DICTIONNAIRE ÉTYMOLOGIQUE ROMAN (DÉROM).

GENÈSE ET ÉVOLUTION DU PROJET

Christoph Groβ,

Mihaela-Mariana Morcov

INTRODUCTION

Dans cet article, nous nous proposons de décrire brièvement le trajet qu'a suivi le DÉRom, en insistant notamment sur l'évolution de la méthodologie et les résultats qu'on a obtenus depuis le moment où ce projet lexicographique a été lancé.

1. GENÈSE DU PROJET

Le projet du *Dictionnaire Étymologique Roman* est présenté pour la première fois en 2007 lors du XXVème Congrès international de linguistique et de philologie romanes d' Innsbruck (Buchi & Schweickard 2010, *cf.* Florescu 2009 : 153), suite aux discussions développées en 1995, autour de la Table ronde *È oggi possibile o augurabile un nuovo REW* (du congrès international de linguistique et de philologie romanes de Palerme). Les débats de 1995 soulignaient la nécéssité d'élaborer un nouveau dictionnaire étymologique dans la mesure où, depuis la publication du REW, plusieurs ouvrages essentiels pour l'étymologie et l'histoire des langues romanes avaient été publiés.

Le 1er janvier 2008, *Le Dictionnaire Étymologique Roman* (DÉRom) est lancé par Éva Buchi et Wolfgang Schweickard, dans le cadre d'une collaboration franco-allemande, financée par la DFG (*Deutsche Forschungsgemeinschaft*) et l'ANR (*Agence Nationale de la Recherche*) et concentrée autour des deux principaux centres de recherche en étymologie romane en Europe : le laboratoire ATILF à Nancy, lieu d'élaboration du *Französisches Etymologisches Wörterbuch*, du *Trésor de la Langue Française* et du projet *TLF-Étym*, et la chaire de linguistique romane de Wolfgang Schweickard, siège du *Lessico Etimologico Italiano* et du *Deonomasticon Italicum*.

2. BUT ET MÉTHODOLOGIE DU PROJET

Dans une première phase, le DÉRom s'est proposé de traiter, à partir d'un inventaire dressé par Iancu Fischer, 488 étymons qui constituent le lexique panroman ou qui ont été continués dans la plupart des langues romanes. Ce matériel linguistique permet au DÉRom d'atteindre son objectif principal, qui consiste à reconstruire l'ancêtre commun des langues romanes ou le protoroman et à en fournir une analyse phonologique, sémantique, stratigraphique et variationnelle sous une forme lexicographique-informatique. Les étymons protoromans postulés constituent les lemmes du DÉRom. Le protoroman a été très peu exploré avant le début du projet en 2008. Les recherches diachroniques en linguistique romane étaient orientées exclusivement vers le latin écrit. On partait de l'étymon latin pour expliquer les lexèmes romans. Cette manière d'analyser le lexique roman caractérise non seulement le REW, mais l'ensemble de l'étymographie prédéromienne.

En revanche, la méthodologie du DÉRom est fondamentalement différente de celle du REW et des autres ouvrages importants de la lexicologie romane.

Le DÉRom propose un nouveau paradigme en étymologie romane, en la refondant sur la base de la grammaire comparée-reconstruction (cf. Buchi & Schweickard 2008 : 352). Cette méthode employée pour la première fois en linguistique romane est déjà couramment appliquée dans d'autres disciplines, par exemple en linguistique indo-européenne, germanique, slave, celtique, bantoue, austronésienne, etc. Le DÉRom applique les principes méthodologiques établis par Jean-Pierre Chambon (2007) et complète pour le lexique les recherches de reconstruction morphosyntaxique du protoroman menées par Robert de Dardel (1996 et 2007, *cf.* Andronache 2010 : 130).

Selon cette méthode, on part des unités lexicales romanes (médiévales et/ou contemporaines) pour reconstruire un ou plusieurs étymons protoromans, en s'appuyant sur les principes de la reconstruction phonologique, sémantique et syntaxique et grâce à la comparaison systématique d'unités romanes du lexique héréditaire commun. La comparaison permet au chercheur de faire une analyse des lexèmes protoromans sur le plan sémantique, morphologique et phonologique. Les recherches du DÉRom fournissent de nouveaux renseignements sur l'histoire et sur la diffusion géographique de différents phénomènes linguistiques (Buchi & Schweickard 2009 : 101). Elles permettent de détecter une stratification du protoroman concernant les formes, les lexèmes ou certains phénomènes linguistiques.

L'intérêt du chercheur se concentre sur les lexèmes de tradition orale (Florescu 2009 : 155). La grammaire comparée-reconstruction cherche à se libérer de la fixation sur la langue écrite.

Étant donné que les lemmes représentent des unités orales reconstruites, ils sont obligatoirement pourvus d'un astérisque. Pour la transcription, on se sert de l'alphabet phonétique international (API, *cf.* Buchi & Schweickard 2010: 63).

La vue panromane du DÉRom agit efficacement contre la fragmentation des disciplines à l'intérieur de la linguistique romane et donne un nouvel élan à la romanistique traditionnelle.

3. ÉVOLUTION ET RÉSULTATS DU PROJET

Des 488 articles projetés au début dans le DÉRom, 73 ont été publiés en ligne (http://www.atilf.fr/DÉRom), d'autres n'attendent que les dernières retouches pour pouvoir être lancés bientôt sur Internet. En outre, le DÉRom a donné lieu à 38 publications et à 68 présentations qui tournent autour des problèmes spécifiques rencontrés lors du processus rédactionnel.

4. STRUCTURE DES ARTICLES

En parcourant le DÉRom, on peut constater que la structure des articles a progressivement évolué, tout en s'adaptant à la complexité et à la diversité des problèmes rencontrés par les rédacteurs. Tout au début, les articles classiques contenaient une seule subdivision suivie d'un paragraphe où les résultats de la reconstruction étaient comparés aux données du latin écrit. À ce type de schéma s'ajoute bientôt un autre qui suppose la présence d'au moins deux subdivisions portant sur la reconstruction phonologique (cas de*/'ɸen-u/ ~ */'ɸɛn-u/), morphologique (cas de */'kad-e-/ ou de*/'pɔnt-e/) ou sémantique (cas de*/'anim-a/) de l'étymon. Dernièrement, le DÉRom s'est enrichi de quelques articles qui prennent simultanément en compte des aspects en rapport avec l'évolution phonétique (ou grammaticale) et sémantique des étymons (*/'ɸamen/,*/'kuer-e-/~*/'kuɛr-e-/, */'lɛβ-a-/ et */'mɛnt-e/).

Les différents types phonologiques ou morphologiques reconstruits sont traités dans des articles séparés s'il s'agit d'unités minimales libres (cas de */as'kʊlt-a-/, */es'kʊlt-a-/, */a'pril-e/, */a'pril-i-u/), mais dans un article unique s'il s'agit de simples variantes (*/'ɛder-a/, */'ɛrb-a/ ~ */'ɛrβ-a/ etc.).

Ce type de structure relève d'une différence importante entre la méthodologie des dictionnaires étymologiques traditionnels et celle du DÉRom. Si dans les premiers, on rattachait à un étymon unique une série entière de cognats romans, dans le DÉRom, les cognats sont souvent mis en rapport avec plusieurs prototypes étymologiques apparentés[1], à même de rendre compte de certaines évolutions successives qui sont intervenues à l'époque protoromane. Dans le REW, par exemple, lat. *hĕrba* est censé expliquer roum. *iarbă*, vegl.

[1]Ces prototypes ont été appelés dans une publication récente du DÉRom (Bastardas i Rufat, Buchi, Cano González 2013: 17) « étimos con continuidà frayada » (étymons à continuité fracturée).

yarba, it. *erba*, log., engad. *erva*, fr. *herbe*, prov., cat. *erba*, esp. *hierba*, port. *herva* etc. En revanche, dans le DÉRom, on reconstruit deux étymons différents : */'ɛrb-a/ et */'ɛrβ-a/. L'étymon */'ɛrb-a/ se reconstruit à partir de dacoroum. *iarbă*, istroroum**.** *iårba*, méglénoroum. *i̯arbă*, aroum. *iarbă*, dalm. *jarba*, istriot. *ǧiérba*, it. *erba*, frioul. *jerbe*, lad. *èrba*, bas-engad. *erba*, fr. *herbe*, frpr. *'ɛrba*, occit.⌜*erbo*⌝, cat. *herba*, tandis que l'étymon */'ɛrβ-a/ se reconstruit à partir d'itcentr./itmérid.⌜*erva*⌝, sard.*èrđa*, haut-engad./surs.⌜*jarva*⌝, gasc.⌜*erba*⌝, l'esp.*hierba*,ast.⌜*yerba*⌝, gal.*herba*/port.*erva*.

Cette diversité peut être observée non seulement au niveau phonologique, mais aussi au niveau morphologique et sémantique.

Au niveau morphologique, on peut évoquer la paire d'étymons */a'pril-e/ et */a'pril-i-u/. L'étymon */a'pril-i-u/ représente un dérivé suffixal à partir de */a'pril-e/ ; il est dépourvu de corrélat en latin écrit. Le REW et le FEW contiennent uniquement l'entrée *aprīlis*, qui correspond au protoroman */a'pril-e/.

Pour ce qui est de la reconstruction sémantique, on constate qu'une partie des sémèmes qu'on a pu reconstruire dans le DÉRom à l'intérieur d'un étymon unique ne connaissent pas de corrélat dans le latin écrit. À part les exemples récemment évoqués par Éva Buchi (*/'anim-a/, */'mɛnt-e/, */ti'tion-e/, *cf*. Buchi 2013), on peut mentionner le cas de l'étymon */'brum-a/, reconstruit dans les sens « hiver », « givre » et « brouillard (surtout brouillard sur mer) ». Le corrélat du latin écrit, *bruma* « hiver », n'est pas enregistré avec les deux derniers sens.

En général, les rédacteurs du DÉRom ont établi également une succession chronologique des principaux types étymologiques reconstruits, en tenant compte des critères suivants : distribution géographique, relation formelle, morphologique ou sémantique entre les bases étymologiques. Dans les rares cas où il est impossible de donner des précisions concernant la succession chronologique, on indique que les types relevés sont « irréductibles en reconstruction » (c'est le cas des étymons */'ɛrb-a/ ~ */'ɛrβ-a/, qui relèvent d'une fluctuation phonématique entre /b/ et /β/) ou que les données dont on dispose ne nous aident pas à éclaircir ce sujet (c'est le cas des sens « crier», « appeler », « proclamer », « nommer » et « s'appeler » que l'on peut reconstruire pour*/'klam-a-/).

5. VARIATION INTERNE

Les résultats du DÉRom démontent le préjugé concernant l'uniformité de la langue reconstruite. Parmi les 73 articles rédigés, il y a peu d'exemples d'étymons uniques. La nature de ces distinctions est elle aussi diverse. On peut parler d'une variation diachronique (*/'ɸak-e-re/ originel vs. une variante syncopée en proclise, */'ɸ-a-re/), d'une variation diatopique (fondée, dans le cas de l'article */'βɪndik-a-/, sur l'opposition entre les sens «

guérir » et « venger »), d'une variation diastratique et à la fois diamésique [les sens « givre » et « brouillard (surtout brouillard sur mer) » de protorom. */'brum-a/ sont spécifiques au langage véhiculé par les marins et les paysans et n'ont pas eu accès au code écrit)] et d'une variation diaphasique (protorom. */'baβ-a/ n'a pas de corrélat dans le latin écrit et appartenait au langage utilisé par les adultes en s'adressant aux enfants).

6. LES ACQUIS DU DÉROM RÉPONDENT À DES PROBLÈMES QUI RESTAIENT INAPERÇUS OU NON RÉSOLUS AU NIVEAU IDIOROMAN

La méthodologie du DÉRom sert à préciser la langue dans laquelle se sont formés les lexèmes romans. Jusqu'à une date récente on pensait que cat. *vinassa* s'était créé en catalan comme dérivé de *vi* (Bastardas i Rufat, Buchi 2012: 26, 29). La démarche comparative aide à comprendre que cat. *vinassa* représente un cognat d' it. *vinaccia*, sard.*vináθθa*, aost. *vinace*, occit. *vinassa*, gasc. *binasso* et participe donc à la reconstruction de protorom. */βi'n-aki-a/ (Bastardas i Rufat, Buchi 2012: 29, *cf.* Delorme 2010–2012 *in* DÉRom *s.v.**/βi'n-aki-a/).

Dans d'autres cas, les résultats de la reconstruction nous permettent de choisir, parmi plusieurs variantes d'un lexème héréditaire, celle qui représente l'issue régulière de l'étymon. C'est le cas de dacoroum. *príer* « avril » (avec l'accent sur l'avant-dernière syllabe) et *priér* « avril » (avec l'accent sur la dernière syllabe). Grâce aux données des autres langues romanes, qui incitent à reconstruire un signifiant paroxyton (*/a'pril-e/), Victor Celac et Éva Buchi ont démontré que l'issue régulière de l'étymon est *príer* et non pas *priér* (Celac, Buchi 2011 : 366).

On peut évoquer encore un autre exemple parlant (qui a déjà été signalé et commenté par Buchi 2010 : 50) en faveur de l'idée que DÉRom peut apporter sa pierre à l'édifice là où l'étymologie idioromane reste impuissante. Pour expliquer l'étymologie d'it. *fieno* «foin», on avait posé plusieurs hypothèses : 1. croisement de protorom. */'ɸen-u-/ s.n. «foin» avec */'ɸlor-e/ s.m. «fleur» (cf. REW3; FEW3, 461a; DELI2) ; 2. «pronunzia rustica *faenum* >*fĕnum*» (Rohlfs 1966 : § 51) ; 3. protorom.*/'ɸen-ul-u/ s.n. «petit foin» > */'ɸlen-u-/ (par métathèse, cf. Rohlfs 1966: § 104 [fĕnulu] et 323 [*foenulum*]). La seule hypothèse valable est la deuxième, car elle s'accorde très bien avec la conclusion qui se dégage à la fin de la démarche comparative : les cognats enregistrés dans les langues romanes nous incitent à reconstruire deux étymons protoromans */'ɸen-u/ et */'ɸɛn-u/ qui remontent à un étymon préprotoroman */'ɸain-u/. Protorom. */'ɸen-u/ et */'ɸɛn-u/ s'expliquent par la réduction de la diphtongue accentuée /–ai-/, qui normalement aboutit, à /-ɛ-/ et plus rarement à /-e-/ (*cf.* MeyerLübkeGLR 1, § 291-292, 637 ; LausbergSprachwissenschaft 1, § 241-242 ; RohlfsGrammStor 1, § 44, 51, 104 ; LloydLatin 105-106).

7. PREMIÈRES DATATIONS

Les recherches du DÉRom ont également permis d'antédater les premières attestations de bon nombre de lexèmes romans. De cette manière, les analyses du DÉRom fournissent également des contributions importantes aux études consacrées à des langues romanes individuelles. Un exemple illustratif en est l'article */'ali-u/, où Jan Reinhardt date cat. *all* depuis 1240, 12 ans avant la première attestation prise en considération par le DECat et par le DCVB (Bastardas i Rufat & Buchi 2012 : 30, *cf.* Reinhardt 2010–2012 *in* DÉRom s.v.*/'ali-u/).

8. LE RÔLE JOUÉ PAR LES DONNÉES DU LATIN ÉCRIT

De ce point de vue, on peut observer également la manière dont la méthodologie employée par les rédacteurs du DÉRom a évolué. Dans une première étape, les étymons traités permettaient uniquement de comparer les données du latin écrit avec les résultats de la reconstruction. La tâche du rédacteur se limitait à chercher et à dater un corrélat latin de l'étymon reconstruit (*/a'gʊst-u/, */'ann-u/, */a'pril-e/, /a'pril-i-u/, */'dɛke/). Cette démarche était vouée essentiellement à injecter les données de la philologie latine dans les articles du DÉRom. Par la suite, les rédacteurs ont commencé, dans certains articles, à tirer explicitement les conclusions de la confrontation des étymons reconstruits avec les données du latin écrit pour la connaissance du latin global[2], en mettant en évidence les types reconstruits qui n'ont pas eu accès au code écrit (cas de */'brum-a/, */'ɛder-a/, */'ɸamen/, */'klam-a-/, */'kuer-e-re/~*/'kuɛr-e-re/, */'laks-a-/), ou inversement les types non reconstructibles, mais connus grâce aux données du latin écrit.

De même, la comparaison avec les données du latin écrit aide aussi à continuer la réflexion étymologique, en soulignant non seulement les rapports existant entre des lexèmes isolés, mais toute une série de transformations et de réorganisations qui ont eu lieu au niveau du système lexical protoroman. Dans l'article */'ɸak-e-/, par exemple, les données du latin écrit et les résultats de la reconstruction mettent en évidence deux phénomènes complexes : d'un côté la neutralisation d'une opposition sémantique attestée pour le latin écrit [entre « faire (une activité considérée dans son exercice continu) » et « faire (une activité considérée à un certain moment) »] à laquelle il est possible d'accéder par le fait que lat. *agere*, qui exprimait l'un des sens opposés, est dépourvu de corrélat en protoroman et, d'un autre côté, la substitution de */'pon-e-/ à */'ɸak-e-/ dans le sens « poser » (Buchi 2009–2013 *in* DÉRom s.v. */'ɸak-e-/). Une confrontation du même type entre la documentation latine et les résultats de la méthode comparative éclaircit le rapport entre protorom. */ka'βall-u/ et lat. *equus* (« lequel

[2]Le concept de « latin global », est défini par Robert de Dardel comme « le latin sous toutes les formes qu'il a pu revêtir depuis la fondation de Rome, dans le monde romain occidental antique et par la suite dans les parlers romans, en Europe, ainsi que dans le latin moderne pratiqué par les sciences et le culte » (Dardel, RLiR 73: 7).

reflète un stade antérieur de la langue parlée ») et représente le point de départ pour « un scénario en trois étapes » qui, selon Ana María Cano González, a eu comme résultat final la diffusion de */ka'βall-a/ « jument » dans une zone limitée de l'Empire Romain (Cano González 2009–2013 *in* DÉRom *s.v.* */ka'βall-u/, */ka'βall-a/).

9. LE PROBLÈME DE LA DISLOCATION ET DE LA PÉRIODISATION DU PROTOROMAN

Même si, dans l'état actuel du projet, il serait prématuré de se prononcer sur ce sujet (voir Buchi & Schweickard 2011 : 305), on devrait retenir quelques aspects qui serviront probablement de base pour l'élaboration d'une hypothèse concernant l'individuation de premières subdivisions à l'intérieur du protoroman. À l'heure actuelle, seules des ébauches de réflexions à ce propos apparaissent dans les articles du DÉRom. On oppose ainsi à un « protoroman *stricto sensu* » (DÉRom s.v.*/'ɸamen/) un « protosarde » ou un « protoroman régional de la Sardaigne » (*/'brum-a/, */'kad-e-/, */'karpin-u/, */sa'gɪtt-a/), un « protoroman continental » (*/'brum-a/,*/'ʊnkt-u/, */'βad-u/, */'βɪndik-a-/, */'ɸen-u/), un « protoroman régional (oriental) » (mentionné seulement dans */'βɪndik-a-/) et un « protoroman nord-occidental » (évoqué jusqu'à présent uniquement dans */βi'n-aki-a/). Une communication récente d'Éva Buchi (Buchi en préparation) tente d'affiner cet aspect du sous-bassement conceptuel et terminologique du projet.

CONCLUSIONS

À partir des données analysées ci-dessus, on peut ébaucher un premier aperçu des tendances générales et des changements qui sont intervenus dans la manière de travail adoptée par les membres du projet. Ces changements se manifestent sous plusieurs formes : ramification des types étymologiques reconstruits, raffinement de l'analyse sur laquelle se fonde la reconstruction phonologique, morphologique et sémantique de l'étymon et la reconsidération du rapport existant entre les résultats de la reconstruction et les données du latin écrit.

La nouveauté de la méthodologie et l'étendue des objectifs poursuivis dans le DÉRom rendent nécessaire un va-et-vient constant entre travail rédactionnel concret et réflexion à la conception du projet tant au niveau théorique qu'au niveau pratique. En tout état de cause, l'intérêt du DÉRom dépasse de loin celui de ses résultats concrets : à notre avis, son apport central réside dans l'enrichissement méthodologique (et méta-méthodologique) qu'il apporte à la linguistique historique en général.

BIBLIOGRAPHIE

Andronache 2010 = Marta Andronache, « Le Dictionnaire Étymologique Roman (DÉRom) : une nouvelle approche de l'étymologie romane ». *Dacoromania* 15, 129-144.

Bastardas i Rufat & Éva Buchi 2012 = María Reina Bastardas i Rufat & Éva Buchi, « Aportacions del DÉRom a l'etimologia catalana ». In : Bürki, Yvette, Cimeli, Manuela & Sánchez, Rosa (éd.) : *Lengua, Llengua, Llingua, Lingua, Langue. Encuentros filológicos (ibero)románicos. Estudios en homenaje a la profesora Beatrice Schmid*, Munich, Peniope, 19-32.

Bastardas i Rufat, Buchi & Cano González 2013 = Maria Reina Bastardas i Rufat, Éva Buchi & Ana María Cano González, « Etimoloxía asturiana ya etimoloxía romance : aportaciones mutues nun contestu de camudamientu metodolóxicu pendiente ». *Lletres Asturianes* 108, 11-39.

Buchi (en préparation) = Éva Buchi, « Qu'est-ce que c'est que le protoroman ? La contribution du DÉRom (*Dictionnaire Étymologique Roman*) ». Otto Winkelmann (éd.): Actes du XXVIIIe Romanistisches Kolloquium (Justus-Liebig-Universität Gießen, 30. Mai - 1. Juni 2013).

Buchi & Schweickard 2008 = Éva Buchi & Wolfgang Schweickard, « Le Dictionnaire Étymologique Roman (DÉRom) : en guise de faire-part de naissance ». *Lexicographica. International Annual for Lexicography* 24, 351-357.

Buchi & Schweickard 2009 = Éva Buchi & Wolfgang Schweickard, « Romanistique et étymologie du fonds lexical héréditaire: du REW au DÉRom (Dictionnaire Étymologique Roman) ». Alén Garabato, Carmen et al. : *La Romanistique dans tous ses états*. p. 97-110. Paris : L'Harmattan.

Buchi, Schweickard 2010 = Éva Buchi, Wolfgang Schweickard, « À la recherche du protoroman : objectifs et méthodes du futur *Dictionnaire Étymologique Roman* (DÉRom) ». Iliescu, Maria / Siller-Runggaldier, Heidi / Danler, Paul: *Actes du XXVe Congrès International de Linguistique et de Philologie Romanes (Innsbruck, 3-8 septembre 2007)*. Berlin / New York : de Gruyter, 61-68.

Buchi & Schweickard 2011 = Éva Buchi & Wolfgang Schweickard, « Sept malentendus dans la perception du DÉRom par Alberto Vàrvaro ». *Revue de linguistique romane* 75, 305-312

Dardel, RLiR 73 = Robert de Dardel (2009), « La valeur ajoutée du latin global », *Revue de linguistique romane* 73, 5-26.

DÉRom = Buchi, Éva & Schweickard, Wolfgang (dir.) (2008–) : *Dictionnaire Étymologique Roman (DÉRom)*. Nancy : ATILF : http://www.atilf.fr/DERom.

Cristina Florescu 2009 = Cristina Florescu, « Limba română în *Dictionnaire Étymologique Roman* DÉRom (< *Romanisches Etymologisches Wörterbuch* REW) ». In :

Botoşineanu, Luminiţa et al. (éd.) : *Distorsionări în comunicarea lingvistică, literară şi etnofolclorică românească şi contextul european*. Iaşi : ALFA / Asociaţia Culturală A. Philippide, 153-159.

Lausberg, Heinrich 1967 (1957), *Romanische Sprachwissenschaft*, vol. 1, *Einleitung und Vokalismus*, Berlin, De Gruyter.

Lloyd, Paul M. 1987, *From Latin to Spanish*, vol. I, *Historical Phonology and Morphology of the Spanish language*, Philadelphie, American Philosophical Society.

Meyer-Lübke, Wilhelm 1890, *Grammaire des langues romanes*, vol. 1 : Phonétique, Paris, Welter.

REW3 = Meyer-Lübke 1930-1935 (1911-1920), *Romanisches Etymologisches Wörterbuch*, Heidelberg, Winter.

Rohlfs, Gerhard 1966 (1949), *Grammatica storica della lingua italiana e dei suoi dialetti, Fonetica*, Turin, Einaudi.