

Jean-Paul Chauveau

1.3. Reconstruction comparative et histoire sémantique

1 Introduction

La reconstruction comparative est une procédure heuristique à laquelle les romanistes diachronistes ont longtemps préféré l'histoire méthodiquement retracée des anciens témoignages jusqu'au présent.¹ Ceux d'entre eux qui ont entrepris de la pratiquer n'en sont plus à leurs débuts dans la recherche étymologique, et ils s'appuient sur les travaux antérieurs et parallèles où domine la perspective historique. La comparaison-reconstruction, censée enjamber les lacunes historiques, ne saurait cependant oublier les acquis de la méthode philologico-historique dans le domaine roman. Elle n'aurait aucun intérêt, non plus, à pratiquer la *tabula rasa* et à reconstruire les étymons à partir des seules données modernes et sans se soucier le moins du monde de la documentation qui subsiste du latin de l'Antiquité (cf. Maggiore/Buchi 2014). Les lexèmes des langues romanes qui sont ses données de base ne sont pas des matériaux bruts, mais des faits linguistiques établis selon les résultats de la recherche historique antérieure et, à défaut de tels travaux, par une recherche neuve qui interroge l'ancienneté, la géographie, les formes, les emplois et les sens des cognats de chaque langue, afin de déterminer l'assiette de chacun de ceux-ci pour les positionner correctement dans le schéma reconstitutif.

De ce fait, si l'on compare les subdivisions sémantiques des grands dictionnaires étymologiques et les quelques sémèmes que retient la reconstruction, on voit beaucoup de similitudes, mais aussi, parfois, des différences. Les premières illustrent qu'il n'y a pas d'opposition fondamentale entre les deux procédures, tandis que les secondes sont fonction de leurs objectifs distincts, bien plus que de points de vue antagonistes.

¹ Ce chapitre fait suite à nos réflexions publiées dans le DÉRom 1 (Chauveau 2014).

2 Sémèmes de création protoromane

L'article consacré à l'adjectif */'gross-u/ (Dworkin/Maggiore 2015/2016 in DÉRom s.v.) distingue deux sens : 'qui dépasse la mesure ordinaire, gros' et 'qui manque de raffinement, grossier', mais n'a qu'une subdivision, car ces deux sens sont donnés comme systématiquement présents partout, sans qu'il soit nécessaire d'en fournir la documentation langue par langue. C'est seulement pour les premières attestations qu'est précisé lequel des deux sens elles documentent. Ainsi le type it. *grosso* est d'abord attesté par aitsept. *grose* f.pl. 'grossières' (1176/1200), tandis qu'esp. *grueso* l'est d'abord par *gruessos* pl. 'gros' (1140), par exemple. Le commentaire précise que ces deux sens sont déjà attestés en latin impérial, tardivement donc, mais suffisamment tôt pour confirmer, si cela était nécessaire – en réalité, le témoignage convergent du sarde, du roumain, du végliote, de l'istriote, de l'italien, du frioulan, du ladin, du romanche, du français, du francoprovençal, de l'occitan, du gascon, du catalan, de l'espagnol, de l'asturien et du galégo-portugais suffit à l'établir –, le caractère protoroman de cet adjectif. L'article correspondant du FEW (von Wartburg 1947 in FEW 4, 274a–283b, GRÖSSUS) comporte une dizaine de subdivisions selon les domaines d'emploi et les sens de l'adjectif. Certains de ces sémèmes sont d'époque récente ou localisée : 'houleux, agité (de la mer)', 'en crue (d'un cours d'eau)', 'âgé (d'un personne)', mais la plupart sont déjà attestés à l'époque médiévale. Ce n'est pas pour autant que ces sens sont traités comme originels. Ils sont distingués pour décrire les innovations sémantiques des langues-objets du FEW (le français, le francoprovençal, l'occitan et le gascon), dont les articulations sont explicitées dans le commentaire. Finalement seuls deux sens sont rapportés dans le commentaire (FEW 4, 280b et 281ab) à la période latine, celui de 'gros' ('dick') et celui de 'grossier' ('grob'), au concret et au figuré, exactement comme dans le DÉRom. Le contenu des articles s'accorde avec les objectifs spécifiques des deux dictionnaires : représenter l'histoire d'un lexème dans un ensemble de branches de la famille romane (FEW) ou bien déterminer le point de départ protoroman à partir duquel les différentes branches romanes ont développé leur propre individualité (DÉRom). S'y ajoute la différence des modes de présentation des matériaux, celui du DÉRom étant plus contraint (cf. Souvay/Renders 2014).

Von Wartburg 1945 in FEW 2, 1277b–1279a, CRASSUS ne distingue dans les continueurs directs et indirects de CRASSUS qu'un sens primaire 'gras' et des sens secondaires. Pour ces derniers, il indique que « die bed. nüancen und ablt. des gallorom. haben weitgehend in den andern rom. sprachen parallelen » ('les nuances sémantiques et les dérivés galloromans disposent en grande partie de

parallèles dans les autres langues romanes’). Dans l’article consacré à l’adjectif */‘grass-u/ du DÉRom (Dworkin/Maggiore 2014/2015 in DÉRom s.v.), au contraire, ces parallélismes conduisent à distinguer deux sens : ‘qui contient de la graisse, gras’ et ‘qui produit beaucoup de végétation utile, fertile’. Le premier est continué dans tous les parlers romans à l’exception de l’espagnol, de l’asturien et du galégo-portugais. Le second, qui résulte d’une évolution sémantique à partir du premier, quoique « moins répandu, a toutefois une distribution aréale assez large (dacoroum. it. frioul. lad. romanch. fr. occit. gasc. cat. arag.) pour qu’on le considère comme héréditaire » (*/‘grass-u/, commentaire). En outre, les deux sens sont attestés déjà chez les auteurs latins de la période classique. La comparaison romane, qui est seulement évoquée dans le FEW, est ici mobilisée pour la reconstruction protoromane. Dans ce cas, c’est le nombre important de parlers concernés et leur appartenance à plusieurs branches romanes, dont le roumain, qui sont considérés comme décisifs pour l’attribution de ce sens secondaire à la période protoromane.

3 Sémèmes de création idioromane diffusés par emprunts intraromans

Mais il ne s’agit pas seulement de constater la présence du sémème secondaire dans une pluralité d’idiomes pour reporter sa formation à la protolange. Encore faut-il écarter les possibilités que cette pluralité relève d’autres explications. Et la reconstitution historique est l’un des outils indispensables dans cette perspective.

L’article consacré au substantif masculin */‘merl-u/ (Albrecht 2016 in DÉRom s.v.) n’affecte à cet étymon que des continuateurs (de genre masculin) du nom de l’oiseau, le merle, à la différence aussi bien de REW₃ s.v. *měřulus* ‘ouverture pratiquée à intervalles réguliers du sommet d’un rempart, créneau’ (‘Mauerzinne’) – article qui fait suite à l’article *měřūla*, -us ‘merle ; merlan’ (‘Amsel ; Meeramsel’) – que de von Wartburg 1967 in FEW 6/2, 38b, *MĚŘŮLUS* ‘créneau’ (‘zinne’), qui en font un sens métaphorique du latin tardif *merulus* ‘merle’ qui se serait maintenu dans des parlers de l’Italo-romania et de la Gallo-romania. La note 5 de l’article du DÉRom réfute la possibilité que ce nouveau sémantisme puisse être attribué au protoroman. C’est une réappréciation de l’histoire de ce type lexical, conduite par l’auteur, Louis Albrecht, avec le concours d’étymologistes expérimentés intervenus comme réviseurs, qui invalide l’étymon de ses devanciers, tout autant d’ailleurs que l’hypothèse d’un germa-

nisme aventurée par Coromines 1990 in DECat 5, 609–611 s.v. *merlet*. Ce type est d'abord attesté par le latin médiéval *merulus* s.m. 'créneau' du 10^e au 12^e siècle en Italie et en Italie seulement (Niermeyer/VanDeKieft). Les équivalents sur le sol de la Galloromania et leurs dérivés ne sont densément attestés que dans le Sud-Est occitan et seulement à partir du 13^e siècle. En français, ils n'apparaissent que ponctuellement, en traduction de l'italien (cf. Huguet 5, 217 s.v. *merle*⁴) ou dans des zones de contacts linguistiques où le français est langue expatriée, en Terre sainte, ou langue administrative, en domaine francoprovençal (cf. Gdf 5, 230c–261a s.v. *merle*¹). C'est encore le cas pour le synonyme *merlon*, lui aussi emprunté à l'italien : it. *merlone* est d'abord attesté dans une traduction de l'italien (Tournes 1593, 68, 72, 75). La présence du sémème 'créneau' dans plusieurs langues de la Romania occidentale relève d'un phénomène de diffusion au contact, à partir du type it. *merlo*, par le relais de l'occitan du sud-est jusqu'en catalan (cf. cat. *merlet*, dp. fin 13^e s., DECat 5, 609b) et en franco-provençal (*merlos* pl., dp. 1358, Philippon 1884, 579–581 ; 1346/1378, Durdilly 1975, 315–334). Propre au domaine italien au départ, ce sémème ne peut pas être inséré dans un schéma de comparaison-reconstruction et sera donc tenu pour idioroman. Cela ne préjuge pas de la date d'apparition de ce sens métaphorique, mais entérine le fait que la reconstruction ne peut apporter aucune information sur ce point et ne peut donc, par méthode, se permettre aucune conjecture. Le désaccord entre le DÉRom et les autres dictionnaires étymologiques ne porte pas sur la méthode, mais sur l'interprétation qui peut être donnée des résultats.

Ce n'est pas parce que l'on constate un accord sur un sémème entre plusieurs parlers romans qu'il est originel. Cette communauté partagée peut être attribuée à d'autres raisons. Dans ce dernier exemple, la communauté est le résultat d'une diffusion d'époque romane, comme la documentation historique permet de l'établir. Mais il est encore d'autres possibilités qui sont susceptibles de créer une uniformité secondaire, telles qu'une évolution sémantique parallèle ou un emprunt sémantique, notamment. Les articles du DÉRom en illustrent quelques exemples.

4 Sémèmes de création parallèle

L'article */βent-u/ (Tamba 2016 in DÉRom s.v.) distingue deux types morphologiques selon le genre, masculin ou neutre, des cognats dans les langues romanes, mais ne leur affecte qu'un seul sens : 'mouvement de l'air, vent'. Une note cependant indique que « les continuateurs de */βent-u/ présentent un certain nombre de sens secondaires [...], dont aucun ne paraît suffisamment

diffusé ni documenté avec suffisamment d'ancienneté pour justifier sa reconstruction en protoroman » (*'/βent-u/ n. 3). Il en est pourtant un qui a des équivalents dans quatre branches et qui est aussi attesté en latin écrit de l'Antiquité, le sens de 'gaz intestinal qui sort de l'anus, pet', documenté en dacoroumain, en italien, en français et en portugais. Mais, malgré ces conditions apparemment favorables, ce sens n'a pas été retenu comme relevant du stade protoroman. Du fait du caractère assez récent des quatre unités lexico-sémantiques concernées, il pourrait tout simplement s'agir d'autant de métaphores indépendantes, comme en anglais ou en allemand, par exemple. En outre, ces sens métaphoriques étant surtout portés par le milieu médical, dont la formation était dispensée en latin jusqu'à l'époque moderne, il est fort possible d'y voir des emprunts sémantiques savants au latin médical. En outre, les parlers romans continuent, pour exprimer le sens 'pet', le substantif */'pedit-u/ (REW₃ s.v. *pēdītum*), ainsi que, pour le verbe correspondant, */'ped-e-/ (REW₃ s.v. *pēdēre*), */'βis-i-/ (REW₃ s.v. *vīssire*) et */'βis-in-a-/ (REW₃ s.v. **vīssinare*). Il n'y a aucun argument fort pour affecter au protoroman un tel sens métaphorique pour */'βent-u/ et, au contraire, il y a des solutions alternatives qui sont plus plausibles. D'ailleurs, Reinhard 1959 in FEW 14, 259a, VENTUS I 1 b e classait les données occitanes et françaises correspondantes parmi les sens secondaires, sans argumenter ce classement.

Là où Schmidt/Schweickard 2015 in DÉRom s.v. */'kant-a-/ ne reconnaissent qu'un seul sens, celui de 'former avec la voix une suite de sons musicaux, chanter', exactement comme von Wartburg 1937 in FEW 2, 220b, CANTARE (cf. 2, 224a), Franchi/Hohnerlein/Pfister 2008 in LEI 10, 1336–1399, CANTĀRE adjoignent à ce sens premier et fondamental « un ampio spettro di significati derivati » (LEI 10, 1397) dont ils signalent surtout les correspondants français et occitans, mais aussi « le accezioni 'comporre versi', 'declamare, recitare' » du latin classique. Ce sont des correspondances qui éclairent, comparativement, le cadre dans lequel s'est diversifié sémantiquement le verbe dans le domaine italien. Mais ces comparaisons latine et romane n'ont pas pour finalité de reconstruire la base qui aurait formé le point de départ des verbes romans. L'article du DÉRom n'établit qu'un seul sens originel qui est prouvé de façon indubitable par la présence de cognats dans tous les parlers romans sans exception. Les autres sens attestés dans plusieurs langues romanes n'impliquent pas nécessairement une communauté originelle. D'une part, le maintien du latin comme langue de la communication ecclésiastique et l'importance liturgique du chant ne permettent pas d'exclure des relatinisations sémantiques : it. *cantare* v.tr. 'commémorer, célébrer avec des chants' s'est maintenu dans l'usage ecclésiastique, est-ce pour autant qu'il a appartenu de manière constante à l'usage

quotidien ? D'autre part, les usages du chant lui valent d'être considéré comme une modalité artistique et intensive de la parole, qui permet le développement de sens tels que 'proclamer', 'célébrer', 'louer', 'exposer sous forme versifiée' ou encore 'avouer' et, en emploi absolu, 'manifester sa joie'. De tels développements peuvent se déterminer potentiellement à toute époque. Ce n'est pas la date de leur première apparition dans le lexique qui pourrait permettre de choisir entre maintien et innovation. Faute de pouvoir trancher avec assurance, la reconstruction doit opter pour la solution qui ne suscite aucun doute.

5 À la recherche de critères

L'article consacré au substantif neutre */'kaput/ (Schmidt/Schweickard 2015/2016 in DÉRom s.v.) connaît trois subdivisions formelles et seulement deux subdivisions sémantiques, 'tête' et 'extrémité'. Le LEI, quant à lui (Hohnerlein/Pfister/Cornagliotti 2009 in LEI 11, 1021–1361, CAPUT/CAPUS), repère six domaines référentiels où l'étymon est employé et dans chacun desquels le sens peut se ramifier et subdiviser. Et il n'a aucune difficulté à documenter des prototypes latins d'un bon nombre des sens qu'il distingue et à en signaler des équivalents dans d'autres branches romanes que l'italien et ses dialectes. De façon implicite, l'article du DÉRom montre la méthode qu'il met en œuvre. Dans la subdivision formelle qui traite les représentants du pluriel */'kapit-a/, sont rassemblées d'abord les données romanes qui maintiennent le sens concret : 'tête'. Les cognats sardes et roumains, qui en forment la majeure partie, sont sémantisés comme 'animal d'élevage en tant qu'unité de mesure d'un troupeau, tête de bétail'. La subdivision consacrée au type */'kap-u/ s.m. au sens concret 'tête' s'appuie sur gasc. *cap* (dp. 1495, 'maison principale d'un ordre religieux') et ast. *cabu* 'tête de bétail' (dp. av. 1144). On voit que ces attestations sont considérées comme des sens métonymiques (l'individu dénommé par sa tête) ou des sens métaphoriques (le lieu principal et directeur envisagé comme constituant la tête qui dirige l'activité de la communauté humaine). Ces sens secondaires sont traités comme le résultat de figures de langue qui sont des potentialités en permanence disponibles pour les locuteurs. Seule l'histoire lexicale est susceptible d'en démontrer l'ancienneté dans les langues romanes dont le passé est bien documenté et leur présence dès la période latine. Elle fournit, en conséquence, des arguments qui autorisent à soutenir le maintien d'une riche gamme sémantique dès la période constitutive des langues romanes. Mais la reconstruction ne peut pas se permettre de s'appuyer sur les données historiques de quelques langues pour combler le silence des autres à date ancienne, et en écartant la

possibilité que ces données soient des emprunts sémantiques au latin classique, dont la connaissance n'avait pas disparu dans certains milieux, ou des innovations idioromanes. Il est de fait que les concurrents des héritiers de protorom. */'kaput/ qu'ont adoptés postérieurement certaines langues romanes, par exemple fr. *tête* ou esp. *cabeza*, ont développé de tels sens secondaires.

À l'inverse, lorsque le corrélat en latin écrit de l'Antiquité de l'étymon protoroman est attesté tardivement, faiblement et sous des formes morphologiquement variables, il est tout à fait improbable que sa diffusion et sa ramification sémantique soient dues à une influence savante. Et s'il a des continuateurs dans presque toutes les langues romanes documentant plusieurs sémèmes qui ne sont pas tous liés entre eux par des évolutions banales, la reconstruction montre tout son profit et son intérêt. C'est le cas de l'article consacré au substantif neutre */'dɔl-u/ (Morcov 2014 in DERom s.v.), pour lequel les données romanes se regroupent sous cinq sémantismes : 'douleur physique', 'douleur morale', 'deuil', 'manifestation de deuil' et 'compassion'. Comme l'explicite le commentaire, « le fait que le sarde partage la totalité des sens de */'dɔl-u/ avec des parlers issus du protoroman continental incite à les reconstruire tous pour la période du protoroman commun ». Certes, les sens 'douleur physique' et 'douleur morale' sont liés par analogie et ceux de 'deuil' et 'manifestation de deuil' par une métonymie. Mais ces sens sont aussi liés en latin écrit, respectivement, pour *dolor* et pour *luctus*, dont les corrélats en protoroman */'do-l-or-e/ et */'luk-t-u/ ont été concurrencés plus ou moins vigoureusement par ceux de */'dɔl-u/. Il est très probable que celui-ci, étant donné la diffusion de ses continuateurs, a assumé très tôt l'ensemble de ces sémantismes. Il y a bien eu une influence savante, mais elle a joué en faveur de la vitalité des représentants de *dolor* et *luctus*, certainement pas au profit de */'dɔl-u/, dont la diffusion et la ramification sémantique peuvent être rapportées au protoroman.

Le recours à l'histoire permet aussi de ne pas mettre sur un seul et même plan tous les matériaux lexicaux similaires : comparaison n'est pas raison, comme on sait.

Maggiore 2012–2015 in DÉRom s.v. */'kuer-e-/ (cf. aussi Maggiore 2014) distingue trois sens, tous transitifs : 'chercher', 'vouloir' et 'demander', qui sont exactement les mêmes que ceux que von Wartburg 1945 in FEW 2, 1409b–1410a, QUÆRERE reconnaît dans les données romanes. Mais l'auteur de l'article du DÉRom évite d'accorder un rôle dans la reconstruction sémantique au verbe d'ancien français et d'ancien occitan *querre* 'vouloir, désirer' et 'demander', dont les dictionnaires donnent de nombreuses attestations. Il adopte tacitement l'opinion synthétisée par von Wartburg dans la note 4 de l'article du FEW (2, 1410b) qui, se fondant sur les contraintes syntaxiques limitant ces emplois, y

voit un développement spontané et autonome à partir de ‘chercher’. Semblablement, Maggiore in DÉRom s.v. */'kuɛr-e-/ n. 14 s'appuie, mais explicitement cette fois, sur les travaux antérieurs pour exclure que le sens ‘aimer’ soit d'époque protoromane.

6 Problèmes d'interprétation de l'absence d'un cognat

Le même article */'kuɛr-e-/ illustre une autre inférence qu'une reconstruction purement mécanique serait tentée de tirer des données. Le sens ‘chercher’ n'est pas documenté en sarde. On pourrait en tirer l'hypothèse, étant donné que le sarde ne connaît pas habituellement les formes et les sens qui se sont développés tardivement en protoroman, que ce sens est plus tardif que les deux autres. Une telle inférence serait en contradiction avec la logique sémantique, qui explique le développement des sens ‘demander’ et ‘vouloir’ dans l'ordre ‘chercher’ > ‘chercher à obtenir, demander’ > ‘vouloir’. Il en résulte que « l'absence du sens ‘chercher’ en sarde [...] s'explique aisément par un évincement idioroman » (Maggiore in DÉRom s.v. */'kuɛr-e-/ n. 13). Comme le montre la disparition totale en catalan ou partielle en français des continuateurs de */'kuɛr-e-/ à date historique, les manques dans telle ou telle langue peuvent parfaitement représenter des phénomènes spécifiques, auxquels on ne doit faire jouer aucun rôle dans la reconstruction.

La présence d'un sens dans une langue donnée peut être attribuée soit à un emprunt sémantique au latin, quand la connaissance et la pratique du latin ont été cultivées, ou à une autre langue romane, s'il y a des contacts étroits entre les deux langues, et si cela se produit dans un contexte favorable à l'emprunt, soit représenter une évolution sémantique endogène, si cette évolution est banale, susceptible de se déclencher de façon spontanée. Moyennant ces précautions, les présences permettent de préciser le scénario reconstitutif. On ne peut en dire autant des manques. L'absence d'un sens dans une langue donnée peut être attribuée à un phénomène d'obsolescence spontanée qu'on peut chercher à expliquer de diverses façons, mais que la reconstruction comparative doit surtout constater et enregistrer, sans chercher à en tirer des informations sur le schéma reconstitutif. De la même manière, qu'un étymon ait un continuateur régulier dans une langue, c'est une information positive sur la vivacité du protoroman. Mais qu'un autre étymon n'ait pas de continuateur régulier dans telle ou telle langue, c'est une information sur ces langues, mais pas sur le protoro-

man : il n'y a qu'une seule explication au maintien d'un continuateur régulier dans une langue, c'est qu'il y était déjà présent dès les débuts de cette langue ; mais, si on n'y repère aucune trace de son maintien, on peut imaginer plusieurs hypothèses pour sa disparition.

C'est ce qui a conduit à corriger une première version de l'article */arbor-e/ s.f. (Álvarez Pérez 2014 in DÉRom 1 s.v.), dont le lemme ne comportait qu'un seul sens : 'plante au tronc ligneux qui se ramifie à une certaine hauteur', tandis que les subdivisions de l'article en reconstruisaient trois : 'plante de tronc ligneux qui se ramifie à une certaine hauteur du sol, arbre', 'espar planté sur le pont ou dans la quille d'un navire, mât' et 'pièce qui sert de support à d'autres pièces animées, pièce maîtresse', et que ces trois sens étaient dits reconstruc-tibles pour le protoroman dans le commentaire. Il s'agissait là d'une question de principe de la reconstruction sémantique restée non résolue au sein du projet que ses directeurs pointaient déjà du doigt, en surlignant la différence de traitement entre les articles */arbor-e/ et */leβ-a-/, par exemple (Buchi/Schweickard 2014, 29–30). La nouvelle version de l'article (Álvarez Pérez 2016 in DÉRom s.v.) mentionne bien les trois sens 'plante au tronc ligneux qui se ramifie à une certaine hauteur ; espar planté sur le pont ou dans la quille d'un navire ; pièce qui sert de support à d'autres pièces animées' dans le lemme.

On peut se demander où se situe la raison de la discordance originelle entre le lemme d'une part et le commentaire de l'autre. À y regarder de près, on s'aperçoit qu'elle était due au fait que la reconstruction détermine un protoroman *stricto sensu*, porteur du sens 'arbre', d'extension panromane et pour lequel deux genres étaient possibles, le féminin originel et le masculin évolué. Les deux autres unités lexico-sémantiques sont dites relever d'un protoroman *lato sensu* plus tardif, du fait qu'elles occupent des aires limitées et qu'elles sont corrélatives du seul genre masculin innovant. Au sens restreint ou au sens large, les trois sémèmes reconstruits relèvent bien du protoroman global, c'est-à-dire de la période au cours de laquelle il n'y a pas encore d'individuation des différents parlers romans, même si la différenciation linguistique est déjà entamée. Mais, ce qui est notable, c'est que la différence entre protoroman *stricto sensu* et protoroman *lato sensu* est fondée exclusivement sur les différences entre les aires couvertes par les trois sens. Chacune des aires est dessinée par la somme des parlers où le sens en cause est ou a été attesté et leur datation relative en est la conséquence :

- toutes les langues attestent le sens 'arbre' → il est donc originel ;
- le sens 'mât' est méditerranéen, mais non sarde → il s'est donc déterminé postérieurement à l'individuation du sarde ;

- le sens ‘pièce maîtresse’ couvre seulement une vaste aire italo-occidentale, mais pas le domaine roumain → il doit être postérieur à la formation du domaine roumain.

Ces conclusions résultaient de deux inférences :

- les présences du sens déterminent les aires ;
- les absences déterminent la datation des aires.

Mais, comme on l’a vu dans l’exemple précédent, les absences n’ont pas, par nécessité, une valeur catégorique. Si le sens manque dans un domaine donné, c’est peut-être parce qu’il n’a jamais été connu dans le domaine géographique où cette langue s’est développée, mais c’est peut-être aussi parce qu’on n’y a jamais eu besoin de ce sens ou que ce sens s’est perdu ou qu’on ne dispose pas de la documentation qui pourrait l’attester. Il faut bien dire que ces sens secondaires sont des sens techniques, typiquement peu susceptibles de passer dans les œuvres littéraires et souvent périmés par un changement technique. Nous pouvons négliger le silence portugais, parce que nous disposons d’une épave ancienne qui documente le sens ‘mât’ dans cette langue et que nous devons uniquement à la vitalité de la marine portugaise : « le portugais connaît un vestige du sens ‘mât’ dans le syntagme *árvore seca* ‘épars d’un navire où toutes les voiles sont attachées’ (2^e m. 15^e s. – 1712/1728, PicoTerminologia 251 ; HouaissGrande ; Bluteau) » (Álvarez Pérez 2014 in DÉRom 1, 352, */arbor-e/ n. 13). Si un terme de marine traditionnel peut se perdre dans la langue d’un peuple dont les marins ont ouvert les principales routes maritimes mondiales, il n’y a pas à s’étonner que la même chose ait pu arriver dans le contexte d’une marine moins vigoureuse. Qu’une langue comme le roumain, qu’on nous décrit comme le moyen d’expression originel de communautés pastorales, ait dû emprunter à l’époque moderne *arbor* pour les sens techniques communs dans les autres langues romanes ne surprend pas non plus. Et ces interprétations d’ordre purement référentiel n’épuisent pas le champ des possibles.

L’interprétation de la présence est univoque, sauf si on peut prouver un emprunt sémantique. L’interprétation de l’absence ne l’est pas, parce qu’il y a toujours plusieurs possibilités distinctes pour l’expliquer. L’aréologie des présences fournit une base sûre à la reconstruction, mais, à elle seule, l’aréologie des absences ne fait que poser question à la reconstruction : l’interprétation des silences est hautement conjecturale.

C’est ce que permet de vérifier la confrontation avec les données latines dans le commentaire : « le corrélat du latin écrit du type I., *arbor*, -is s.f., est usuel durant toute l’Antiquité dans les sens ‘arbre’ (dp. Plaute [* *ca* 254 –

† 184], IEEDLatin ; cf. TLL 2, 419–427) et ‘arbre du pressoir’ (dp. Caton [* 234 – † 149], TLL 2, 427) et attesté depuis Virgile (* 70 – † 19) dans les sens ‘mât’ et ‘poutre’ (tous les deux TLL 2, 427) » (Álvarez Pérez 2014 in DÉRom 1, 355).

Les trois sens sont attestés dans la littérature latine la plus classique. Le sens de ‘mât’ est plus tardif, mais les Romains furent des terriens avant de devenir des marins. En tout cas, dès le premier siècle de notre ère, les trois sens existent en latin, c’est-à-dire dans la période où s’établit l’Empire romain dont le latin aura été la langue véhiculaire, avant d’en devenir, quelques siècles plus tard, la langue commune dont le protoroman est l’émanation. Il n’y a aucun élément qui interdirait d’attribuer ces trois sens au protoroman et qui obligerait à étager, dans ce cas, la formation des différents sens.

Nous sommes conscient que, ce faisant, nous recommandons une approche différente de la reconstruction sémantique par rapport à celle qui prévaut dans le domaine de la reconstruction phonologique ou morphologique, par exemple. À notre sens, cette différence de traitement se justifie par le fait que, tandis que les variantes phonologiques et morphologiques sont des unités d’un système clos ou très contraint, les variantes lexicales et sémantiques sont des unités d’une collection lâche et illimitée : les emprunts de phonèmes sont des raretés, ceux de morphèmes peuvent se compter, on ne peut qu’évaluer le volume des emprunts lexicaux et sémantiques. Ce qui vaut pour les gains vaut aussi pour les pertes. Les critères sur lesquels se fonde la reconstruction sémantique demandent à être pondérés.

7 Conclusion

Reconstruction comparative et histoire sémantiques sont des points de vue méthodologiques distincts sur le même objet, mais certainement pas des méthodes exclusives. La restitution du passé repose sur l’interprétation des faits qu’on a établis, à partir de la documentation historique, au moyen des outils de la philologie et de la linguistique diachronique. Quelle que soit la méthode adoptée, l’établissement des données est un prérequis. La comparaison-reconstruction se singularise par certaines exigences. Là où l’histoire invoque parfois l’ancienneté pour bâtir, hypothétiquement, un pont au-dessus d’un vide documentaire de quelques siècles jusqu’à la communauté linguistique protoromane, la reconstruction ne peut pas attribuer, sans examen, un sens spécifique, sur la foi de son ancienneté dans l’une ou l’autre langue, à une origine protoromane. Les différentes langues ont commencé très tôt leurs développements autonomes, sur le plan de la sémantique lexicale comme sur les autres plans.

Comme on le sait, fr. *empereur* (dp. mil. 11^e s. [*empereür*, *emperethur*], Alexis) et occit. *emperador* (dp. ca 1060, SFOYHA 292), par exemple, ont beau être documentés dès les premiers monuments de ces deux langues, ils n'en sont pas moins tenus pour des emprunts, parce que leur forme démontre qu'ils ne peuvent pas être héréditaires. Les emprunts au latin écrit ont été une possibilité précoce, et les emprunts sémantiques ont été d'autant plus faciles que les clercs étaient bilingues et que l'intercompréhension romane pouvait contribuer à égaliser là où s'étaient déterminées des lacunes ou des différenciations. Le nombre des cognats, leur dispersion à travers l'espace roman et leur pleine insertion dans le lexique de chacune des langues en cause sont des critères qui se conjuguent pour établir la continuité d'un sémème depuis l'origine, dans la mesure, toutefois, où il n'y a pas d'indice qu'il ait connu une interruption, dans l'usage commun des langues qui le maintiennent, qui aurait pu être compensée par une reviviscence savante ou bien une recreation par une innovation selon un procédé sémantique banal.

8 Bibliographie

- Buchi, Éva/Schweickard, Wolfgang, *Conception du projet*, in : Éva Buchi/Wolfgang Schweickard (edd.), *Dictionnaire Étymologique Roman (DÉRom). Genèse, méthodes et résultats*, Berlin, De Gruyter, 2014, 5–38.
- Chauveau, Jean-Paul, *Reconstruction sémantique*, in : Éva Buchi/Wolfgang Schweickard (edd.), *Dictionnaire Étymologique Roman (DÉRom). Genèse, méthodes et résultats*, Berlin, De Gruyter, 2014, 199–209.
- DECat = Coromines, Joan, *Diccionari etimològic i complementari de la llengua catalana*, 10 vol., Barcelone, Curial, 1980–2001.
- DÉRom = Buchi, Éva/Schweickard, Wolfgang (dir.), *Dictionnaire Étymologique Roman (DÉRom)*, Nancy, ATILF, <<http://www.atilf.fr./DERom>>, 2008–.
- DÉRom 1 = Buchi, Éva/Schweickard, Wolfgang (edd.), *Dictionnaire Étymologique Roman (DÉRom). Genèse, méthodes et résultats*, Berlin, De Gruyter, 2014.
- Durdilly, Paulette (ed.), *Documents linguistiques du Lyonnais (1225–1425)*, Paris, Éditions du CNRS, 1975.
- FEW = Wartburg, Walther von et al., *Französisches Etymologisches Wörterbuch. Eine darstellung des galloromanischen sprachschatzes*, 25 vol., Bonn/Heidelberg/Leipzig-Berlin/Bâle, Klopp/Winter/Teubner/Zbinden, 1922–2002.
- Gdf = Godefroy, Frédéric, *Dictionnaire de l'ancienne langue française et de tous ses dialectes du IX^e au XV^e siècle*, 8 vol., Paris, Vieweg, 1881–1895.
- Huguet = Huguet, Edmond, *Dictionnaire de la langue française du seizième siècle*, 7 vol., Paris, Champion/Didier, 1925–1967.
- LEI = Pfister, Max/Schweickard, Wolfgang (dir.), *Lessico Etimologico Italiano*, Wiesbaden, Reichert, 1979–.

- Maggiore, Marco, *Note di etimologia romanza a margine dell'articolo */'kuɛr-e-/ (« quærere ») del « Dictionnaire Étymologique Roman »*, in : Piera Molinelli/Pierluigi Cuzzolin/Chiara Fedriani (edd.), *Latin vulgaire – Latin tardif X. Actes du X^e colloque international sur le latin vulgaire et tardif (Bergamo, 5–9 septembre 2012)*, vol. 2, Bergame, Sestante, 2014, 599–614.
- Maggiore, Marco/Buchi, Éva, *Le statut du latin écrit de l'Antiquité en étymologie héréditaire française et romane*, in : Franck Neveu et al. (edd.), *Actes du Congrès Mondial de Linguistique Française 2014 (Berlin, 19–23 juillet 2014)*, Paris, Institut de Linguistique Française, <http://dx.doi.org/10.1051/shsconf/20140801161>, 2014, 313–325.
- Niermeyer/VanDeKieft = Niermeyer, Jan Frederik/Kieft, Co van de, *Mediae Latinitatis lexicon minus. Lexique latin médiéval – Medieval Latin Dictionary – Mittellateinisches Wörterbuch*, 2 vol., Leiden/Boston, Brill, ²2002 [¹1976].
- Philippon, Édouard, *Phonétique lyonnaise au XIV^e siècle*, *Romania* 13 (1884), 542–590.
- REW₃ = Meyer-Lübke, Wilhelm, *Romanisches Etymologisches Wörterbuch*, Heidelberg, Winter, ³1930–1935 [¹1911–1920].
- Souvay, Gilles/Renders, Pascale, *Traitement informatique*, in : Éva Buchi/Wolfgang Schweickard (edd.), *Dictionnaire Étymologique Roman (DÉRom). Genèse, méthodes et résultats*, Berlin, De Gruyter, 2014, 247–257.
- Tournes, Jean de (trad.), *Le Capitaine de Jerosme Cataneo, Contenant la manière de fortifier places, assaillir et defendre [...] mis en françois*, Lyon, Jaques Roussin, 1593.

