

Latin oral et latin écrit en étymologie romane : l'exemple du DÉRom (*Dictionnaire Étymologique Roman*)

Myriam BENARROCH, Université Paris-Sorbonne

Il existe aujourd'hui un consensus sur l'idée que les langues romanes sont nées du latin oral et non du latin écrit¹. Mais qu'est-ce que ce latin oral ? Est-ce une langue différente du latin que nous ont transmis les textes ? Ou bien un ensemble de variétés appartenant à un diasystème auquel ressortit aussi le latin dit *classique* ? Quel accès peut-on avoir aujourd'hui à cette langue orale ? Les textes latins, qu'ils soient rédigés dans une langue classique ou bien dans un latin plus relâché sont-ils suffisants pour nous permettre d'appréhender les changements existant entre une norme écrite, plus ou moins figée, et la manière dont les locuteurs ordinaires s'exprimaient oralement ? Y a-t-il d'autres moyens d'accéder, tout au moins en partie, à ce latin oral ? Nous nous proposons de montrer, à travers quelques exemples pris dans le *Dictionnaire Étymologique Roman* (DÉRom), comment la méthode mise en œuvre dans ce dictionnaire permet d'envisager un ensemble de variétés du latin oral et d'accéder, quoique très partiellement, aux caractéristiques phoniques, morphologiques et sémantiques de son lexique et, ce faisant, à mesurer la distance séparant ces variétés orales du latin écrit de l'Antiquité.

1 Latin et étymologie romane

1.1 Le diasystème du latin

Aux trois dimensions traditionnelles de la variation linguistique auxquelles Coseriu (1998) fait appel pour décrire le diasystème d'une langue historique, dimensions diatopique, diastratique et diaphasique, Koch et Oesterreicher ajoutent une quatrième dimension, celle de proximité / distance. Ils affirment que « pour la variation d'une langue donnée, le point de référence est un continuum universel qui s'étend entre les pôles de l'oralité et de la scripturalité, ou plus exactement entre l'immédiat communicatif et la distance communicative » (2008 : 2575). Ils voient ainsi un « continuum conceptionnel » à mettre en relation avec les deux formes de réalisation médiales phonique et graphique, ce qui aboutit à un schéma où apparaissent quatre domaines : « immédiat phonique (A) et distance phonique (B), immédiat graphique (C) et distance graphique (D) ». On peut alors se poser cette question : ce schéma s'applique-t-il au latin ? Faut-il formuler cette question comme le fait l'un de ces deux auteurs : « Le latin – une langue pas tout à fait comme les autres ? » (Koch 2008 : 43). À cette question, Koch répond que oui, le latin est une langue comme les autres et il le justifie ainsi : « [...] la langue historique latine, comme toutes les autres langues historiques, forme un espace variationnel englobant les quatre dimensions de la variation linguistique [...] : immédiat-distance [...], diaphasie [...], diastratie [...] et diatopie », (Koch 2008 : 62). Gaetano Berruto propose un schéma de cercles concentriques pour expliquer sa conception de la hiérarchie des variétés avec, de l'intérieur vers l'extérieur, les variétés diamésique, diaphasique, diastratique et diatopique. Pour Kabatek (2000 : 318), la variation diamésique n'existe pas comme dimension propre, pour la langue historique, elle fait partie de la langue diaphasique. L'approche la plus éclairante et la plus précise du diasystème latin est, à nos yeux, l'« approche sociolinguistique systématique » de Piera Molinelli, qui reprend le modèle théorique que Berruto avait utilisé pour l'italien en l'appliquant à la situation

¹ Nous remercions chaleureusement Jean-Pierre Chambon, Jérémie Delorme, Cristina Florescu, Gilles Petrequin et Jan Reinhardt qui ont eu la gentillesse de relire la première version de cet article.

linguistique du latin dans les deux premiers siècles de l'Empire, c'est-à-dire avant la crise du 3^e siècle ap. J.C. Elle dessine un schéma à quatre axes : diastratique (vertical), diamésique (horizontal), diaphasique (oblique), le quatrième « axe », diatopique, étant constitué d'une ellipse qui englobe les trois autres. Le premier sous-tend toutes les autres dimensions mais c'est du côté de l'oral qu'il est le plus chargé. Molinelli voit dans le diasystème latin huit variétés : (1) standard, (2) littéraire, (3) familier (*colloquiale*), (4) populaire, (5) argotique, (6) juridique, (7) technico-scientifique, (8) chrétien. Sur l'axe diastratique se trouve, en bas, le latin populaire, tel qu'il est reflété dans les documents pompéiens ou encore dans les tablettes d'exécration (Molinelli 2006 : 472). Pour le latin familier (3), l'auteur convient que nous ne pouvons en avoir qu'un reflet à travers des textes comme les comédies ou les lettres privées. L'axe diaphasique va du plus formel, le latin juridique (6) au moins formel, le latin argotique (5). Le latin familier et le latin populaire se recoupent mais présentent une différence diastratique, le second étant réservé aux couches populaires tandis que le premier est aussi parlé par des couches plus cultivées de la population : « una lettera alla famiglia scritta da Cicerone si differenzia da una del soldato Claudio Terenziano non rispetto al mezzo o alla situazione, ma rispetto al loro differente livello socio-culturale » (Molinelli 2006 : 473-474). Sans surprise, le latin standard se trouve au centre de tous les axes, mais on voit qu'il est quand même tiré vers le haut, sur l'axe diastratique et vers la gauche, sur les axes diaphasique et diamésique. Le schéma de Molinelli traduit bien la complexité de la situation du diasystème latin, tout en ne contredisant pas celui de Koch et Oesterreicher. Toutefois, parmi les huit variétés évoquées, seules trois ressortissent exclusivement à la langue orale ; or, ce sont numériquement les plus importantes : il ne faut pas oublier qu'à l'époque envisagée, moins de 20 % de la population avait accès à la langue écrite.

1.2 Le latin dans l'étymologie romane traditionnelle

1.2.1 Latin classique et latin vulgaire dans l'étymologie romane traditionnelle

Dans les dictionnaires étymologiques traditionnels idioromans², l'étymon proposé pour un mot appartenant au lexique héréditaire correspond, en général, à la forme du latin écrit de l'Antiquité. S'il s'agit d'un substantif, cette forme est au nominatif ou à l'accusatif et revêt sa graphie traditionnelle, les voyelles portant la marque de quantité :

(1) « *Agosto*, del lat. *augūstus*, íd., nombre puesto en memoria del emperador Octavio Augusto », DCECH.

Cette marque de quantité est généralement conservée dans la graphie du latin dit « vulgaire » :

(2) « *Agosto*, s. Do lat. *Augūstu-* [...] a pronúncia vulgar era *Agūstu-*», DELP.

(3) « *agust* [...] Lat.pop. *agūstus*, formă redusă de la *augūstus* », DER.

(4) « 786. *augūstus* 'Monat August', 2. *agūstus* », REW.

Or Väänänen (1959 : 18) tire de son analyse des inscriptions pompéiennes la conclusion que l'évolution ayant conduit à passer de l'opposition quantitative (durée) à l'opposition qualitative (timbre) des voyelles remonte à une époque antérieure à la destruction de Pompéi (en 79 de notre ère).

Dans le parler populaire, ce développement [le passage de l'opposition quantitative à l'opposition qualitative des voyelles] semble être amorcé à une date relativement reculée. En effet, nos matériaux pompéiens renferment déjà des particularités phonétiques qui dénoncent l'état périmé du rythme quantitatif. Cette preuve capitale, croyons-nous, est livrée notamment par nombre d'exemples attestant la confusion entre

² Le terme *idioroman* a été proposé pour la première fois par Buchi et Schweickard (2009 : 101) à propos de la lexicographie : « la lexicographie que l'on pourrait appeler idioromane, c'est-à-dire concernant les différentes langues romanes prises isolément ». *Idioroman* signifie donc : « relatif à une langue romane particulière » et s'oppose à *panroman* (cf. Buchi 2010a : 43 ; Benarroch à paraître b).

ae et *ē*, ainsi que, d'une manière moins probante, par quelques inscriptions en vers négligeant la quantité des voyelles.

Est-il donc pertinent de représenter les voyelles du latin « vulgaire », ou « populaire » avec la marque de quantité ?

Le recours à la langue parlée est souvent réduit à la portion congrue et la terminologie utilisée dans ce domaine est souvent peu claire et manque parfois de cohérence, comme on peut le voir dans les articles *août* et *février* du TLF :

(5) « Du lat. *augustus (mensis)*, d'où *augustus* subst. [...]. *Augustus* devenu en b.-lat. *agustus* (IV^e s.) », TLF.

(6) « Du lat. pop. *febrarius*, lat. class. *februarius* », TLF.

Les exemples ci-dessus inspirent plusieurs remarques et suscitent autant de questions : (i) le latin écrit de l'Antiquité est désigné dans (5) sous le nom de *latin* tout court, et dans (6), sous *latin classique*, les deux exemples étant empruntés au TLF ; (ii) à côté de ce latin dit *classique*, apparaît, dans (5) un *bas-latin*, et dans (6) un *latin populaire* ; (iii) l'ordre dans lequel sont données les deux variétés de latin n'est pas le même ; (iv), dans les deux cas, la forme du latin non classique est donnée au nominatif. Que comprendre de l'abréviation « lat. » (c.-à-d. *latin* tout court) dans le cas de *augustus*? Qu'il s'agit de l'ensemble du diasystème latin ou seulement du latin *classique*? Et dans ce cas, pourquoi ce manque de cohérence avec le glottonyme attribué à *februarius*? Pourquoi pour *agustus* a-t-on une dénomination d'ordre diachronique, accompagnée de la date (IV^e s.), alors que pour *febrarius*, la désignation est clairement diastatique? Pourquoi cette désignation de *bas-latin* alors que l'on peut aisément vérifier (TLL) que, du point de vue diachronique, *augustus* est attesté au tout début de notre ère (avant l'an 8, chez Ovide), mais que des formes présentant la réduction de la diphtongue /au/ en /a/ se trouvent fréquemment dans des inscriptions dès la première moitié du premier siècle (cf. DÉRom, s.v. */a'gust-u).

Des cinq dictionnaires cités ci-dessus, seuls le REW et le DER attribuent explicitement, pour la désignation du mois d'août, les issues romanes à la forme du latin oral³. Les autres dictionnaires ne mentionnent que la forme du latin écrit (DCECH) ou, dans le meilleur des cas, ils évoquent une forme *vulgaire*, de *bas-latin* ou *populaire* sans préciser la filiation de l'issue romane. Ces dictionnaires prennent massivement en compte le témoignage des textes écrits (comment faire autrement pour une langue *morte*?) et sont gênés aux entournures lorsqu'apparaît un lexème qui ne correspond pas à la forme canonique du latin. Comment alors le désigner? Est-ce du latin *vulgaire*? Du *bas-latin*? Du latin *populaire*?

1.2.2 Qu'est-ce que le latin vulgaire ?

Nous ne prétendons pas ici traiter en profondeur la question du latin vulgaire qui a fait et continue à faire l'objet de nombreuses publications, de nombreux colloques, en particulier la série des colloques internationaux « Latin vulgaire, latin tardif » dont la dixième édition, organisée par Piera Molinelli, s'est tenue à Bergame en septembre 2012. Nous nous contenterons d'évoquer la définition que donnent de ce terme deux spécialistes de la question dans des publications déjà anciennes, József Herman, organisateur du premier colloque de cette série et Veikko Väänänen, auteur de *l'Introduction au latin vulgaire*, vite devenu l'ouvrage de référence sur la question. Herman (1967 : 16) définit le latin vulgaire comme « la langue parlée des couches peu influencées ou non influencées par l'enseignement scolaire et par les modèles littéraires ». Il précise que cette appellation ne comporte aucune limite chronologique et que le latin vulgaire « était constitué d'un ensemble de faits complexes et

³ En effet, c'est à *agustus* que le REW fait remonter les issues romanes qu'il énumère dans la suite de l'article.

mouvants », qu'il avait évolué dans le temps, qu'il comportait des « variations locales », « divers 'styles' et divers argots de métiers » (Herman 1967 : 17). Väänänen (1963 : 5) fait remarquer que l'expression *latin vulgaire* remonte à Cicéron, qui désigne sous le nom de *sermo plebeius* des tours populaires qu'il n'hésitait pas à employer lui-même dans sa correspondance intime et pour lequel « *vulgari sermone* veut dire 'en des termes de la langue courante' » de même que *vulgare genus orationis* a le sens neutre de 'manière courante de s'exprimer' ». Il considère le « terme » *latin vulgaire* comme « peu adéquat mais consacré par la tradition » et indique avoir fait le choix de le conserver « pour faire ressortir [son] objectif spécial » qui est d'envisager « toutes les particularités et les tendances plus ou moins vivantes, propres à la langue populaire et familière, et qui se soustraient à la norme classique et, en général, littéraire » et de chercher à saisir « la langue vivante et réelle » (Väänänen 1963 : 5-6). Il précise, à son tour que « Le latin vulgaire [...] comprend les états successifs depuis la fixation du latin commun, à l'issue de la période archaïque, jusqu'à la veille des premières consignations par écrit de textes en langue romane ; il n'exclut ni les variations sociales, ni mêmes régionales. » (Väänänen 1963 : 6). Nous voyons que tant Herman que Väänänen considèrent le latin vulgaire essentiellement comme une variété diastratique et diaphasique du latin, parlée sur une longue période et comportant des « variations locales ». Toutefois, lorsque Väänänen entre dans la description interne de ce latin vulgaire, et en particulier dans la description de son système vocalique, il admet « (Väänänen 1963 : 30) que le système à sept voyelles qu'il propose, et qui sera considéré par la suite comme le système vocalique du « latin vulgaire », n'est pas valable pour le sarde, le roumain et le sicilien. Ce fameux latin vulgaire ne peut donc prétendre être l'ancêtre de toutes les langues romanes. C'est pourquoi il vaut mieux parler de latin oral ou de latin parlé et considérer cette entité linguistique dans l'ensemble du diasystème latin avec toute la variation à laquelle elle a été soumise : diastratique et diaphasique, certes, mais aussi diachronique diatopique et diamésique.

1.2.3 Le latin oral, source de l'étymologie romane

La prise de conscience que le latin écrit de l'Antiquité ne peut être donné comme étant à l'origine directe des langues romanes et que l'ancêtre commun des langues romanes ne peut être que le latin oral ne date pas d'hier. Dès 1925, Meillet écrivait :

[...] la langue qui a survécu n'est pas celle qui s'écrivait. Entre le latin littéraire, qui est conservé par les textes des écrivains, et le latin parlé, que continuent les langues romanes, il y avait des différences, variables suivant les individus et suivant leur degré de culture. Or les langues romanes ne continuent pas le latin littéraire. La « bouche » se nommait en latin écrit *ōs* ; mais ce qui a subsisté, c'est le nom du vulgaire *bucca*. Et ainsi dans une foule de cas. (Meillet 1925 : 8-9)

Mais comment avoir accès aujourd'hui à une langue qui n'existe plus ? Parce que « même dans les meilleurs cas, la langue écrite est bien loin d'enregistrer exactement les changements successifs de la langue parlée » (Meillet 1925 : 7) et que « tout l'essentiel des changements linguistiques a lieu hors de notre vue » (Meillet 1925 : 9), Meillet prône la comparaison comme « le seul instrument efficace » pour faire l'histoire des langues en général (Meillet 1925 : 11). Il s'agit, pour lui, de comparer des états de langue les uns aux autres, en étudiant « la valeur probante » de combinaisons « rigoureuses et précises » (Meillet, 1925 : 11). Dès 1901, Meyer-Lübke considérait que, étant donné la différence importante qui existait à l'époque de l'Empire romain entre le latin écrit et le latin parlé et le fait que la langue parlée évoluait beaucoup plus rapidement que la langue écrite, la source principale d'accès au latin « vulgaire » devait être les langues romanes (Meyer-Lübke 1901 : 82).

1.3 Le changement de paradigme

1.3.1 La méthode de la grammaire comparée-reconstruction

Robert de Dardel, qui s'est très tôt intéressé à la comparaison dans le cadre des langues romanes, a, dans un premier temps, pratiqué un comparatisme historique à partir des données du latin écrit. Il considère, par la suite, que ce fut une erreur car « les formes du latin écrit sont, du moins en première analyse, des faits de parole, dont la présence, même simultanée, peut-être fortuite » (Dardel, 2007 : 335). Dès lors, pour lui, la seule méthode sûre est de « reconstruire le protoroman à partir des parlers romans [...] parce que le protoroman obtenu par cette voie consiste en faits de langue et constitue un système » (Dardel 2007 : 335 ; cf. *infra* pour la définition de *protoroman*). En effet, la méthode de la grammaire comparée – reconstruction est abondamment utilisée pour de nombreuses familles de langues (familles indo-européenne, bantoue, austronésienne) et, à l'intérieur de la famille indo-européenne, pour les branches germanique, slave, celtique. Une branche fait toutefois exception, celle des langues romanes, et l'argument – fallacieux – évoqué est que l'on n'a pas besoin de cette méthode puisqu'on a la chance de connaître l'ancêtre commun, le latin, et que celui-ci bénéficie d'une vaste documentation. Retour à la case départ : le latin écrit, source de l'étymologie romane !

Jean-Pierre Chambon renforce la thèse de Dardel et va plus loin. Il dénonce « [une] conception et [une] pratique graphocentristes » à l'œuvre dans l'étymologie romane traditionnelle, qui « sont étrangères au point de vue de la linguistique [...] et entrent en contradiction flagrante avec la notion de 'mot héréditaire' pourtant universellement reçue » (Chambon 2010 : 64). Il prône une « rupture épistémologique bien tempérée » (Chambon, à paraître a) qui se traduise par un changement de paradigme. En quoi consiste ce changement de paradigme ? Il s'agit de substituer à la démarche prospective traditionnelle – du latin aux langues romanes –, une démarche rétrospective – des parlers romans au latin –, qui est celle de la grammaire comparée – reconstruction : « il conviendrait en premier lieu [...] d'appliquer aux parlers romans la méthode comparative : rien qu'elle, telle quelle, et, pour tout dire, dans sa sèche simplicité » (Chambon 2007 : 68). Ce qui conduit à chercher l'ancêtre commun, qui ne peut être le latin écrit de l'Antiquité :

[...] les mots du latin écrit de l'Antiquité ne sauraient être placés à l'origine des mots héréditaires du français ou des autres langues (gallo)romanes [...]. Le seul moyen de faire venir à l'existence l'étymon (oral) d'un mot héréditaire est de le *reconstruire* sur la base de la *comparaison* entre formes *orales* affines, c'est-à-dire dont on a montré qu'elles étaient reliées par un ensemble de correspondances phoniques régulières (Chambon 2010, 64).

1.3.2 Le protoroman : reflet partiel du latin parlé

Revenons brièvement à Väänänen pour continuer à le citer lorsqu'il justifie son choix de conserver la dénomination de latin vulgaire. Il rejette le terme de *protoroman* et s'en explique :

Des termes comme « roman commun » ou « protoroman » ne serviraient pas à notre propos, car ils s'appliquent, en principe, à une forme de langue reconstituée d'une part, et d'autre part ils excluent l'étude des phénomènes qui faisaient partie du fond populaire du latin, mais n'ont pas survécu en roman. (Väänänen 1963 : 6)

Son point de vue est celui d'un latiniste. Du point de vue de l'étymologie romane, seuls nous intéressent les phénomènes qui ont justement survécu et cette « forme de langue reconstituée » – ou mieux, reconstruite – permet d'avoir accès, quoique partiellement, certes, à cette langue orale qui lui tient tant à cœur. Nous avons vu que le phonétisme vocalique qu'il décrit, dit « du latin vulgaire » n'est pas valable pour l'ensemble des langues romanes. Ce système vocalique à deux classes de localisation, quatre degrés d'aperture et sept phonèmes correspond à l'évolution majoritaire mais ne représente pas l'ancêtre du sarde ni celui du

roumain, ni même celui de certains parlers de l'Italie méridionale (Sicile, Basilicate, Calabre septentrionale). Or la comparaison du système vocalique sous l'accent du sarde et du toscan permet de reconstruire le système phonologique de l'ancêtre commun dans son entier. Il aboutit à un système de neuf protophonèmes, répartis en deux classes de localisation et cinq degrés d'aperture :

/i/	/u/
/ɪ/	/ʊ/
/e/	/o/
/ɛ/	/ɔ/
/a/	

Ce système vocalique est celui de l'ancêtre commun à tous les idiomes romans, le protoroman, qui remonte donc à avant la première séparation du tronc commun, celle du sarde (survenue au plus tard au 2^e s. ap. J.C., Straka 1956 : 256). Ce système à cinq degrés d'aperture, parce que rare, est instable, c'est pourquoi il n'a survécu dans aucun idiome roman. Il ne peut donc pratiquement pas transparaitre dans les données écrites. Seule la reconstruction permet d'y avoir accès. La méthode de la grammaire comparée-reconstruction permet d'accéder non seulement à la phonologie et au lexique du protoroman mais aussi à certains aspects morphologiques (cf. Dardel 1965 ; 1976 ; 1996 ; Dardel et Wüest 1993).

L'utilisation du terme *protoroman* et l'utilisation de la méthode de la grammaire comparée-reconstruction qui conduit à cette protolangue ne font pas l'unanimité : « Perchè [...] dobbiamo usare un nome e ipotizzare una entità che non ha alcun fondamento storico ? » (Vàrvaro 2011b : 626). La réponse se trouve dans un article publié un an auparavant :

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.) ». (Buchi *et al.* 2010, 114)

Le protoroman est donc l'ancêtre commun des parlers romans tel qu'il est reconstruit par la méthode de la grammaire comparée – reconstruction. S'il appartient au diasystème latin, il se distingue du latin écrit de l'Antiquité par des aspects phonétiques, morphologiques, syntaxiques et lexicaux, comme nous allons le voir à travers quelques exemples du DÉRom. Le protoroman des premiers temps s'est par la suite divisé en protoromans régionaux. On connaît encore mal la chronologie des séparations successives du tronc commun (ce qui est un comble pour un sous-groupe de langues dont on dit connaître l'ancêtre commun !). On a toutefois les travaux de Straka, Rosetti, Stefenelli, Chambon, Greub, Seidl qui balisent le terrain et la reconstruction des 58 étymons du DÉRom semble commencer à dessiner certains contours de cette chronologie du protoroman (cf. Buchi/Schweickard à paraître). Mais il faudra attendre d'avoir reconstruit un nombre beaucoup plus important d'étymons pour pouvoir se faire une idée plus précise.

L'utilisation de la méthode de la grammaire comparée – reconstruction en étymologie du lexique héréditaire roman ne signifie nullement l'élimination des autres méthodes jusque là utilisées, bien au contraire. En effet, elle ne prétend nullement reconstruire le système linguistique, ici roman, dans son intégralité. Ses résultats doivent être confrontés à ceux de la philologie latine – et donc aux textes écrits –, de la chronologie relative, de la géolinguistique

ainsi qu'aux données extralinguistiques telles le processus de latinisation de la Romania en corrélation avec la conquête romaine (cf. Buchi 2010b, 163). Si la primauté revient à la méthode de la grammaire comparée – reconstruction, c'est l'ensemble de ces méthodes qui est mis en œuvre dans le DÉRom.

2 Le DÉRom et la place du latin écrit

2.1 Présentation du DÉRom

Le *Dictionnaire Étymologique Roman* (DÉRom) est le fruit d'un projet franco-allemand, cofinancé par l'ANR et la *Deutsche Forschungsgemeinschaft*. Il a été inauguré lors du XXV^e Congrès International de Linguistique et de Philologie Romanes (Innsbruck, 2007). L'équipe du DÉRom, dirigée par Éva Buchi et Wolfgang Schweickard, est composée de 58 membres originaires de quinze pays, dont 55 linguistes romanistes, rédacteurs et/ou réviseurs, un informaticien et deux documentalistes. Le dictionnaire est publié sur internet ; il est en accès libre. Le DÉRom, qui comptait quinze membres au départ, a rencontré une grande adhésion dans la communauté des romanistes. Il a aussi été l'objet d'un âpre débat opposant les directeurs du dictionnaire à Alberto Várvaro (cf. Buchi et Schweickard 2011a ; 2011b ; Várvaro 2011a ; 2011b). Le point de départ de la nomenclature est la liste des 488 bases étymologiques communes (ou presque) à l'ensemble des langues romanes et recensées par Iancu Fischer (1969: 113). Le dictionnaire compte à ce jour 58 articles publiés⁴ et environ deux cents sont en cours de rédaction.

L'objectif du DÉRom est de rédiger un nouveau dictionnaire étymologique panroman (après le REW de Meyer-Lübke) en reconsidérant l'étymologie du lexique héréditaire commun à l'ensemble des idiomes romans, à la lumière des avancées de la linguistique historique romane. Les entrées sont constituées des étymons protoromans, reconstruits par la méthode de la grammaire comparée – reconstruction (cf. Chambon 2007; 2010; Dardel 2007; 2009) à partir des cognats romans («lexèmes reliés par un réseau de correspondances phoniques et sémiques», Buchi 2010: 44). Ils sont présentés en notation phonologique et portent l'astérisque, ex. */ka'ten-a/. Ce «dépayement épistémologique» (Chambon, à paraître a) ne rend pas plus difficile la consultation du dictionnaire : tous les articles peuvent être interrogés à partir des corrélats latins (*catena*) ou des entrées du REW (*catēna*), ou même des issues romanes (dacorum. *cătină*, frioul. *ciadène*, port. *cadeia*, par ex.).

La métalangue du DÉRom est le français. Les normes rédactionnelles, la bibliographie ainsi que d'autres outils nécessaires à la rédaction et à la révision des articles du DÉRom sont réunis dans un ouvrage intitulé *Livre bleu*, mis à jour une à deux fois par an dans sa version papier et de manière continue sous sa forme électronique. Conformément aux normes rédactionnelles, les articles du DÉRom comportent obligatoirement les parties suivantes:

(1) «Lemme» : signifiant, catégorie grammaticale et signifié de l'étymon protoroman. Le signifiant est présenté entre barres obliques et précédé d'un astérisque. Il comporte le signe d'accent d'intensité et utilise le trait d'union pour segmenter les constituants morphématiques. L'astérisque ne signifie pas, comme dans les dictionnaires traditionnels, que l'étymon n'est

⁴ */a'gost-u/; */ali-u/; */anim-a/; */ann-u/; */a'pril-e/; */a'pril-i-u/; */as'kolt-a /; */aud-i-/; */baβ-a/; */βad-u/; */'barb-a/₁; */'barb-a/₂; */'batt-e-/; */βi'n-aki-a/; */βin-u/; */'biβ-e/; */βindika-/; */'deke/; */'dent-e/; */'dorm-i-/; */'eder-a/; */'erβ-a/ ~ */'erb-a/; */es'kolt-a-/; */'eks-i-/; */'Φak-e-/; */'Φe'βr-ar-i-u/; */'Φen-u/ ~ */'Φen-u/; */'φili-u/; */ka'ball-a/; */ka'ball-u/; */'kad-e-/; */'karn-e/; */'karpin-u/; */'kasi-u/; */kas'tani-a/ ~ */kas'tni-a/; */ka'ten-a/; */'kresk-e-/; */'kul-u/; */la'brusk-a/ ~ */la'brosk-a/; */'laks-a-/; */'lakt-e/; */'laur-u/; */'lεβ-a-/; */'lök-u/; */'mai-u/; */'mart-i-u/; */'ment-e/; */'mönt-e/; */'most-u/; */'nap-u/; */'niβ-e/; */'pan-e/; */'part-e/; */'pönt-e/; */ro'tond-u/; */sa'gitt-a/; */'sal-e/; */'salβi-a/.

pas attesté, mais qu'il a été reconstruit à partir du témoignage des idiomes romans selon les principes de la grammaire comparée – reconstruction. Le signifié est présenté sous la forme d'une définition componentielle. Exemple : « */'baβ-a/ s.f. 'salive visqueuse qui s'échappe de la bouche d'une personne ou de la gueule d'un animal' ».

(2) « Matériaux » : contient l'ensemble des lexèmes appartenant aux différents idiomes romans et permettant de reconstruire l'étymon protoroman ; vingt idiomes – non seulement des langues standard mais aussi des dialectes, essentiels pour la reconstruction (cf. Andronache à paraître) – sont obligatoirement cités s'ils comportent des continuateurs de l'étymon ; ils sont énumérés successivement de l'est (dacoroum.) à l'ouest (port.) de la Romania⁵ ; pour chaque idiome sont donnés obligatoirement l'abréviation du glottonyme, le signifiant, la catégorie grammaticale, le signifié (s'il est différent de celui de l'idiome précédent), la date de première attestation et les références bibliographiques ainsi que la forme de la variante attestée si elle est différente de celle du signifiant. Exemple :

occit. *vinassa* « boisson alcoolisée obtenue par distillation de la râpe, marc de vin » (dp. 1442, Affre, RLaR 15, 15 ; Pansier 3 ; Pansier 5 ; Levy ; FEW 14, 479b ; ALF 1620 p 776, 787, 833 [« marc de raisin »] ; ALLOr 882, 887 [*bináso* « id. ; marc de vin »] ; MoutierDialectes [*vinasso* « lie de vin ; gros vin »]) (Delorme, s.v. */βi'n-aki-a/).

Les références sont prises dans la « bibliographie de consultation et de citation obligatoire » qui comporte 130 titres ainsi que dans d'autres ouvrages, si cela s'avère nécessaire ; les sigles renvoient, par un lien hypertexte, à la référence complète du document cité. Les matériaux peuvent comporter plusieurs subdivisions, d'ordre formel, morphologique, sémantique ou – dans une moindre mesure – syntaxique, lorsque « les parlers romans ont maintenu un lexème du protoroman sous plusieurs réalisations flexionnelles » ou encore « si la comparaison des cognats romans incite à penser que l'étymon protoroman connaissait plusieurs sens (LB : 55).

(3) « Commentaire » : « [il] explicite l'analyse des données réunies dans la section consacrée aux matériaux qui conduit à poser l'étymon cité dans l'entrée de l'article. [...] Il pose la question de l'existence d'un corrélat latin et y répond » (LB: 56). Le commentaire peut être très bref (*/ka'ten-a/) ou assez long (*/'ϕak-e-/) ;

(4) « Bibliographie » : « [elle] cite, par ordre chronologique, les publications [...] qui présentent un intérêt général (c'est-à-dire plus ou moins panroman) pour l'article » (LB: 60).

(5) « Signatures » : [elle] présente l'ensemble des personnes ayant contribué à l'élaboration de l'article, à la rédaction, à la révision ou par des « contributions ponctuelles » ;

(6) « Date de mise en ligne de cet article » : y sont données la date de la première version et celle de la « version actuelle » ;

(7) « Notes » : « Les notes règlent des points de détail appelant des remarques trop longues ou trop spécifiques pour être insérées dans le texte. En revanche [...], les évolutions phonétiques régulières ne sont jamais expliquées (leur explication se déduit indirectement des références données en bibliographie » (LB : 62).

2.2 La place du latin écrit dans le DÉRom

Faire le choix de présenter comme lemme de l'article du DÉRom non un lexème latin mais la forme phonologique protoromane reconstruite relève d'une démarche visant à privilégier la langue parlée par rapport à la langue écrite et caractérisée par le fait que l'étymon ainsi reconstitué n'est pas le point de départ de l'analyse étymologique mais l'aboutissement du travail de reconstruction (cf. Chauveau 2010: 1). Le latin écrit de l'Antiquité n'est donc pas utilisé dans la reconstruction de l'étymon protoroman. Cela ne signifie aucunement qu'il n'a

⁵ Dacoroumain, istroroumain, méglénoroumain, aroumain (aroum.), dalmate, istriote, italien, sarde, frioulan, ladin (lad.), romanche, français, francoprovençal (frpr.), occitan, gascon, catalan, espagnol, asturien, galicien, portugais. Est actuellement à l'étude l'introduction de l'aragonais et du sicilien, parmi les idiomes cités obligatoirement.

pas sa place dans ce dictionnaire. D'une part, comme nous l'avons vu, on peut interroger le lexème latin, s'il existe, à partir du menu et ceci de deux manières : sous les rubriques « Corrélat latins » et « Entrées du REW₃ ». D'autre part, la forme latine est toujours citée, lorsqu'elle existe, dans la partie « Commentaire » de l'article. En effet, dans la métalangue du DÉRom, les données du latin écrit sont proposées comme des corrélats des données protoromanes : « le modèle sous-jacent retenu est celui d'une diglossie – et d'une simultanée – entre le protoroman et le latin classique » (Buchi/Schweickard 2009 : 105-106). La partie « Commentaire » contient obligatoirement un paragraphe entier consacré à la comparaison de l'étymon protoroman avec son éventuel corrélat du latin écrit de l'Antiquité. « Le corrélat est défini comme le correspondant le plus proche de l'étymon protoroman reconstruit dans le latin écrit de l'Antiquité : (1) signifiant le plus proche, (2) catégorie grammaticale la plus proche, (3) signifié le plus proche » (LB 57). Les corrélats latins sont notés sous leur forme citationnelle (verbes à l'infinitif ; noms au nominatif singulier suivi de la désinence du génitif, etc.). En principe, les quantités vocaliques ne sont pas notées sauf, exceptionnellement, si l'intelligibilité du texte le requiert. Ainsi, s.v. */'kad-e-/ , on distingue *cadere*, connu durant toute l'Antiquité, de *cādere*, qui n'est attesté qu'à partir du 4^e siècle ap. J.C. Les corrélats latins sont datés. On considère que le latin écrit de l'Antiquité va jusqu'en 600. Les éventuelles attestations en latin tardif et médiéval (à partir de 600) ne sont pas mentionnées. Si la première attestation du corrélat latin est antérieure à 81 av. J.C. (Cicéron), on utilise la formule « Le corrélat du latin écrit, [...], est connu durant toute l'Antiquité », comme nous venons de le voir pour *cadere*. Si elle est postérieure à cette date, on cite l'auteur et la date du texte (ex. s.v. */'a'gost-u/) ou les dates de vie et de mort de l'auteur. Les articles dont les matériaux présentent des subdivisions traitent la question du corrélat pour chaque subdivision. Dans les cas où un étymon protoroman ou un étymon intermédiaire ne connaît pas de corrélat en latin écrit de l'Antiquité, on le précise explicitement. (cf. *infra*).

3 Les étymons reconstruits du DÉRom : la confrontation avec le latin écrit

Dans la partie qui suit, nous allons nous livrer à une comparaison, à travers quelques exemples, entre les étymons reconstruits du protoroman et leurs corrélats latins respectifs, en cherchant à mettre en évidence ce qui les sépare et en essayant de présenter la diversité de cas qui s'offrent à nous.

3.1 Les étymons protoromans ne présentant pas de corrélat dans le latin écrit de l'Antiquité

Certains étymons protoromans ne connaissent pas de corrélat dans le latin écrit de l'Antiquité. Deux situations se présentent. Soit l'étymon-lemme, c'est-à-dire celui qui chapeaute l'ensemble des issues romanes (et qui constitue « l'entrée » du dictionnaire), n'a pas de corrélat. Soit l'étymon-lemme a un corrélat mais il n'est pas l'étymon direct de tous les cognats romans, l'article contenant alors autant de subdivisions que d'étymons directs des cognats qu'il regroupe ; dans ce cas, certains étymons directs peuvent ne pas avoir de corrélat en latin écrit de l'Antiquité, tandis que d'autres en ont.

3.1.1 Les étymons-lemmes sans corrélat

Pour quatre articles, l'étymon-lemme ne connaît pas de corrélat : */'baβ-a/, */'a'pril-i-u/, */'barb-a₂/, /es'kolt-a-/. Le premier, */'baβ-a/, est un diaphasme, dans la mesure où c'est, à l'origine, un mot utilisé par les adultes pour s'adresser aux enfants, « avec les caractéristiques du langage enfantin : redoublement syllabique, alternance consonnes / voyelles, prépondérance des labiales (et des dentales) et de la voyelle /a/ » (cf. Gross, s.v. */'baβ-a/). */'baβ-a/) ne possède aucun corrélat dans le latin écrit de l'Antiquité. Toutefois, la littérature

latine témoigne de différents indices de son existence dans le langage populaire. Ainsi, chez Sénèque [1^{er} s. ap. J.C.], on trouve le nom de personne *Baba* qui, d'après l'OLD, a le sens de « name of a fool » ; des dérivés comme *babulus* s.m. « bavard » et *bavosus* adj. « stupide » sont attestés respectivement au 2^e s. ap. J.C. (chez Apulée) et *ca* 550 apr. J.C. (cf. LEI 4, 99) (cf. Gross, s.v. */'baβ-a/).

Les trois autres étymons sont nés d'articles rédigés précédemment, la reconstruction romane ayant conduit à l'impossibilité de remonter à un étymon unique et donc à la nécessité de rédiger deux articles distincts correspondant à deux unités lexicales distinctes. À côté du protoroman */a'pril-e/ dont le corrélat latin *aprilis*, -is est généralement donné comme étymon du français *avril* ainsi que de tous les lexèmes désignant ce mois dans les langues romanes, existait un protoroman */a'pril-i-u/, dérivé synonymique du précédent, qui a donné le français [a'vri:Λ], forme héréditaire sûre, bien implantée dans la langue ancienne (attestée de *ca* 1100 à 1881), mais aussi les cognats *varīyi* en vénétien méridional, *april* dans le dialecte des Marches septentrionales et *pril'* en méglénoroumain (le seul de ces quatre idiomes à ne pas connaître de descendant de */a'pril-e/) (cf. Celac, s.v. */a'pril-i-u/). Là encore, le latin écrit de l'Antiquité ne connaît pas de corrélat de ce lexème. Toutefois, on trouve en latin médiéval les formes *aprilus* (attesté en 845) et *aprelius* (apr. 872). D'où il ressort qu'il n'est pas sûr que l'étymon de fr. *avril* soit */a'pril-e/ : « on pourrait y voir le représentant de */a'pril-i-u/ avec une réinterprétation en /-l/ du graphème < -l > pour /-l̥/ » (cf. Celac, s.v. */a'pril-e/). L'étymon */'barb-a/₂ s. m. « oncle » n'est attesté que dans un petit nombre d'idiomes (dalmate, istriote, italien septentrional, frioulan, ladin et romanche). Schmidt, s.v. */'barb-a/₂ s.m., considère cet étymon comme le résultat d'une masculinisation de */'barb-a/₁ s.f. « barbe », intervenue en protoroman régional ; autrement dit, il s'agirait d'une métonymie, la barbe étant conçue comme l'expression de l'autorité et de la virilité. Il faut attendre 643 pour voir apparaître en latin tardif une forme *barba*, -anis s.m. « vieil homme ; oncle » qui représente un reflet indirect des lexèmes romans. Quant à l'étymon */es'kolt-a-/ , il est issu, par greffe préfixale /es-/ , de */as'kolt-a-/ (cf. Chauveau in FEW 25 : p. 1060a) qu'il a complètement évincé dans certains idiomes regroupés dans une aire cohérente occitane, gasconne et catalane dès l'époque pré littéraire. Il s'agit donc d'une innovation, née très probablement dans la Narbonnaise (cf. Schmidt, s.v. */es'kolt-a-/). La forme moderne du portugais *escutar* est issue de cet étymon tandis que l'ancien portugais *ascoitar*, *ascuitar* a pour étymon */as'kolt-a-/.

3.1.2 Les étymons non-lemmes sans corrélat

Les articles comportant des subdivisions sont ceux pour lesquels la reconstruction n'a pas permis de remonter directement à un étymon unique. On aboutit alors à un protoroman déjà différencié, et donc à deux ou plusieurs étymons distincts, chacun d'entre eux étant l'étymon direct d'un ensemble de cognats et ayant lui-même pour étymon direct l'étymon-lemme. Chacun de ces étymons directs ne possède pas nécessairement de corrélat du latin écrit de l'Antiquité.

28 articles présentent des subdivisions qui peuvent être d'ordre formel – affectant la forme phonique – (10 articles), morphologique (12) sémantique (10) et, dans une moindre mesure syntaxique (2)⁶. Un même article peut aussi combiner plusieurs types de subdivisions. L'article */'eder-a/ « lierre » présente trois subdivisions phoniques : I. */'eder-a/ ; II. */'eler-a/ ; III. */'elen-a/. Seul le premier type possède un corrélat *hedera*, -ae, connu durant toute l'Antiquité (dp. Laberius [*105 – † 43]). Les types II. et III., plus récents, sans nul doute diastratiquement – et probablement diaphasiquement – marqués, n'ont pas laissé de traces

⁶ Cet aspect n'apparaît que marginalement et lié à un aspect sémantique, à propos de la valence du verbe protoroman */'kresk-e-/ qui peut être intransitif, avec le sens de « croître », ou transitif, avec celui d'« accroître ».

dans le code écrit. L'article */põnt-e/ a été subdivisé selon trois types morphologiques correspondant à des changements de genre : I. Masculin originel, conservé par le sarde ; II. Féminin innové tardivement dans des aires latérales et dans des aires isolées ; III. Masculin restauré, innovant, sur une vaste aire compacte de la Romania centrale. Si le corrélat du masculin originel, *pons, -tis*, est attesté durant toute l'Antiquité, le féminin ne l'est que dans un toponyme du 3^e siècle ap. J.C. L'article */'anim-a/ connaît deux divisions sémantiques : I. sens abstrait « âme » ; II. sens concret « cœur ; estomac ». Seul le sens abstrait connaît un corrélat en latin écrit *anima, -ae* s.f. « bouffée d'air ; âme ; vie », connu durant toute l'Antiquité (dp. Nénius [3^e s. av. JC.]), tandis que le latin écrit de l'Antiquité ne connaît pas de corrélat de II.

3.2 Ce que le protoroman reconstruit dit du latin oral

Nous nous attacherons à présent aux unités lexicales protoromanes qui connaissent un corrélat dans le latin écrit de l'Antiquité pour montrer comment le protoroman peut rendre compte de phénomènes oraux inaccessibles par la seule documentation latine, autrement dit à l'immédiat communicatif au sens de Koch et Oesterreicher (2008). Il s'agit donc de mettre ici en évidence ce qui distingue le code oral et le code écrit du diasystème latin, c'est-à-dire la variation diamésique du latin.

3.2.1 Du point de vue phonologique

3.2.1.1 Les voyelles /ɪ/ et /ʊ/

Le système vocalique du protoroman ne connaissait pas d'opposition de quantité (nous avons vu que Väänänen 1959 : 18 le montre pour le latin des inscriptions pompéiennes; cf. aussi Straka 1956 : 255) et qu'il comportait neuf voyelles et non les sept voyelles traditionnelles du latin *vulgaire*. Ces deux phonèmes vocaliques spécifiques du protoroman sont /ɪ/ et /ʊ/. Prenons, pour illustrer ces deux phonèmes, les exemples des verbes *boire* et *écouter* et voyons comment se comportent leurs cognats dans cinq autres idiomes romans, le roumain, l'italien, le sarde, l'espagnol et le portugais.

Latin REW ₃	Roumain	Italien	Sarde	Français	Espagnol	Portugais	Phon. voc.	Proto-roman
<i>bībēre</i>	<i>bea</i>	<i>bevere</i>	<i>bībere</i>	<i>boire</i> afr. <i>boivre</i>	<i>beber</i>	<i>beber</i>	/ɪ/	*/'bɪβ-e-/
<i>auscūltāre</i> (« pop. » * <i>ascūltāre</i>)	<i>asculta</i>	<i>ascoltare</i>	<i>askultare</i>	afr. ↑ <i>ascouter</i> ⁷	aesp. <i>ascuchar</i>	aport. <i>ascoitar</i> /ascuitar	/ʊ/	*/askʊl't-a-/

Partir du latin écrit, comme les dictionnaires qui notent *bībēre* et *auscūltāre* (ou « pop. » **ascūltāre*), ne nous dit rien sur la manière dont se prononçaient ces voyelles dans le latin oral à l'époque du changement d'ère, par exemple. Si l'on suit la doxa sur l'évolution menant du latin classique au latin dit « vulgaire », /ĩ/ > /e/ et /ũ/ > /o/. Or, pour le verbe *boire*, le sarde *bībere* (mais aussi l'istriote *bívi*), ne présente pas la voyelle /e/ mais la voyelle /i/⁷. Il s'ensuit que la voyelle de l'ancêtre commun de ces six idiomes ne peut avoir qu'un timbre intermédiaire entre /e/ et /i/ : c'est la voyelle /ɪ/. D'où le protoroman reconstruit */'bɪβ-e-/. De même, pour le verbe *écouter*, l'italien (mais aussi le frioulan *scoltâ*) possède la voyelle /o/, tandis que le roumain, le sarde, le français et l'espagnol présentent la voyelle /u/. L'ancêtre

⁷ Nous laissons de côté, naturellement, les évolutions intraromanes, plus tardives, qui aboutissent, par exemple, au français /wa/ (début du 14^e s., cf. Bourciez 1967 :72).

commun protoroman est donc un intermédiaire entre /u/ et /o/ : c'est la voyelle /ʊ/, et la forme protoromane reconstruite est */askol't-a-/.

Il arrive toutefois que, du point de vue phonologique, la reconstruction des cognats romans de l'unité lexicale ne permette pas de remonter à un phonème vocalique unique et que l'on ne puisse pas trancher entre deux voyelles qui s'opposent, par exemple, par leur degré d'aperture. Dans ce cas, on aboutit à deux résultats divergents. C'est le cas de /e/ ~ /ɛ/ dans l'article à double lemme */'ϕen-u/ ~ */'ϕɛn-u/ du DÉRom : l'étymon commun est antérieur au protoroman, il serait à rechercher dans un latin archaïque */'ϕain-u/ ; le type I (*/'ϕen-u/) couvre la presque totalité de la Romania⁸, tandis que le type II se réduit à une aire s'étendant du ligurien au sarde et comportant le toscan, et donc l'italien standard (cf. Reinhardt, s.v.). On ne peut donc considérer la forme latine *fēnum* donnée par le REW₃ (n° 3247) comme l'ancêtre commun à l'ensemble de ces idiomes. Un autre exemple est celui de l'opposition /a/ ~ /ɪ/ dans le double lemme */kas'tani-a/ ~ */kas'tni-a/ où le type I connaît des issues dans les parlers de toutes les branches romanes, à l'exception du dalmate, tandis que le type II est attesté en ladin, frioulan, italien septentrional et dans quelques variétés centro-méridionales de l'italien⁹. Cette fois-ci, le REW₃ (n° 1742) donne bien deux étymons latins, *castanea* et *castinea*, mais nous avons vu que le < i > du second étymon correspond en réalité à un /i/ protoroman.

3.2.1.2 Le cas du proto-phonème consonantique /ϕ/

Le tableau suivant montre les cognats romans du lexème signifiant « fils » en aroumain, en italien, en français, en gascon et en espagnol.

Latin REW ₃	Aroumain	Italien	Français	Gascon	Espagnol	Phon. cons.	Proto-roman
<i>filius</i>	<i>hil'u</i>	<i>figlio</i>	<i>fils</i>	<i>hilh</i>	<i>hijo</i>	/ϕ/	*/'ϕili-u/

Si, dans la plupart des langues romanes – représentées ici par le français et l'italien –, le graphème < f > du latin correspond à la fricative bilabiale /f/, nous voyons que ce n'est pas le cas de l'aroumain, du gascon ni de l'espagnol, où il correspond, respectivement aux phonèmes /ç/, /h/ et ø. Dans ces idiomes, le phonème consonantique initial ne peut pas remonter au supposé /f/ du latin *filius*. Alors comment se prononçait ce phonème en latin ? Maniet (1975 : 26- 27) montre que le graphème < f > que l'on trouve dans les textes latins correspondait dans un premier temps à une articulation bilabiale. Il en donne pour preuve la présence, dans des graphies archaïques, du graphème < m > devant < f >. Meyer-Lübke (1890-1906, 1 : 353-354) s'est intéressé à la valeur phonique du < f > espagnol :

Les plus anciens monuments de la littérature espagnole écrivent encore presque sans exception *f* [...]. Mais des métathèses orthographiques et des cas isolés où *h* apparaît prouvent que déjà à cette époque le signe *f* ne représentait pas une labio-dentale mais tout au plus une spirante sourde bilabiale et peut-être seulement la simple aspiration.¹⁰

C'est pourquoi le DÉRom postule un proto-phonème bilabial /ϕ/. Ce postulat a suscité la critique d'Alberto Vârvaro (2011b : 627), à laquelle les directeurs du DÉRom ont répondu :

Ce qui est symptomatique, c'est que personne n'a jamais exigé des explications concernant la valeur labiodentale ou bilabiale de l'unité phonique que les dictionnaires étymologiques notent par le graphème < f > (ainsi *facère* dans le REW) – et que 99%

⁸ Les formes régulières et/ou héréditaires correspondant à ce lexème ne se sont pas toujours maintenues dans les idiomes contemporains : la forme régulière du fr. *fein* a été évincée par *foin*, qui est aussi héréditaire mais emprunté aux dialectes de l'Est ; en portugais, l'issue héréditaire *feo* a été évincée par l'emprunt savant *feno*.

⁹ D'après l'auteur de l'article, cette variation vocalique « ne fait que refléter le traitement fluctuant de gr[ec] /ã/, passé par apophonie à /ɪ/ dans certaines variétés de la langue emprunteuse » (cf. Medori, s.v.).

¹⁰ Cf. aussi Cano 2004 : 314.

des lecteurs auront oralisée en /f/ : le système graphique du latin ne permettant pas de distinguer /ϕ/ et /f/, la lecture /f/ ne va en réalité pas plus de soi que /ϕ/. (Buchi et Schweickard 2011b : 630)

L'étymon commun à tous les cognats romans signifiant « fils » n'est donc pas le latin écrit *filius* mais une forme orale à laquelle nous avons accès par la reconstruction : */ϕili-u/.

3.2.2 Du point de vue morphologique

Nous avons déjà évoqué (cf. *supra* 3.1.1) le cas de */as'kult-a-/ et */es'kult-a-/ qui sont deux bases étymologiques distinctes, la seconde étant formée par greffe préfixale sur la première. Si */as'kult-a-/ est représenté dans toutes les branches romanes à l'exception du dalmate, de l'occitan, du gascon et du catalan (surtout dans des formes romanes médiévales mais aussi dans des formes contemporaines telles dacoroum. *asculta*, ast. *ascuchar*), */es'kult-a-/ n'est continué que sur une aire qui, bien qu'assez vaste, est plus réduite (fr., frpr, occit., gasc., cat., esp., ast., gal. et port.). Or les dictionnaires étymologiques idioromans ne proposent généralement qu'un seul étymon, *auscultare* (cf. par ex., le DELP₃). Et si le REW₃ (n° 802) propose, à côté du lemme latin *auscūltāre* l'étymon marqué d'un astérisque – c'est-à-dire non attesté dans le latin écrit de l'Antiquité – **ascūltāre*, il considère ce dernier comme base étymologique unique de tous les cognats romans, sans prendre en compte les différences morphologiques existant entre eux.

Un autre exemple est celui de l'ancêtre commun des cognats romans désignant le verbe *faire*. Les dictionnaires étymologiques, à commencer par le REW₃ (n° 3128), donnent comme étymon commun le latin *facĕre*. Outre la question de la nature du phonème initial évoquée plus haut, se pose un problème morphologique : comment la forme latine *facĕre* peut-elle être considérée comme l'étymon direct d'issues du type du dalmate *fur*, de l'istriote *fa*, de l'italien *fare* ou de l'ancien espagnol et de l'ancien catalan *far* (mais aussi dans les formes du futur en galicien et en portugais), pour n'en citer que quelques-unes ? L'article */ϕak-e-/ du DÉRom distingue deux bases étymologiques, un type originel */ϕak-e-re/ et un type évolué */ϕ-a-re/. Le second type, syncopé, est plus récent (2^e moitié ou fin du 3^e s.) et issu du premier dont il est une variante combinatoire, la syncope de /-ke-/ s'expliquant par l'« usure due à la grande fréquence du verbe, et cela notamment dans la position proclitique du type le plus ancien du futur roman */ϕ-a-re-'aβ-e-/ » (Buchi, s.v. */ϕak-e-/).

3.2.3 Du point de vue sémantique

Dix articles du dictionnaire sur 58 connaissent des subdivisions sémantiques. Cela signifie que l'étymon protoroman peut avoir plusieurs signifiés, tout en conservant (ou pas) le même signifiant. Tous ces signifiés ne sont pas nécessairement attestés dans le latin écrit de l'Antiquité et ne sont donc que très rarement répertoriés par les dictionnaires étymologiques. Nous avons vu plus haut (cf. 3.1.2) le cas de */anim-a/ (corrélât : *anima*, -ae) qui a le sens d'« âme » mais aussi celui de « cœur, estomac ». Donnons deux autres exemples. */barb-a/₁ « barbe », connu dans tous les parlars de la Romania et durant toute l'Antiquité (corrélât : *barba*, -ae), a aussi le sens de « menton » dans treize idiomes (dacoroum., istroroum., dalm., istriot., it., sard., frioul., occit., cat., esp., ast., gal. et port.) ; ce second sens n'est attesté dans les textes latins qu'à partir du 3^e s. ap. J.C. (cf. Schmidt, s.v.). */βi'n-aki-a/ signifie « marc de raisin » ; c'est le sens majoritaire, connu en latin (*uinacea/uinacia*, -ae s.f.) depuis Varron (*116 – † 27) ; mais le frioulan et le romanche connaissent une issue de cet étymon avec le sens d'« oseille ; fruit de l'épine-vinette », qui n'a pas de corrélât latin (cf. Delorme, s.v.).

4 Quelques aspects de la variation interne du protoroman

La méthode de la grammaire comparée-reconstruction à l'œuvre dans le DÉRom permet d'avoir accès à la variation interne du latin parlé, à travers les formes protoromanes reconstruites. Nous distinguerons ici les variations diachronique, diatopique, diastratique et diaphasique.

4.1 La variation diachronique

Que sait-on du protoroman du point de vue diachronique ? Au sens strict, le terme *protoroman* désigne, nous l'avons vu, l'ancêtre commun à *tous* les idiomes romans. Mais cette langue a évolué dans le temps et, ce faisant, elle a perdu de son caractère panroman. Buchi et Schweickard (à paraître) distinguent, à partir de la reconstruction des étymons du DÉRom, quatre phases du protoroman :

1. Le « protoroman le plus ancien » : il date d'avant la seconde moitié du 2^e siècle ap. J.C., c'est-à-dire avant la séparation du protosarde (Straka 1956 : 256). Les étymons du DÉRom datant de cette période sont tous panromans¹¹ ou ne manquent que dans une seule branche romane, comme */'most-u/ absent du dalmate ou */'anim-a/, inconnu du frioulan.

2. Le « protoroman postérieur à la séparation du sarde » (mais antérieur à la séparation du protoroumain, qui a lieu vers la fin [Straka 1956 : 258] ou la seconde moitié [Rosetti 1986 : 184] du 3^e siècle. ap. J.C.). Les étymons relevant de cette période n'ont pas de continuateur en sarde mais ils sont représentés dans la branche roumaine. C'est le cas, par exemple de */'ɛder-a/ ou */'salβi-a/, ce dernier étant aussi absent du dalmate.

3. Le « protoroman postérieur à la séparation du roumain » (mais antérieur au 5^e ou au 6^e siècles, ce *terminus ad quem* correspondant à la séparation du français [début du 5^e s.], du gascon [au plus tard *ca* 600], du galloitalien [autour de 600] et du francoprovençal [probablement acquise au plus tard à la fin du 6^e s.], cf. LB : 15). C'est le cas de */'ϕ-a-re/, variante combinatoire de */'ϕak-e-re/ en position proclitique, à l'origine des futurs à base analytique (Buchi et Schweickard à paraître), qui n'a pas de continuateur en sarde ni dans les idiomes roumains.

4. Le « protoroman postérieur à la séparation du français, du gascon, du galloitalien et du francoprovençal », c'est-à-dire après le 5^e/6^e siècle. Nous avons le cas de l'étymon */'es'kult-a-/, dérivé par greffe préfixale (de */'es-/ de */'as'kult-a-/ qui n'a des continuateurs qu'en français, francoprovençal, occitan, gascon, catalan, espagnol, asturien et galégo-portugais.

De ces quatre phases, seule la première, « protoroman le plus ancien », correspond à la définition au sens strict du terme *protoroman*. Tout en conservant un aspect largement diaroman, les autres phases couvrent des aires de plus en plus réduites.

4.2 Les variétés diatopiques

Les étymons reconstruits du DÉRom n'ont pas tous eu le même degré de diffusion dans la Romania. Si un grand nombre d'entre eux sont panromans, d'autres sont représentés dans toutes les branches romanes ou presque mais pas dans certains parlars, d'autres encore n'ont des issues que dans un nombre plus ou moins réduit de parlars romans. Nous allons voir comment se répartissent géographiquement ces étymons.

4.2.1 Les étymons panromans

Sur les 58 étymons-lemmes du DÉRom, 25 ont des issues dans tous les parlars romans sans exception : */'ali-u/; */'ann-u/; */'barb-a₁/; */'batt-e-/; */'biβ-e/; */'βin-u/; */'dɛke/; */'dɔrm-i-/; */'ɛrβ-a/ ~ */'ɛrb-a/; */'ϕak-e-/; */'ϕɛn-u/ ~ */'ϕɛn-u/; */'ϕili-u/; */'karn-e/;

¹¹ Cf. *infra* sous 4.2.1.

*/'kresk-e-/; */'kul-u/; */'laks-a-/; */'lakt-e/; */'lɛβ-a-/; */'lɔk-u/; */'mai-u/; */'mɔnt-e/; */'nɪβ-e/; */'pan-e/; */'part-e/; */'pɔnt-e/. Ce sont les étymons appartenant à la strate la plus ancienne du protoroman, datant d'avant la séparation du protoroman de Sardaigne.

4.2.2 Les étymons représentés dans toutes les branches romanes ou presque

Certains étymons présentent des continuateurs dans toutes les branches romanes mais pas dans tous les parlars. Quatre étymons sont dans ce cas : */'ka'βall-u/ (manque en istriote)¹² ; */'ka'ten-a/ (manque en istroroumain, méglénoroumain et en istriote; */'ment-e/ (absent de l'istroroumain, du frioulan et du français [sauf comme suffixe formateur des adverbes de manière en *-ment*]) et */'sa'gitt-a/ (inconnu de l'istroroumain et du méglénoroumain). C'est donc essentiellement dans les dialectes roumains et en istriote que ces étymons ne sont pas représentés. 21 étymons sont présents dans presque toutes les branches romanes. Les branches qui manquent le plus souvent à l'appel sont le dalmate (pour quinze étymons), le ladin (pour sept) et le sarde (pour six). L'absence de continuateurs en sarde implique, essentiellement, que les étymons appartiennent à un protoroman postérieur à la séparation du sarde. L'absence en dalmate et en ladin n'est pas significative, étant donné l'aspect extrêmement lacunaire des sources dont nous disposons pour ces deux idiomes.

4.2.3 Les étymons continués dans certains parlars seulement

Enfin, sept étymons protoromains, ne sont continués que dans un petit nombre de parlars romans. Certains sont dispersés à travers la Romania, comme */'a'pril-i-u/, représenté en méglénoroumain, dans deux parlars italiens septentrionaux et en français. D'autres couvrent des aires plus compactes, comme */'barb-a/₂ dont l'extension (dalm., istriot., it. septentrional, frioul., lad. et romanch.) est comprise dans l'aire panromane, de */'barb-a/₁, ou encore */'es'kult-a-/ (fr., frpr., occit., gasc., cat., esp., ast., gal. et port.) et */'ka'βall-a/ (dalm., it., lad. et romanch.)¹³. On ne s'étonnera pas de retrouver trois des quatre étymons ne comportant pas de corrélat dans le latin écrit de l'Antiquité, auxquels on pourrait ajouter */'ka'βall-a/ dont le corrélat latin n'est attesté qu'au 6^e siècle (cf. Cano, *s.v.*). Quant à */'laur-u/ et */'nap-u/, ils ont tous deux été concurrencés ou évincés par d'autres lexèmes, le premier par le dérivé */'lau'r-ari-u/ (cf. Reinhardt, *s.v.*) et le second peut-être par */'rap-u/ et */'rap-a/ (cf. Delorme, *s.v.*).

4.3 Les variétés diastratiques

Nous avons vu que l'étymon protoroman */'anim-a/ a deux significés : I. « âme » ; II. « cœur ; estomac ». Ce second sens ne connaissant pas de corrélat dans le latin écrit semble n'avoir donc été utilisé qu'à l'oral. De plus, il ne connaît d'issues qu'en roumain (dans ses quatre variétés) et en sarde, qui sont les deux premières branches à s'être séparées du tronc commun roman. Nous suivons Buchi et Schweickard (à paraître, 6) qui pensent que le sens concret devait être employé dans des variétés basilectales tandis que l'acrolecte ne connaissait que le sens abstrait.

¹² Le statut exact de l'istriote n'est pas clair. Provisoirement, dans DÉRom, il n'est pas considéré comme une branche – ou plutôt comme un rameau, au sens où les langues romanes sont déjà une sous-branche de la branche italique (cf. LB : 58). Toutefois, dans son étude du vocalisme accentué dans les parlars istriotes, Jean-Pierre Chambon conclut que « l'istriote/istroroman constitue un rameau indépendant de la branche romane, au sens où cet idiome a connu des innovations qui lui sont propres (abaissement du vocalisme) et non au sens où il aurait refusé ensuite tout changement en commun avec d'autres groupes de parlars » (Chambon 2011 : 299-300).

¹³ Sard. *kaddina* s.f. « caprice », esp. *caballa* « maquereau », gal. *cabala* et port. *cavala* « id. » sont des formations idioromanes sur les lexèmes désignant le mâle et non des issues héréditaires de */'ka'βall-a/ (cf. Cano, *s.v.*).

L'article */laks-a-/ présente deux étymons protoromans distincts correspondant à deux variantes diastratiques, l'une, */laks-a-/, majoritaire et acrolectale, l'autre, */laks-i-a-/, minoritaire et basilectale, qui connaît des issues en italien septentrional, en toscan, en ladin et en romanche. Voici ce qui en est dit dans le commentaire de l'article :

Ce type évolutif s'analyse avec profit [...] comme issu de I. par ajout de l'interfixe postradical-préfixif /-i-/, diastratiquement marqué : sans contenu sémantique dénotatif, cet interfixe apporte une connotation plus basilectale aux lexèmes qu'il frappe, par opposition à leurs correspondants non interfixés, qui relèvent davantage de variétés acrolectales. Par la suite, les idiomes issus des branches connaissant les deux variantes ont sélectionné soit l'une, soit l'autre, de sorte que toutes les deux ont perdu leur connotation diastratique originelle (Florescu, s.v.).

Un autre exemple encore est celui de l'article */kad-e-/ qui présente une forme acrolectale */kad-e-re/, et une forme basilectale */ka'd-e-re/ (cf. Buchi, s.v. ; 2010b : 3).

4.4 Les variétés diaphasiques

Si */baβ-a/ est clairement un diaphasisme (utilisé quand on s'adresse à des enfants), Jan Reinhardt, auteur des articles */'ɛder-a/ et */'ϕen-u/ ~ */'ϕen-u/, se pose aussi la question de la diaphasie pour les variantes */'ɛler-a/ et */'ɛlen-a/ du premier, non attestées en latin écrit, ainsi que pour la variation interne du second.

5 Conclusion

La méthode mise en œuvre dans le DÉRom, celle de la grammaire comparée-reconstruction aboutit à des résultats déterminants en matière d'étymologie romane sur le rôle et la place qu'y occupent le latin écrit et le latin oral, en remettant en cause plusieurs idées communément acceptées. Elle montre que les formes du latin classique présentées dans les dictionnaires ne peuvent en aucun cas être les étymons directs des lexèmes romans qui leur sont rattachés et que c'est dans des formes reconstruites en protoroman que l'on peut trouver l'ancêtre commun à l'ensemble des parlers romans. Elle démontre que le système phonologique à sept voyelles qui était jusque là considéré comme celui du latin dit *vulgaire* est insuffisant pour rendre compte du système vocalique de l'ancêtre commun, qui est constitué de neuf et non de sept voyelles. L'analyse des articles du DÉRom nous a permis d'accéder à certaines caractéristiques de la variation du latin parlé à différentes époques et dans différentes aires de la Romania, dont nous serions privés si nous nous contentions du recours habituel aux seuls textes écrits. Les articles */a'pril-i-u/, */baβ-a/, */barb-a/₂, */es'kult-a-/ montrent que ces étymons protoromans n'avaient pas de corrélat dans le latin écrit de l'Antiquité. */'ɛder-a/, */'pɔnt-e/ et */'anim-a/ mettent en lumière des variantes, respectivement, formelles, morphologiques et sémantiques inconnues du latin écrit. Du point de vue phonique */as'kult-a/ et */'bɪβ-e/ donnent des indications sur la manière dont se prononçaient en latin parlé les deux phonèmes vocaliques que le latin écrit représente par < ũ > et < ĩ >, tandis que */'ϕili-u/ explique le phonétisme de la consonne initiale dans les issues aroumaine, gasconne et espagnole, ce que ne peut en aucun cas faire le graphème < f >. Les articles */'ϕen-u/ ~ */'ϕen-u/ et */kas'tani-a/ ~ */kas'tmi-a/ mettent en évidence des protoromans régionaux couvrant des aires cohérentes. Du point de vue de la généalogie des langues romanes, l'étude des articles du DÉRom donnent l'image d'un protoroman qui s'est différencié en quatre stades successifs allant d'avant la seconde moitié du 2^e siècle à après le 6^e siècle. Enfin, ont été mis en évidence, à partir d'un choix d'étymons du DÉRom l'existence et la distribution de variétés diastratiques et diaphasiques du latin parlé. Certes, toutes ces avancées sur la connaissance du latin oral ne donnent qu'une image très partielle de ce

qu'était cet ensemble de variétés du diasystème latin et il n'est pas sûr que l'on pourra un jour appréhender cette langue multiforme dans son intégralité. Alors laissons le dernier mot à Graziado Isaia Ascoli qui dit en exergue de la *Grammaire des langues romanes* de Meyer-Lübke: « Ogni nuovo avanzamento ridonda in nuovo onore dei maestri che ci hanno aperto e spianato la via ardua e buona ».

Références bibliographiques

Dictionnaires

- DCECH = COROMINAS, Joan et PASCUAL, José Antonio, 1980-1991 : *Diccionario crítico etimológico castellano e hispánico*, 6 volumes, Madrid, Gredos.
- DELP₃ = MACHADO, José Pedro (1977³ [1952¹]) : *Dicionário etimológico da língua portuguesa*, 5 volumes, Lisbonne, Horizonte.
- DER = CIORĂNESCU, Alexandru, 2005 : *Dictionarul etimologic al limbii române*, Bucarest, Saeculum I. O.
- DÉRom = BUCHI, Éva et SCHWEICKARD, Wolfgang (dir.), 2008- : *Dictionnaire Étymologique Roman (DÉRom)*. Nancy, ATILF (<http://www.atilf.fr/DERom>).
- FEW = WARTBURG, Walther von *et al.* (1922–2002) : *Französisches Etymologisches Wörterbuch. Eine Darstellung des galloromanischen sprachschatzes*, 25 volumes, Bonn / Heidelberg / Berlin / Bâle, Klopp / Winter / Teubner / Zbinden.
- GAFFIOT, Félix et FLOBERT, Pierre, 2001 : *Le Grand Gaffiot, dictionnaire latin-français*, Paris, Hachette.
- LEI = Pfister (Max)/Schweickard (Wolfgang) (dir.), 1979- . *Lessico Etimologico Italiano*, Wiesbaden, Reichert.
- OLD = GLARE, P. G. W. (éd.), 1968–1982 : *Oxford Latin Dictionary*, Oxford, Clarendon.
- REW₃ = MEYER-LÜBKE, Wilhelm, 1930–1935³ [1911–1920¹] : *Romanisches Etymologisches Wörterbuch*, Heidelberg, Winter.
- TLF = TLF = IMBS, Paul, et QUEMADA, Bernard (dir.), 1971–1994 : *Trésor de la langue française. Dictionnaire de la langue du XIX^e et du XX^e siècle (1789–1960)*, 16 volumes, Paris, Éditions du CNRS/Gallimard, site internet : <http://atilf.atilf.fr/tlf.htm>.
- TLL = 1900- . *Thesaurus Linguae Latinae*, Leipzig, Teubner.

Ouvrages et articles

- ANDRONACHE, Marta, à paraître : « Le statut des langues romanes standardisées contemporaines dans le DÉRom », in Casanova, Emili *et al.* (éd.), *Actes del 26^e Congrés Internacional de Lingüística i Filologia Romàniques (València 2010)*, Berlin/NewYork, De Gruyter.
- BENARROCH, Myriam, à paraître a : « L'apport du DÉRom à l'étymologie portugaise », in : Casanova, Emili *et al.* (éd.) : *Actes del 26^e Congrés Internacional de Lingüística i Filologia Romàniques (València 2010)*, Berlin/New York, De Gruyter.
- BENARROCH, Myriam, à paraître b : « L'étymologie du lexique héréditaire : en quoi l'étymologie panromane est-elle plus puissante que l'étymologie idioromane ? L'exemple du DÉRom », in Chabrolle-Cerretini, Anne-Marie (dir.), *Actes du colloque international « Romania : réalité(s) et concepts »*, Nancy, 6-7 octobre 2011, Paris, Lambert Lucas.
- BOURCIEZ, Édouard et BOURCIEZ, Jean (1967) : *Phonétique française*, Paris, Klincksieck.
- BUCHI, Éva, 2010a : « Pourquoi la linguistique romane n'est pas soluble en linguistiques idioromanes. Le témoignage du *Dictionnaire Étymologique Roman (DÉRom)* », in C. Alén Garabato *et al.* (éd.), *Quelle linguistique romane au XXI^e siècle ?*, Paris, L'Harmattan, p. 43-60.
- BUCHI, Éva, 2010b : « Where Caesar's Latin does not belong : a comparative grammar based approach to Romance etymology », in : Brewer, Charlotte (éd.), *Selected Proceedings of the Fifth International Conference on Historical Lexicography and Lexicology held at St Anne's College, Oxford, 16-18 June 2010*, Oxford, Oxford University Research Archive. Version électronique (<http://ora.ox.ac.uk/objects/uuid%3A237856e6-a327-448b-898c-cb1860766e59>), p. 1-7.
- BUCHI, Éva, à paraître : « Cent ans après Meyer-Lübke : le *Dictionnaire Étymologique Roman (DÉRom)* en tant que tentative d'arrimage de l'étymologie romane à la linguistique générale », in : Casanova, Emili *et al.* (éd.) : *Actes del 26^e Congrés Internacional de Lingüística i Filologia Romàniques (València 2010)*, Berlin/New York, de Gruyter.
- BUCHI, Éva, CHAUVEAU, Jean-Paul, GOUVERT, Xavier et GREUB, Yan, 2010 : « Quand la linguistique française ne saurait que se faire romane : du neuf dans le traitement étymologique du lexique héréditaire », in : Neveu, Franck *et al.* (éd.), *Congrés Mondial de Linguistique Française – CMLF 2010*, Paris, Institut de Linguistique Française, publication électronique (<http://dx.doi.org/10.1051/cmlf/2010025>), p. 111-123.
- BUCHI, Éva / SCHWEICKARD, Wolfgang (2009) : « Romanistique et étymologie du fonds lexical héréditaire : du REW au DÉRom (*Dictionnaire Étymologique Roman*) », in : Alén Garabato, Carmen *et al.* (éd.) : *La Romanistique dans tous ses états*, Paris, L'Harmattan, 97-110.
- BUCHI, Éva et SCHWEICKARD, Wolfgang, 2011a : « Sept malentendus dans la perception du DÉRom par Alberto Vårvaro », *Revue de linguistique romane* 75, 305-312.

- BUCHI, Éva et SCHWEICKARD, Wolfgang, 2011b: « Ce qui oppose vraiment deux conceptions de l'étymologie romane. Réponse à Alberto Vårvaro et contribution à un débat méthodologique en cours », *Revue de linguistique romane* 75, p. 628-635.
- BUCHI, Éva et SCHWEICKARD, Wolfgang, à paraître : « Per un'etimologia romanza saldamente ancorata alla linguistica variazionale : riflessioni fondate sull'esperienza del DÉRom (*Dictionnaire Étymologique Roman*) », in : Boutier, Marie-Guy, Hadermann, Pascale & Van Acker, Marieke (éd.), *Variation et changement en langue et en discours*, Helsinki, Société Néophilologique.
- CANO, Rafael, 2004 : *História de la lengua española*, Barcelona, Ariel.
- CELAC, Victor et BUCHI, Éva (à paraître) : « Étymologie-origine et étymologie-histoire dans le DÉRom (*Dictionnaire Étymologique Roman*) : coup de projecteur sur quelques trouvailles du domaine roumain », in Overbeck, Anja, Schweickard, Wolfgang & Völker, Harald (éd.), « *Alba pratalia araba...* ». *Romanistische Studien zum 65. Geburtstag von Günter Holtus*, Berlin/New York, de Gruyter.
- CHAMBON, Jean-Pierre, 2007 : « Remarques sur la grammaire comparée – reconstruction en linguistique romane (situation, perspectives) », *Mémoires de la Société de linguistique de Paris* 15, p. 57-72.
- CHAMBON, Jean-Pierre, 2010 : « Pratique étymologique en domaine (gallo-)roman et grammaire comparée-reconstruction. À propos du traitement des mots héréditaires dans le TLF et le FEW », in Choi-Jonin, Injoo, Duval, Marc et Soutet, Olivier, *Typologie et comparatisme. Hommages offerts à Alain Lemaréchal*, Louvain/Paris/Walpole, Peeters, p. 61-75.
- CHAMBON, Jean-Pierre, 2011 : « Note sur la diachronie du vocalisme accentué en istriote/istroroman et sur la place de ce groupe de parlers au sein de la branche romane », *Bulletin de la Société de Linguistique de Paris* 106, 1, p. 293-303.
- CHAMBON, Jean-Pierre, à paraître a : « Étymologie lexicale, étymologie onomastique : quoi de neuf ? », in : Casanova, Emili et al. (éd.) : *Actes del 26^e Congrès Internacional de Lingüística i Filologia Romàniques (València 2010)*, Berlin/New York, de Gruyter.
- CHAMBON, Jean-Pierre (à paraître b) : « Réflexions sur la reconstruction comparative en étymologie romane (entre Meillet et Herman) », in *Actes du Colloque en l'honneur du 80^e anniversaire de Max Pfister, Zurich, 19-21 avril 2012*.
- CHAMBON, Jean-Pierre / GREUB, Yan, 2002 : « Note sur l'âge du (proto)gascon », *Revue de linguistique romane* 66, p. 473-495.
- CHAUVEAU, Jean-Paul, 2010, « Structuration des articles et rédaction du commentaire », in: DÉRom, Nancy: ATILF, (<http://www.atilf.fr/DÉRom>), sous « Actualités et historique », 26-30 juillet 2010.
- COSERIU, Eugenio, 1998 : « Le double problème des unités 'dia-s', in *Les cahiers dia. Études sur la diachronie et la variation linguistique* 1, Gand, p. 9-16.
- COSERIU, Eugenio, 1987 : « Le latin vulgaire et le type linguistique roman », in Herman (József) (éd.) : *Latin vulgaire – latin tardif. Actes du I^{er} Colloque international sur le latin vulgaire et tardif (Pécs, 2-5 septembre 1985)*, Tübingen, Niemeyer, p. 53-64.
- DARDEL, Robert de, 1965 : *Recherches sur le genre roman des substantifs de la troisième déclinaison*, Genève, Droz.
- DARDEL, Robert de, 1976 : « Une analyse spatio-temporelle du roman reconstruit (à propos du genre) », in : Vårvaro, Alberto (éd.), *XIV Congresso internazionale di linguistica et filologia romanza, Napoli 15-20 aprile 1974*, Naples/Amsterdam, Macchiaroli/Benjamins, 14/2, p. 75-82.
- DARDEL, Robert de / WÜEST, Jakob, 1993 : « Les systèmes casuels du protoroman. Les deux cycles de simplification », *Vox Romanica* 52, p. 25-65.
- DARDEL, Robert de, 1996 : *À la recherche du protoroman*, Tübingen, Niemeyer.
- DARDEL, Robert de, 2007 : « Une mise au point et une autocritique relatives au protoroman », *Revue de linguistique romane* 71, p. 329-358.
- DARDEL, Robert de, 2009 : « La valeur ajoutée du latin global ». *Revue de linguistique romane* 73, 5-26.
- ERNOUT, Alfred / MEILLET, Antoine (1959⁴ [1932¹]) : *Dictionnaire étymologique de la langue latine. Histoire des mots*, Paris, Klincksieck.
- FISCHER, Iancu, 1969 : « III. Lexicul. 1. Fondul panroman », in Rosetti, Alexandru et al., *Istoria limbii române*, Bucarest, Editura Academiei Republicii Socialiste România, 2, p. 110-116.
- GREUB, Jan « La fragmentation de la Romania et la formation de l'espace linguistique francoprovençal : le témoignage des monnaies mérovingiennes », in *Actes de la Conférence annuelle sur l'activité scientifique du Centre d'études francoprovençales. Aux racines du francoprovençal (Saint-Nicolas, 20-21 décembre 2003)*, Aoste, Centre d'études francoprovençales "René Willien", p. 15-22.
- HERMAN, Joseph, 1967 : *Le Latin vulgaire* Paris, Presses Universitaires de France.
- HERMAN, Joseph, 1996 : « Les variétés du latin », in Holtus, Günter, Metzeltin, Michael, Schmitt, Christian (éds.), *Lexikon der Romanistischen Linguistik. Latein und Romanisch : Historisch-vergleichende Grammatik der romanischen Sprachen*, 2/1, Tübingen, Niemeyer, p. 44-61.
- HERMAN, József, 2006 : « La chronologie de la transition du latin aux langues romanes : un dossier revisité », in Arias Abellán, Carmen (éd.) : *Latin vulgaire – latin tardif. Actes du VI^e Colloque international sur le latin vulgaire et tardif (Séville, 2-6 septembre 2003)*, Séville, Université de Séville, p. 361-368.
- KABATEK, Johannes, 2000 : « L'oral et l'écrit – quelques aspects théoriques d'un 'nouveau' paradigme dans le

- canon de la linguistique romane », in Dahmen, Wolfgang, Holtus, Günter, Kramer, Johannes, Metzeltin, Michael, Schweickard, Wolfgang et Winkelmann, Otto (éd.), *Kanonbildung in der Romanistik und in den Nachbardisziplinen. Romanistisches Kolloquium XIV*, Tübingen, Narr, p. 305-320.
- KOCH, Peter, 2008 : « Le latin – une langue pas tout à fait comme les autres ? Le problème de la diglossie en Gaule septentrionale », in Van Acker, Marieke, Van Deyck, Rika et Van Uyftanghe, Marc, *Latin écrit – Roman oral ? De la dichotomisation à la continuité*, Turnhout, Brepols, p. 43-67.
- KOCH, Peter et OESTERREICHER, Wulf, 2008 : « Comparaison historique de l'architecture des langues romanes » in Ernst, Gerhard, Gleßgen, Martin-Dietrich, Schmitt, Christian, et Schweickard, Wolfgang (éd.), *Romanische Sprachgeschichte. Ein internationales Handbuch zur Geschichte der romanischen Sprachen*, Berlin/New York, De Gruyter, 3, p. 2575-2610.
- LB, 2011 = *Livre bleu, version 6, réalisée à l'occasion du septième Atelier DÉRom, Sarrebruck, 18/19 novembre 2011*, Nancy, ATILF, site internet (<http://www.atilf.fr/DÉRom>).
- LLOYD, Paul M., 1993 : *Del latin al español. I. Fonología y morfología históricas de la lengua española*, Madrid, Gredos.
- MANIET, Albert, 1975⁵ [1950¹] : *La Phonétique historique du latin dans le cadre des langues indo-européennes*, Paris, Klincksieck.
- MEILLET, Antoine, 1925 : *La méthode comparative en linguistique historique*, Paris, Champion.
- MEYER-LÜBKE, Wilhelm, 1890–1906 : *Grammaire des langues romanes*, 4 volumes, Paris, Welter.
- MEYER-LÜBKE, Wilhelm, 1901 : *Einführung in das Studium der romanischen Sprachwissenschaft*, I, Heidelberg, Winter.
- MOLINELLI, Piera, 2006 : « Per una sociolinguistica del latino », in Carmen Arias Abellán (éd.), *Latin vulgaire – latin tardif VII. Actes du VIIème Colloque International sur le Latin Vulgaire et Tardif (Séville, 2-6 septembre 2003)*, Séville : Universidad de Sevilla, 463-474.
- ROSETTI, Alexandru, 1986. *Istoria limbii române. De la origini și pînă la începutul secolului al XVII-lea*, Bucarest, Editura Științifică și Enciclopedică.
- SCHWEICKARD, Wolfgang (2010): « Die Arbeitsgrundlagen der romanischen etymologischen Forschung : vom REW zum DÉRom », *Romanistik in Geschichte und Gegenwart* 16: 3-13.
- SEIDL, Christian, 2003 : « Les variétés du latin », in Ernst, Gerhard, Gleßgen, Martin-Dietrich, Schmitt, Christian, et Schweickard, Wolfgang (éd.), *Romanische Sprachgeschichte. Ein internationales Handbuch zur Geschichte der romanischen Sprachen*, Berlin/New York, De Gruyter, 3, p. 515-530.
- SEIDL, Christian, 2003 : « Latein + Griechisch + Gotisch = Galloitalisch. Zur Herausbildung der norditalienischen Dialekte im 6. Jahrhundert », in Graf (Martin H.)/Moser (Christian) (éd.), *Strenarum lanx. Beiträge zur Philologie und Geschichte des Mittelalters und der Frühen Neuzeit. Festgabe für Peter Stotz zum 40-jährigen Jubiläum des Mittellateinischen Seminars der Universität Zürich*, Zug, Archius, p. 9-38.
- SERBAT, Guy, 1986. *Les Structures du latin*, Paris, Picard.
- STEFENELLI, Arnulf, 1996. « Thesen zur Entstehung und Ausgliederung der romanischen Sprachen », in *Lexikon der Romanistischen Linguistik*, Tübingen, Niemeyer, 2/1, p. 73-90.
- STRAKA, Georges, 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, p. 249-267.
- VÄÄNÄNEN, Veikko, 1963 : *Introduction au latin vulgaire*, Paris, Klincksieck.
- VÄÄNÄNEN, Veikko, 1959 : *Le latin vulgaire des inscriptions pompéiennes*. Nouvelle édition revue et augmentée, Berlin, Akademie.
- VÁRVARO, Alberto, 2011a : « Il DÉRom : un nuovo REW ? », *Revue de linguistique romane* 75, p. 297-304.
- VÁRVARO, Alberto, 2011b : « La « rupture épistémologique » del DÉRom. Ancora sul metodo dell'etimologia romanza », *Revue de linguistique romane* 75, p. 623-627.

Myriam BENARROCH est agrégée de Portugais et maître de conférences à l'Université Paris-Sorbonne où elle enseigne l'histoire de la langue portugaise, la linguistique romane et la traduction. Elle est rattachée à l'équipe de recherche « Linguistique et lexicographie latines et romanes » (Paris-Sorbonne, dir. Michèle Fruyt) et membre de l'équipe internationale DÉRom (*Dictionnaire Étymologique Roman*), en tant que réviseur et rédacteur. Elle est l'auteur d'une thèse sur les arabismes du portugais au 16^e s. (2000). Ses travaux portent sur l'étymologie du lexique portugais, l'étymologie du lexique héréditaire roman et le lexique dans la lexicographie portugaise et romane classique et contemporaine. Parmi les dernières publications : « L'apport des dictionnaires de Jerónimo Cardoso (XVI^e siècle) à la datation du *Dicionário Houaiss* (2010) » ; « L'apport du DÉRom à l'étymologie portugaise » (CILPR 2010, à paraître), « Quelques réflexions sur les arabismes portugais et espagnols dans deux dictionnaires de 1611, le *Dictionarium* d'Agostinho Barbosa et le *Tesoro* de Sebastián de Covarrubias » (2011) ; « L'étymologie du lexique héréditaire : en quoi l'étymologie panromane est-elle plus puissante que l'étymologie idioromane ? L'exemple du DÉRom » (colloque 2011, à paraître).