, quels qu'ils soient. Le glottonyme le plus approprié pour désigner cette langue est *protoroman* (cf. de Dardel 1996). Le protoroman est une protolangue dans le sens (2) délimité par Campbell (2004 : 125) :

Proto-language : (1) the once spoken ancestral language from which daughter languages descend ; (2) the language reconstructed by the comparative method which represents the ancestral language from which the compared languages descend. (To the extent that the reconstruction by the comparative method is accurate and complete, (1) and (2) should coincide.)

2.2 Signifiant

L'un des arguments couramment avancés pour nier la nécessité, voire la possibilité même d'une reconstruction du protoroman – et, par conséquent, pour placer le latin à l'origine génétique du français – est le suivant : la méthode comparative, employée seule, n'est pas en mesure de restituer la réalité de l'ancêtre commun d'une famille de langues dans toutes ses dimensions. Ainsi formulée, la proposition est

inattaquable : elle n'a d'ailleurs jamais été contestée par les théoriciens ni par les praticiens de la grammaire comparée-reconstruction de quelque famille linguistique que ce soit (cf. Fox 1995 : 11, 137-142 ; Harrison 2003 ; Campbell 2004 : 166-167). Nul ne prétend reconstruire dans son intégralité le lexique du latin global – non plus que celui du proto-indo-européen ou du protobantou – tel qu'il a existé à toutes les étapes de l'histoire et dans toutes les variétés géographiques et sociolinguistiques de cette langue. Mais du fait, universellement admis, que l'on ne peut reconstruire l'intégralité de la langue ancestrale, on ne devrait pas tirer la conclusion que le latin écrit de l'Antiquité serait en mesure de combler les lacunes de la reconstruction ou plutôt, par un étonnant renversement méthodologique, que le rôle de la reconstruction serait justement de suppléer le témoignage du latin écrit lorsque celui-ci fait défaut (cf. les formes à astérisque des dictionnaires étymologiques traditionnels). En un mot, la reconstruction de la protolange serait une perte de temps puisqu'« on la possède déjà sous une forme attestée » (comme si les textes latins antiques donnaient eux-mêmes une représentation exhaustive de la langue latine !). Dans le même ordre d'idées, on reproche à la méthode reconstructive son incapacité à restituer la profondeur temporelle de la langue ancestrale et à saisir ses états synchroniques successifs. Un tel reproche traduit une criante méconnaissance des techniques et des résultats du comparatisme, notamment en domaine roman.

Un coup d'œil sur l'exemple de fr. *foin*⁷ nous fournira l'illustration de ce que, dans bien des cas, la reconstruction protoromane permet non seulement d'atteindre la réalité de la protolange dans l'état immédiatement antérieur à sa fragmentation, mais également de saisir son évolution diachronique et, par contrecoup, d'éclairer d'un nouveau jour certains problèmes touchant à l'histoire du latin classique lui-même. En consultant l'article *foin* des dictionnaires étymologiques du français, on remarque que le signifiant de l'étymon proposé revêt des formes différentes selon les auteurs. Pour Gamillscheg (1928¹–1969²), Bloch & Wartburg (1932¹–1968⁵), Dauzat, Dubois & Mitterand (1938¹–2001) et Picoche (1971¹–2002), la forme posée est *fēnum* ; pour le Robert historique, c'est simplement *fenum* ; selon le TLF (1980), il s'agit de *faenum*, *fēnum* ; Baumgartner & Ménard (1996) donnent *foenum*, *fenum*. On voit que quatre notations différentes sont proposées, quelquefois simultanément, pour le même étymon, *fenum*, *fēnum*, *foenum* et *faenum*, sans qu'à aucun moment ne soit débattue la question de leur pertinence en phonétique historique. Or, le problème mériterait qu'on s'y arrête, puisque le signifiant et l'étymologie du lexème latin désignant le foin sont un sujet de débat chez les latinistes eux-mêmes (cf. Ernout & Meillet 1959 et de Vaan 2008 s.v. *fēnum*).

Du côté des francisants, on remarquera, pour commencer, que la graphie <foenum>/<faenum>, mentionnée par le TLF et par Baumgartner & Ménard (1996), est fautive et résulte probablement d'une mélecture : la forme à diphtongue attestée en latin écrit de l'Antiquité est bien <faenum> et non <foenum> (qui ne se trouve que dans des manuscrits médiévaux). Quant à la question de savoir si la forme latine primitive est /'faenum/ ou /'fe:num/, elle n'a, à ce jour, pas été tranchée par les latinistes ; elle n'est pas négligeable, en ce qu'elle intéresse également, par ricochet, les comparatistes-reconstructeurs de l'indo-européen. Elle n'est pourtant guère soulevée par les dictionnaires étymologiques français, à l'exception notable de Bloch & Wartburg (1932¹–1968⁵), qui adoptent assez systématiquement une perspective romane. Or, dans ce débat, l'apport de la reconstruction du protoroman s'avère décisif. La comparaison romane (cf. Reinhardt in DÉRom s.v. */ϕen-u/ ~ */ϕɛn-u/) permet en effet de poser à l'origine de fr. *foin* et de ses cognats – dacorum. *fān*, dalm. [fɪn], frioul. [fɛn], it. *fieno*, sard. *fenu*, occit. *fen*, cat. *fe*, esp. *heno*, gal. *feo* etc. – non pas une forme « *fēn-u-m » que la reconstruction n'autorise pas à postuler, mais deux bases étymologiques parallèles : (1) */ϕen-u/, dont sont issues toutes les formes romanes, à l'exception de l'italien, du frioulan et du sarde, dont les corrélats se ramènent à (2) */ϕɛn-u/. Ce résultat – la restitution d'une double base étymologique – n'est que la première étape de la reconstruction, ce que l'on peut appeler la reconstruction mécanique. Il incombe au comparatiste de rendre compte de cet état de fait par une deuxième procédure, la reconstruction interne : il s'agit d'élaborer un scénario plausible pour rendre compte d'une anomalie observée dans l'état de la protolange reconstruite et de ramener l'irrégularité apparente à une régularité implicite. Dans le cas */ϕen-u/ ~ */ϕɛn-u/, l'hypothèse proposée, la plus vraisemblable au regard des tendances générales de

l'évolution phonétique, est que la concurrence */e/ ~ */ɛ/ en protoroman est le résultat de la réduction d'une diphtongue plus ancienne (pré-protoromane). On reconstruit cette dernière comme */ai/, qui aura abouti en général à */ɛ/, plus rarement à */e/ (cf. Meyer-Lübke 1890 : 1 : 240-241 ; 524-525 ; Lausberg 1963 : 1 : 190-191 ; Rohlf's 1949 : 1 : 115-116 ; 122-123 ; 183-184 ; Lloyd 1987 : 105-106). L'étape supplémentaire franchie par la reconstruction interne – laquelle, redisons-le, ne fait pas intervenir les données du latin écrit – aboutit à un résultat remarquable, puisqu'elle permet d'orienter le changement phonétique dans le sens */ai/ > */e/ ~ */ɛ/ et donc d'affirmer que le stade de la diphtongue est antérieur à celui de la voyelle simple. Il est dès lors peu risqué de supposer, par extrapolation, que le latin écrit reflète la même évolution phonétique et que la forme <faenum> ne représente pas un hypercorrectisme, mais bien un archaïsme graphique. On a donc ici un cas remarquable où la grammaire comparée-reconstruction romane s'avère de beaucoup plus efficace à restituer la profondeur temporelle que le témoignage des textes latins, parce qu'elle ne passe pas par l'écrit, même interprété phoniquement, mais va de l'oral à l'oral. Ce n'est pas le latin qui vient au secours du romaniste, mais, dans ce cas au moins, le protoroman qui offre son aide au latiniste.

Un exemple plus complexe, mais tout aussi éclairant, est fourni par fr. *faire*. On sait que la base latine *facere* ne peut rendre compte régulièrement du paradigme verbal français. À latclass. <facis>, <facit>, <facer-> devraient répondre respectivement *<faiz> (ind. prés. 2), *<faist> (ind. prés. 3), *<fairai>, *<fairas> (ind. fut.) etc., toutes formes inexistantes en ancien français. Pour tenter de résoudre cette irrégularité, une explication monogénétique *ad hoc* a été défendue notamment par Meyer-Lübke (1894 : 435) et reprise, non sans distorsions, par Dauzat (1938¹), puis par Picoche (1971¹-2002) : il faudrait poser un « latin vulgaire » « **fagere* », « altéré à l'inf. [...] d'après *agere* » (Dauzat, Dubois & Mitterrand 1938¹-2001 ; Picoche 1971¹-2002 : « lat. vulg. **fagère*, du class. *facere* »). Or, si */fagere/ peut effectivement rendre compte de /faiʁə/ (mais */fakere/ le peut tout autant), cette hypothèse n'explique pas davantage les formes /fais/ et /fait/ du présent, ni le radical /fər-/ du futur et du conditionnel ; pis, elle rend inexplicables toutes les formes qui, elles, réclament une base /fak-/ (<faisons>, <faisais> etc.) ! Dauzat (mais non Picoche) se voit donc contraint de se replier sur une troisième hypothèse et de postuler que « le futur et le conditionnel reposent sur une forme abrégée **fare* (**fare-habeo* > *ferai* ; cf. it. *fare*) ».

Sans même parler des inconséquences et de l'arbitraire de la formulation étymologique, l'explication du phonétisme des représentants de */ʰak-e-/ par l'action analogique d'un lexème non transmis aux parlers romans (*agere*) est gratuite. La méthode comparative aboutit en revanche à une explication cohérente (cf. Buchi in DÉRom s.v. */ʰak-e-/). Les matériaux romans, d'une part dacorum. *face*, aémil. *facere*, sard. *fakere*, occit. *faire*, cat. *fer*, esp. *hacer*, port. *fazer*, d'autre part dalm. *fur*, istriot. *fa*, it. *fare*, frioul. *fã*, conduisent sans doute possible à poser deux étymons distincts : une base primitive */ʰak-e-re/ et une base syncopée */ʰ-a-re/. La seconde est issue de la première en raison d'un traitement phonétique particulier dû à la haute fréquence du verbe, notamment dans la position proclitique du type le plus ancien du futur roman, */ʰ-a-re-'aβ-e-/ (cf. Rydberg 1893 : 55, 66-67). De prime abord, l'aire occupée par ce type de futur analytique ne se superpose pas complètement avec la zone occupée par la forme syncopée : l'homotopie n'est complète que si l'on accepte la thèse de Tekavčić (1976/1977), lequel plaide, contre la *communis opinio* (notamment Bartoli 2000 [1906] : 451-452 § 482 et Doria 1989 : 523), mais avec de bons arguments, pour une origine périphrastique du futur dalmate. De façon bien inattendue, l'analyse stratigraphique des verbes issus de protorom. */ʰak-e-/ proposée ici apporte ainsi un argument supplémentaire à la thèse de Tekavčić.

Cet exemple est intéressant dans la double mesure où, d'un côté, la morphologie flexionnelle historique apporte ici sa pierre à l'étymologie lexicale et où, en retour, c'est l'étymologie lexicale qui éclaire la formation du système morphologique. On se trouve même dans un cas de figure où, en quelque sorte, c'est la phonétique historique du dalmate qui vient au secours de la lexicologie du français... Ce mouvement dialectique donne une éclatante illustration aux propos de Durkin (2009 : 32) :

Finally, words form part of a system, the lexis of a language, with numerous links to its grammar also. Any change in our understanding of one part of that system may have echoes or repercussions in another, possibly quite distantly removed, part of the same large system, and we must always be alert to such implications in our own or others' work. Sometimes, one changed etymology can open the way to a whole set of new solutions to old problems.

2.3 Signifié

La part sémantique de l'étymologie vue par les dictionnaires examinés n'est que différentielle (à l'exception du TLF), et leur silence à ce sujet signifie donc simplement l'identité sémantique entre l'étymon et son résultat français, l'immobilité du sens au cours du déroulement historique. Le TLF⁸, le Robert historique⁹ et Baumgartner & Ménard (1996)¹⁰ sont les seuls qui introduisent, au moins implicitement, la question sémantique, par la mention du sens du lexème latin ; celle-ci insiste sur la différence de sens entre lat. *facere* et fr. *faire*, mais ne prend que la forme d'une description, ou d'une histoire, du sens à l'intérieur du latin. Il n'y a donc aucune étude de la façon dont on passe du sémantisme latin (supposé originaire) au sémantisme français, ni du moment auquel ce changement prend place. Or la comparaison romane montre que le sens français est aussi celui qu'on peut reconstruire en protoroman ; celui-ci apparaît donc, de ce point de vue, comme tout à fait séparé du latin classique tel qu'il est attesté dès avant la séparation du sarde du tronc commun : sens « produire un effet à travers un travail manuel ou intellectuel » généralisé et champ sémantique organisé différemment (absence du parasynonyme *agere*).

La comparaison romane donne ainsi une datation (au moins relative) des changements sémantiques ayant affecté *faire* et ses antécédents ; elle est la seule méthode qui le fasse.

2.4 Combinatoire restreinte

L'exemplification de la composante morphosyntaxique de l'étymon se fera à partir du traitement du nom d'arbre *charme*. Le point commun le plus remarquable des étymologisations des sept dictionnaires considérés, c'est ce qui leur manque, c'est l'absence de toute caractérisation morphologique de l'étymon, qu'ils présentent unanimement comme « lat. *carpīnus* ». L'analyse d'un signe linguistique, et donc d'un étymon, doit prendre en compte, en plus de son signifiant et de son signifié, sa valeur morphosyntaxique ou bien, pour utiliser la terminologie de la théorie sens-texte, sa combinatoire restreinte (Polguère 2008 : 40¹¹) ; il s'agit de traits inhérents, stockés, d'une unité lexicale, comme le signifiant et le signifié.

Le lexème français étant un substantif, il est implicite que l'étymon en est un, lui aussi. Mais la catégorie grammaticale ne suffit pas pour caractériser précisément un lexème du latin, langue qui se signale par une morphologie variée et au rôle important. Tel qu'est mentionné cet étymon latin, on peut hésiter sur le type flexionnel dont il relève (2^e ou 4^e déclinaisons ?) et sur son genre (masculin ou féminin ?). Le lecteur non prévenu pourrait penser que, puisque le nom français est de genre masculin, qu'il doit en aller de même pour le nom latin, à défaut d'indication contraire. Ce qui est faux, puisque le latin *carpīnus* est un nom féminin de la seconde déclinaison. Bien sûr, les auteurs de ces dictionnaires ne l'ignoraient pas. S'ils n'ont pas jugé utile de le signaler, c'est que (1) la morphologie lexicale n'entre pas dans leur conception de l'étymologie et que (2) leurs étymons sont de simples étiquettes. Le latin qui leur fournit leurs étymons n'est pas un système linguistique, mais un réservoir d'unités dégagées de tout usage langagier, puisque dépourvues des traits morphosyntaxiques inhérents, tels que le genre du substantif et son type flexionnel.

Dans la reconstruction menée dans le cadre ici proposé (cf. Medori *in* DÉRom s.v. */karpn-u/), les formes qui relèvent d'une continuité ne peuvent se ramener, par les règles de l'évolution phonique, à un seul point de départ. Il y en a obligatoirement deux, distingués par leurs traits phonologiques et morphologiques : */karpn-u/ s.m. et */karpn-a/ s.f. La première vertu de la reconstruction, c'est d'obliger à faire une double constatation : (1) la dénomination protoromane du charme n'est pas unique ; (2) les deux formes reconstruites ne peuvent pas se superposer à celles du latin classique.

Si l'on replace le cas au sein des dénominations romanes des arbres, où l'on a d'autres exemples d'une variation de genre, il devient possible d'interpréter les deux prototypes obtenus comme deux réfections de sens contraire d'un même point de départ, visant à accorder le genre avec la phonie de la finale selon le modèle suivi par la très grande majorité des substantifs correspondant à ces finales : le masculin pour une finale */-u/ et le féminin pour une finale */-a/. Seul un féminin */karpin-u/ peut devenir */karpin-u/ s.m., par mise en accord du genre avec le suffixe flexionnel, et */karpin-a/ s.f., par mise en accord du suffixe flexionnel avec le genre. Or les textes latins nous révèlent un nom d'arbre *carpīnus*, -nī s.f. qui est le répondant exact, formel et sémantique, du lexème */karpin-u/ s.f. que la reconstruction a établi comme le point de départ des formes romanes. Mais ce roman commun des premiers siècles de notre ère est déjà travaillé par des tendances qui vont affecter profondément sa morphologie flexionnelle. Le parallélisme entre la répartition du genre et les marques flexionnelles est déjà bien entamé, puisque la masculinisation de */karpin-u/ s'est presque généralisée. Et c'est ce protorom. */karpin-u/ s.m., ce lexème identifié phonologiquement et morphologiquement et situé parmi les variantes concurrentes du latin parlé au sein de l'Empire romain, que continue sans interruption ni discontinuité le français *charme*.

Le maintien, uniquement en domaine francoprovençal, de */karpin-a/ s.f., avec conservation du genre féminin et marquage formel de celui-ci, pour un nom d'arbre n'est pas une curiosité anecdotique. En effet, le francoprovençal est bien connu comme ayant conservé les formes du neutre singulier pour quelques noms de fruits (cf. Gardette 1983 : 698-706). Le maintien du féminin pour le nom d'arbre et celui du neutre pour le nom de fruit, selon un modèle indo-européen fidèlement conservé par le latin, voilà qui signale une zone conservatrice qui s'est déterminée très tôt à l'intérieur de la Romania. Semblablement il faut replacer dans cette perspective les quelques hésitations marginales pour les noms d'arbres en français dont témoigne, par exemple, le maintien des deux genres pour l'ancien français *coudre* « noisetier », *sauz* « saule » ou le français régional *ver(g)ne* « aulne ».

L'étymologie n'est pas seulement une quête d'origines, mais aussi la recherche de la documentation utile pour reconstituer précisément des évolutions globales. Pour cela, il est nécessaire de ne pas limiter l'étymologie à des relations interlexicales, mais de reconstituer des parcours où soient délimitées : (1) les évolutions qui ne rompent pas la continuité, affectant le signifiant (comme la palatalisation de /k/ devant /a/, la syncope de la voyelle posttonique etc.) et le signifié (comme la métonymie régulière en français : « arbre » > « bois de l'arbre », *un charme/du charme*) ; (2) les discontinuités évolutives, les ruptures affectant (a) le signifiant, comme le changement de genre qui démotive le genre dans un domaine lexical où il était encore motivé, et (b) les relations sémantiques : par l'abandon de lat. *carpineus* adj. « de bois de charme », la relation « arbre »/« bois de l'arbre » sort du lexique, ne passe plus par la morphologie constructionnelle, mais recourt à une solution syntaxique : *de/en charme*. Ce sont ces discontinuités qui déterminent la rupture et justifient le choix de */karpin-u/ s.f. comme ancêtre commun des prototypes romans, comme étymon.

2.5 Insertion dans le cadre roman

La comparaison des dictionnaires étymologiques courants du français laisse craindre que l'attention aux langues romanes tende à diminuer avec le temps : le seul dictionnaire qui cite systématiquement des congénères romans (par des renvois bibliographiques) est le plus ancien, Gamillscheg (1928¹–1969²). Bloch & Wartburg (1932¹) traite (s.v. *faire*) ou traite mieux (s.v. *charme*) les autres langues romanes que ses versions ultérieures (1949²–1968⁵) ; Dauzat, Dubois & Mitterand (1964–2001) s.v. *faire* suppriment, par rapport à Dauzat (1938¹), la mention du parallèle italien *fare*, utile à l'intelligence de l'origine des formes du futur-conditionnel du français ; les dictionnaires les plus récents (Picoche 1971¹–2002 ; Baumgartner & Ménard 1996) omettent complètement de citer les autres langues romanes. La seule exception à cette dégradation est le Robert historique, mais l'information donnée n'est pas d'une grande qualité : s.v. *charme*, « l'italien *carpino* et l'espagnol *carpe* remontent au même étymon » n'est ainsi que partiellement juste : des considérations d'ordre phonétique, chronologique et géobotanique (le charme

n'est pas autochtone sur la péninsule ibérique) incitent en effet à suivre Corominas et Pascual *in* DCECH, qui considèrent esp. *carpe* comme un emprunt (à l'occitan ?).

3 Proposition de réécriture

Après avoir détaillé les questions posées par l'étiquette glottonymique, le signifiant, le signifié, la combinatoire restreinte et l'insertion dans le cadre roman des étymons de *charme*, *faire* et *foin*, nous proposons ci-dessous le traitement de l'étymologie-origine de ces trois lexèmes telle qu'elle pourrait s'insérer par exemple dans le TLF, à travers la mise à jour de ses notices étymologiques que constitue le programme de recherche TLF-Étym. Dans le TLF, la notice étymologique fait suite immédiatement à la rubrique phonétique.

3.1 S.v. *charme*

Continueur de protorom. */karpin-u/ s.m. « arbre de la famille des bétulacées, au bois blanc et dur, aux fruits entourés de bractées trilobées, aux feuilles caduques, ovales et dentées (*Carpinus betulus* L.), charme » (maintenu par roum. *carpen*, it. *carpino*, frioul. *ciârpín(t)*, frpr. *charpeno*, occit. *carpre*, gasc. *carpe*) par mise en accord du genre avec la phonie de la finale d'un plus ancien */karpin-u/ s.f. dont le corrélat en latin écrit est *carpīnus*, -nī s.f. (TLL 3, 491), à la différence d'une mise en accord de la phonie de la finale avec le genre attestée par protorom. */karpin-a/ s.f. continué par frpr. *charpena* s.f. Cf. REW³ s.v. *carpīnus* ; von Wartburg *in* FEW 2, 406b-408a, CARPINUS ; Medori 2008–2010 *in* DÉRom s.v. */karpin-u/.

3.2 S.v. *faire*

Continueur de protorom. */fak-e-/ v.tr. « produire un effet à travers un travail manuel ou intellectuel, faire » (maintenu dans toutes les langues romanes, au moins dans une partie du paradigme flexionnel) et, pour le futur-conditionnel, de sa variante combinatoire */f-a-/ (qui est maintenue dans le futur-conditionnel de toutes les langues romanes à part le sarde et le roumain et dans l'infinitif de certaines d'entre elles, ainsi occit. *far* et it. *fare*), cf. lat. *facere* (TLL 6/1, 82-133). Le dégagement de */f-a-/ est en rapport direct avec l'apparition du futur périphrastique de type */f-a-re'aβ-e-/, inconnu en sarde et en roumain. Cf. Rydberg 1893 ; REW³ s.v. *facēre* ; von Wartburg *in* FEW 3, 346b-354b, FACÈRE ; Buchi 2009/2010 *in* DÉRom s.v. */fak-e-/.

3.3 S.v. *foin*

Continueur de protorom. */fen-u/ (~ */fēn-u/) s.n. « herbe fauchée puis séchée, foin » (maintenu dans toutes les langues romanes), cf. lat. *fænum* ~ *fēnum* (TLL 6/1, 165-166). L'alternance vocalique protoromane */e/ ~ */ē/ (reflétée par les issues du type esp. *heno* versus it. *fieno*) doit résulter de la réduction différentielle d'une diphtongue plus ancienne */ai/, dont l'hésitation entre les deux graphies latines est le reflet. Cf. Meyer-Lübke 1890 : 1, § 89 ; Bourciez & Bourciez 1982 : § 60, r. I ; REW³ s.v. *fēnum* ; von Wartburg *in* FEW 3, 455a-461b, FENUM ; Boutier *in* FEW 25, 1214ab, AVENA ; Reinhardt 2008–2010 *in* DÉRom s.v. */fen-u/ ~ */fēn-u/.

4 Conclusion

L'analyse des trois articles-témoins des sept dictionnaires étymologiques consacrés au seul français ne remplit pas d'espoir : les plus récents présentent les pratiques les plus contestables. On en viendrait à se

demander si l'étymologie française n'est pas une science régressive. L'étymologie héréditaire française a besoin d'être refondée, et cette refondation doit se faire dans le cadre de la linguistique romane.

La solution étymologique, si on veut à tout prix déterminer un étymon, va devenir propre à chaque cas, par sa datation et par sa motivation. Pour *charme*, c'est à un lexème protoroman des débuts de l'Empire romain qu'on va attribuer le rôle de l'étymon, pour des raisons morphologiques. Dans le cas de *foin*, la coexistence des points de départ des formes romanes */ ϕ en-u/ et */ ϕ en-u/ invite à choisir */ ϕ ain-u/ comme ancêtre commun, forme appartenant à un état de langue possédant une diphtongue */ai/ étrangère au système phonologique protoroman tel qu'on l'a reconstruit par ailleurs. Pour *faire*, si l'on veut établir une rupture, c'est au niveau sémantique qu'on devra se placer : le point de départ est un verbe polysémique « v.tr. poser, placer ; faire ; v.intr. se mettre, se placer » (cf. Ernout & Meillet 1959 s.v. *faciō*) dont les langues romanes n'ont retenu que le sens de « faire », prépondérant en latin classique, mais non exclusif. C'est donc à l'intérieur de la période latine que se déterminera la rupture.

La notion d'étymon change de nature, puisqu'elle n'est plus mise en rapport avec les étiquettes glottonymiques. Elle devient abstraite, du fait que le terme en vient à désigner le lexème immédiatement antérieur à la première rupture, de quelque nature qu'elle soit, constatée dans la continuité d'usage langagier. Mais elle acquiert une valeur heuristique : elle oblige à chercher les lignes de faille dans le paysage, les périodes de rupture, les cascades qui interrompent le trajet sur le long fleuve tranquille... et donc à repérer les faits intéressants. En tout cas, dire que fr. *charme* vient de lat. *carpīnus*, c'est manquer l'étape décisive, dans ce cas, de la rupture entre le latin et les langues romanes. Sur le plan de l'érudition, cette solution étymologique n'est qu'une variante notationnelle de celle qui est proposée ici. Du point de vue de la linguistique diachronique, il y a deux étymologies irréductibles l'une à l'autre, visant, l'une, un simple apport d'information et, l'autre, l'établissement de données et de processus linguistiques.

Nous espérons avoir montré, à travers une réflexion sur les pratiques actuelles, mais aussi sur la base de propositions concrètes de réécriture des étymons de *charme*, *faire* et *foin*, que la nouvelle approche de l'étymologie du lexique héréditaire français préconisée ici est en mesure de renouveler utilement la pratique lexicographique. Selon la vérité d'évidence dont ont fait leur titre S. Nielsen et S. Tarp : « Nothing is more practical than a good theory » (Nielsen & Tarp 2009 : IX).

Références bibliographiques

- Baldinger, K. (1959). L'étymologie hier et aujourd'hui. *Cahiers de l'association internationale des études françaises*, 11, 233-264.
- Bartoli, M. G. (2000 [1906]). *Il Dalmatico : resti di un'antica lingua romanza parlata da Veglia e Ragusa e sua collocazione nella Romània appennino-balcanica*. Rome : Istituto della Enciclopedia Italiana.
- Baumgartner, E. & Ménard, Ph. (1996). *Dictionnaire étymologique et historique de la langue française*. Paris : Librairie Générale Française.
- Bloch, O. & Wartburg, W. von (1968⁵ [1932¹]). *Dictionnaire étymologique de la langue française*. Paris : Presses Universitaires de France.
- Bourciez, É. & Bourciez, J. (1982 [1967]). *Phonétique française. Étude historique*. Paris : Klincksieck.
- Buchi, É. & Schweickard, W. (2008). Le *Dictionnaire Étymologique Roman* (DÉRom) : en guise de faire-part de naissance. *Lexicographica. International Annual for Lexicography*, 24, 351-357.
- Buchi, É. & Schweickard, W. (2009). Romanistique et étymologie du fonds lexical héréditaire : du REW au DÉRom (*Dictionnaire Étymologique Roman*). In : Alén Garabato, C. et al. (éd.) : *La Romanistique dans tous ses états*. Paris : L'Harmattan : 97-110.
- Campbell, L. (2004² [1998¹]). *Historical Linguistics. An Introduction*. Cambridge : MIT Press.
- Chambon, J.-P. (2007). Remarques sur la grammaire comparée-reconstruction en linguistique romane (situation, perspectives). *Mémoires de la Société de linguistique de Paris*, 15, 57-72.

- Chambon, J.-P. (à paraître). Pratique de la reconstruction en domaine (gallo)roman et grammaire comparée-reconstruction. À propos du traitement des mots héréditaires dans le TLF et le FEW. In : I. Choi-Jonin *et al.* (éd.) *Typologie et comparatisme. Hommages offerts à Alain Lemaréchal*. Louvain : Peeters.
- Dardel, R. de (1985). Le sarde représente-t-il un état précoce du roman commun ? *Revue de linguistique romane*, 49, 263-269.
- Dardel, R. de (1996). *À la recherche du protoroman*. Tübingen : Niemeyer.
- Dardel, R. de (2009). La valeur ajoutée du latin global. *Revue de linguistique romane*, 73, 5-26.
- Dauzat, A., Dubois, J. & Mitterand, H. (2001 [1938¹]). *Dictionnaire étymologique et historique du français*. Paris : Larousse.
- DCECH = Corominas, J. & Pascual, J. A. (1980–1991). *Diccionario crítico etimológico castellano e hispánico* (6 vol.). Madrid : Gredos.
- DÉRom = Buchi, É. & Schweickard, W. (dir.) (2008–). *Dictionnaire Étymologique Roman (DÉRom)*. Nancy : ATILF : site Internet (<http://www.atilf.fr/DERom>).
- Doria, M. (1989). Dalmatico. Storia linguistica interna. In : Holtus, G. *et al.* (éd.) : *Lexikon der Romanistischen Linguistik (LRL)*. Tübingen : Niemeyer : 3 : 522-530.
- Durkin, Ph. (2009). *The Oxford guide to etymology*. Oxford : Oxford University Press.
- Ernout, A. & Meillet, Antoine (1959⁴ [1932¹]). *Dictionnaire étymologique de la langue latine. Histoire des mots*. Paris : Klincksieck.
- FEW = Wartburg, W. von *et al.* (1922–2002). *Französisches Etymologisches Wörterbuch. Eine darstellung des galloromanischen sprachschatzes* (25 vol.). Bonn/Heidelberg/Leipzig-Berlin/Bâle : Klopp/Winter/Teubner/Zbinden.
- Florescu, C. (2009). Limba română în *Dictionnaire Étymologique Roman DÉRom* (< *Romanisches Etymologisches Wörterbuch REW*). In : Botoșineanu, L. *et al.* (éd.) : *Distorsionări în comunicarea lingvistică, literară și etnofolclorică românească și contextul european*. Iași : ALFA/Asociația Culturală "A. Philippide" : 153-159.
- Fox, A. (1995). *Linguistic Reconstruction. An Introduction to Theory and Method*. Oxford : Oxford University Press.
- Gamillscheg, E. (1969² [1928¹]). *Etymologisches Wörterbuch der französischen Sprache*. Heidelberg : Winter.
- Gardette, P. (1983). *Études de géographie linguistique*. Strasbourg : Société de Linguistique romane.
- Geckeler, H. & Dietrich, W. (2007⁴ [1995¹]). *Einführung in die französische Sprachwissenschaft*. Berlin : Schmidt.
- Harrison, S. P. (2003). On the Limits of the Comparative Method. In : Joseph, B. D. & Janda, R. D. (éd.) : *The Handbook of Historical Linguistics*. Malden/Oxford/Victoria : Blackwell : 213-243.
- Koch, P. & Oesterreicher, W. (2008). Comparaison historique de l'architecture des langues romanes. In : Ernst, G. *et al.* (éd.) : *Histoire linguistique de la Romania. Manuel international d'histoire linguistique de la Romania*. Berlin/New York : de Gruyter : 3 : 2575-2610.
- Lausberg, H. (1963–1972² [1957–1962¹]). *Romanische Sprachwissenschaft* (3 vol.). Berlin : de Gruyter.
- Lloyd, Paul M. (1987). *From Latin to Spanish. Vol. I: Historical Phonology and Morphology of the Spanish Language*. Philadelphie : American Philosophical Society.
- Mańczak, W. (1991). Origine des langues romanes. *Revue des langues romanes*, 95, 159-170.
- Meyer-Lübke, W. (1890–1906). *Grammaire des langues romanes* (4 vol.). Paris : Welter.
- Meyer-Lübke, W. (1894). Compte rendu de Rydberg 1893. *Zeitschrift für romanische Philologie*, 18, 434-439.
- Nielsen, S. & Tarp, S. (2009). Introduction. Nothing is more practical than a good theory. In : *Lexicography in the 21st Century. In honour of Henning Bergenholtz*. Amsterdam/Philadelphia : Benjamins : IX-XI.
- Pfister, M. & Lupis, A. (2001). *Introduzione all'etimologia romanza*. Soveria Mannelli : Rubbettino.
- Picoche, J. (2002 [1971¹]). *Dictionnaire étymologique du français*. Paris : Le Robert.

- Polguère, A. (2008² [2003¹]). *Lexicologie et sémantique lexicale. Notions fondamentales*. Montréal : Les Presses de l'Université de Montréal.
- REW³ = Meyer-Lübke, Wilhelm (1930–1935³ [1911–1920¹]). *Romanisches Etymologisches Wörterbuch*. Heidelberg : Winter.
- Rey-Debove, J. & Rey, A. (dir.) (2009). *Le nouveau Petit Robert. Dictionnaire alphabétique et analogique de la langue française*. Paris : Le Robert.
- Robert historique = Rey, A. (dir.) (1998² [1992¹]). *Dictionnaire historique de la langue française*. Paris : Le Robert.
- Rohlf, G. (1949–1954). *Historische Grammatik der italienischen Sprache und ihrer Mundarten* (3 vol.). Berne : Francke.
- Rosetti, Al. (1986). *Istoria limbii române. De la origini și pînă la începutul secolului al XVII-lea*. Bucarest : Editura Științifică și Enciclopedică.
- Rydberg, G. (1893). *Le Développement de facere dans les langues romanes*. Paris : Noblet.
- Straka, G. (1956). La dislocation linguistique de la Romania et la formation des langues romanes à la lumière de la chronologie relative des changements phonétiques. *Revue de linguistique romane*, 20, 249-267.
- Tekavčić, P. (1976/1977). Sulla forma verbale vegliota *féro* e sull'origine del futuro veglioto. *Incontri linguistici*, 3, 71-89.
- TLF = Imbs, P. & Quemada, B. (dir.) (1971–1994). *Trésor de la langue française. Dictionnaire de la langue du XIX^e et du XX^e siècle (1789–1960)* (16 vol.). Paris : Éditions du CNRS/Gallimard.
- TLF-Étym = Steinfeld, N. (dir.) (2005–). *Programme de recherche TLF-Étym*. Nancy : ATILF : site Internet (<http://www.atilf.fr/tlf-etym>).
- TLL = (1900–). *Thesaurus linguae latinae*. Leipzig : Teubner.
- Vaan, Michiel de (2008). *Etymological dictionary of Latin and the other Italic languages*. Leiden/Boston : Brill.
- Van Acker, M. (2007). Quelques réflexions d'ordre conceptuel et terminologique relatives à la transition latin/langues romanes à partir de la notion de « latin vulgaire ». *Zeitschrift für romanische Philologie*, 123, 593-617.

¹ Ce processus s'appuie sur l'habitude répandue de penser la langue à partir de sa notation par l'écriture et de la tradition de celle-ci, distorsion qu'on pourrait appeler grapho-centrique.

² Que nous remercions par ailleurs de sa relecture d'une première version de ce texte. Celui-ci a également bénéficié des commentaires de deux relecteurs anonymes.

³ Pfister & Lupis (2001 : 37) : « In questo modo, tra gli studiosi di lingue indoeuropee e quelli di lingue romane si è determinata quasi una sorta di spartizione tacita del lavoro etimologico ; per cui i romanisti si accontentano normalmente dell'etimo latino o della eventuale base di sostrato o superstrato e non arretrano nella loro ricerca fina alle radici indoeuropee ».

⁴ De Dardel (2009 : 7) : « le latin global est le latin sous toutes les formes qu'il a pu revêtir depuis la fondation de Rome, dans le monde romain occidental antique et par la suite dans les parlers romans, en Europe, ainsi que dans le latin moderne pratiqué par les sciences et le culte ».

⁵ À titre d'exemple, on peut citer la formule nuancée de Lausberg (1963 : 1 : 67) : « Eine auf die Dauer bedeutendere Rolle als die Schriftsprache der römischen Oberschicht spielte bei der schließlich auch die tieferen Schichten erfassenden Romanisierung des Imperiums die Umgangssprache des gemeinen Mannes, das sog. Vulgärlatein (*sermo vulgaris, plebeius, quotidianus, rusticus*) : die Sprache des Bauern, des Soldaten, des Händlers, des Sklaven. Auch die tägliche Umgangssprache der Gebildeten selbst stand nicht auf dem grammatisch-rhetorischen Niveau der Schriftsprache ». Cf. aussi Bourciez & Bourciez (1982 : 1).

⁶ Par exemple dans le manuel de Geckeler & Dietrich (2007, 76-77). La formulation n'a guère varié depuis Bourciez & Bourciez (1982 : 24) :

Le vocalisme du latin vulgaire ou parlé a éprouvé de graves modifications pendant la période impériale. Les voyelles ont cessé peu à peu d'être prononcées longues ou brèves, et en sont venues à différer entre elles seulement par le timbre. [...] Il en

résulte qu'au terme de l'évolution, on a obtenu en tout sept voyelles (*a, e, ē, o, i, u*), dont la correspondance avec celles du latin classique est la suivante :

<i>a</i> vulgaire	=	<i>ā, ǣ</i> classiques	(<i>cāru, mǣre</i>)
<i>e</i>	—	= <i>ě</i>	(<i>měl</i>)
<i>ē</i>	—	= <i>ē, ĭ</i>	(<i>tēla, pīlu</i>)
<i>i</i>	—	= <i>ī</i>	(<i>filu</i>)
<i>o</i>	—	= <i>ō</i>	(<i>mōla</i>)
<i>o</i>	—	= <i>ō, ū</i>	(<i>flōre, gūla</i>)
<i>u</i>	—	= <i>ū</i>	(<i>mūru</i>).

Comme on sait, il ne s'agit pas là du vocalisme de l'ancêtre des langues romanes dans leur ensemble, mais seulement d'une partie d'entre elles. Lausberg (1963) délimite ainsi quatre systèmes vocaliques du roman commun : (a) le « sog[enantes] 'vulgärlateinische[s]' System » (Lausberg 1963 : 1 : 144-146) : **ī/ > *i/, *ī/ ~ *ē/ > *e/, *ē/ > *e/, *ā/ ~ *ǣ/ > *a/, *ō/ > *o/, *ō/ ~ *ū/ > *o/, *ū/ > *u/* ; (b) le « archaische[s] System in Sardinien, Lukanien und Afrika » (*ibid.* 146-147) : **ī/ ~ *ī/ > *i/, *ē/ ~ *ē/ > *e/, > *ā/ ~ *ā/ > *a/, *ō/ ~ *ō/ > *o/, *ū/ ~ *ū/ > *u/* ; (c) le « Kompromißsystem in Ostlukanien und im Balkanromanischen » (*ibid.* 148-149) : **ī/ > *i/, *ī/ ~ *ē/ > *e/, *ē/ > *e/, *ā/ ~ *ā/ > *a/, *ō/ ~ *ō/ > *o/, *ū/ ~ *ū/ > *u/* ; (d) le « sizilianische[s] System » (*ibid.* 149) : **ī/ ~ *ī/ ~ *ē/ > *i/, *ē/ > *e/, *ā/ ~ *ā/ > *a/, *ō/ > *o/, *ō/ ~ *ū/ ~ *ū/ > *u/*.

Il ne fait pas de doute que le latin vulgaire des étymologistes-lexicographes francisants est identique au latin vulgaire de Lausberg, que l'on pourrait appeler latin vulgaire continental occidental. Ce latin vulgaire, caractérisé par le double syncrétisme de **ī/* et de **ē/* et de **ō/* et de **ū/*, fait partie du bagage universitaire de tout francisant qui se respecte – y compris des signataires de ces lignes. Or, si ce latin vulgaire continental représente bien l'ancêtre du français – et de l'italien, de l'espagnol etc. –, il ne s'agit pas de l'ancêtre du sarde, par exemple.

Force est de constater, dès lors, que le latin vulgaire véhiculé par les linguistes francisants représente une variété de latin marquée non seulement sur l'axe diastratique et diaphasique, mais aussi sur l'axe diachronique et diatopique : il s'agit de la variété du diasystème latin caractéristique de l'époque d'après la séparation du sarde du tronc commun (que l'on datera au plus tard du 2^e siècle, cf. Straka 1956 : 256 ; de Dardel 1985 : 268) et de celle du roumain (au 3^e siècle, cf. Straka 1956 : 258 ; Rosetti 1986 : 184), et qui englobe la seule Romania continentale à l'exception de la Dacie, de la Sicile, de la Basilicate et de la Calabre septentrionale.

⁷ Les formes romanes citées sont le plus souvent les entrées des dictionnaires que nous citons, directement ou indirectement ; dans ce dernier cas, on y accède par l'article correspondant du DÉRom.

⁸ « Issu du lat. class. *facere* 'réaliser quelque chose ; créer, commettre' et servant de substitut à un verbe précédemment exprimé ».

⁹ « À l'origine, *facere* signifiait 'placer, poser' puis a été employé de façon technique, par exemple en religion dans *sacrum facere* 'placer un sacrifice (sur l'autel)', d'où 'faire (un sacrifice)'. Le sens de 'placer' est pris par *ponere* [...] et *facere* signifie alors 'causer', 'exciter', 'travailler', 'faire artificiellement' ; il s'emploie aussi en litote et comme substitut d'un autre verbe et entre, à l'époque impériale, dans des emplois impersonnels ».

¹⁰ « Du lat. *facere* 'réaliser, accomplir' ».

¹¹ « La *combinatoire restreinte* d'un signe linguistique est l'ensemble des contraintes propres à ce signe qui limitent sa capacité de se combiner avec d'autres signes linguistiques et qui ne peuvent être déduites ni de son signifié ni de son signifiant. Par opposition, la *combinatoire libre* d'un signe linguistique est sa capacité de se combiner avec d'autres signes linguistiques qui est directement héritée de son signifié et de son signifiant ».