

Marie-Thérèse Kneib

1.8. Le DÉRom expliqué aux lecteurs non spécialistes, mais dotés d'une saine curiosité¹

1 Introduction

Les êtres humains semblent éprouver un besoin naturel de retrouver leurs racines, de comprendre leur culture, d'être fiers de ce qu'ils sont en considérant avec gratitude tout ce qui leur a été donné en héritage. Nous avons tous reçu un trésor, et ce trésor, qui ne vient pas de nous-mêmes – acceptons humblement ce fait : nous ne sommes pas des *self made persons* – compte pour une grande part dans la construction de notre identité, nous donne des clefs pour connaître le réel, l'aimer et le transformer, pour agir en hommes. Pour cette raison, nous avons le devoir de l'aimer, de beaucoup l'aimer, et de le faire fructifier en l'enrichissant nous-mêmes à notre petit niveau, avant de le transmettre à ceux qui nous succéderont : cet héritage est un bien commun et une tradition.

Les langues (ou, pour utiliser la métalangue du DÉRom, qui entend insister sur le fait qu'il ne s'agit pas seulement – et en réalité pas du tout – des langues standardisées,² mais des variétés vernaculaires primaires : les parlers ou les idiomes) tiennent une énorme place dans cette tradition, par cela qu'elles sont à la fois objet transmis et moyen de transmission : transmission d'elles-mêmes, transmission de la connaissance, avec toute la finesse qu'elles peuvent donner à la pensée, transmission même de la délicatesse, qu'elles impriment dans les âmes par les mots qu'elles fournissent, par exemple par les formules de politesse. C'est tout l'homme qui produit le langage, et c'est tout l'homme que le

1 Ce chapitre représente une version amendée de la première partie d'un mémoire de Master intitulé *Mode d'emploi d'un dictionnaire étymologique innovant dans le paysage roman : le « Dictionnaire Étymologique Roman » (DÉRom)*, préparé dans le cadre du European Master in Lexicography (EMLex, cf. <http://www.atilf.fr/emlex>), sous la direction d'Éva Buchi et avec la participation de María Dolores Sánchez Palomino (Université de La Corogne) en tant qu'expert étranger, et soutenu le 26 mai 2016 à l'Université de Lorraine. Nos remerciements les plus chaleureux s'adressent à Wolfgang Schweickard (Sarrebruck) et à Pierre Swiggers (Leuven) pour leurs remarques stimulantes sur une première version de ce chapitre.

2 Pour cette question, cf. Andronache 2013.

langage forme à sa manière. Aimer notre culture et aimer notre langue sont deux choses qui doivent aller de pair.

Comme tout ce qui relève de la tradition, les langues actuelles ne sont pas arrivées jusqu'à aujourd'hui sans connaître une évolution qui est nécessairement influencée par les circonstances et le développement de la pensée humaine. L'histoire des langues, parce qu'elles en sont le reflet, nous dit quelque chose sur l'histoire des hommes et de leurs idées, et voilà pourquoi les langues nous aident à comprendre qui nous sommes aujourd'hui : la linguistique historique (particulièrement l'étymologie) est profondément une science humaine. Nous avons dit que l'amour de notre culture et l'amour de notre langue se répondent mutuellement. Nous ajoutons que l'amour de notre langue ne va pas sans la soif de connaître au moins globalement ses origines.

Les peuples qui parlent des langues romanes ont la chance d'avoir la possibilité de connaître des variétés écrites du latin, langue historique dont chacune est issue, et ce, d'une façon relativement précise et complète si l'on compare le latin aux langues-mères des autres familles de langues, de par les nombreux textes d'époque qui nous sont parvenus. Ces peuples peuvent donc avoir accès à une bonne partie de la culture qui les a moulés. Mais le latin n'est pas l'héritage de ces seuls peuples : il a longtemps été la langue internationale, la langue dans laquelle on parlait d'un pays à l'autre pour se comprendre, la langue des sciences. Il est l'héritage de tous les fils de l'Église catholique, présente sur tous les continents. Voilà ce qui fait de lui une langue digne du plus haut intérêt : ce n'est pas une langue morte et enterrée, de laquelle on veut se souvenir par attachement au passé, c'est un peu la langue de tous les peuples, c'est une langue universelle, en laquelle tout le monde peut retrouver quelque chose qui lui appartient.

De tous temps l'on a trouvé des linguistes et des philologues romanistes qui se sont employés amoureusement à établir le « dépôt » laissé par nos ancêtres, sur lequel se sont greffés tous les enrichissements qui ont contribué à la particularisation des langues romanes, ainsi que des lexicographes qui ont eu le souci de découvrir aux yeux du grand public les résultats de ces investigations. Aujourd'hui encore, de nombreux chercheurs se lancent dans l'aventure. En 2008, un nouveau projet, mis en place à l'initiative d'Éva Buchi et de Wolfgang Schweickard, a vu le jour. Il s'agit de la réalisation d'un nouveau dictionnaire étymologique, le *Dictionnaire Étymologique Roman*, abrégé DÉRom. Est-ce qu'après tous ces siècles de recherches, il resterait encore des choses à découvrir en étymologie romane ? Les personnes impliquées dans le projet en sont persuadées et ont déjà fourni des preuves de la pertinence de leur travail en

apportant à la science de nouvelles connaissances.³ Comment ? En changeant leur approche et leur méthode, en optant pour la reconstruction comparative.

Nous nous intéresserons ici à ce dictionnaire innovant. Ce chapitre a pour vocation de donner quelques clefs qui doivent aider le lecteur à le comprendre. Nous verrons dans une première partie ce qui fait la particularité du DÉRom par rapport à ce qui s'est toujours fait jusqu'ici dans le domaine de l'étymologie romane. Dans une deuxième partie, nous nous pencherons sur la méthodologie qui fait de lui un dictionnaire sans égal en la matière. Enfin, dans une troisième partie, nous l'ouvrirons pour regarder la structure des articles et apprendre à en tirer les informations qu'il veut nous communiquer.

2 Le DÉRom en général : ce qu'il est, pourquoi il est, ce qu'il dit

2.1 Ce qu'il est

Avant même d'ouvrir le DÉRom, le lecteur trouve dans son titre de précieux renseignements sur sa nature et le type d'informations qu'il pourra y trouver.

2.1.1 Il s'agit d'un dictionnaire...

Il faut être bien conscient qu'un dictionnaire ne nous donne pas une langue dans toute la complexité de son système, mais en fournit uniquement le lexique, c'est-à-dire les unités qui pourront ensuite être assemblées en énoncés en suivant des règles de grammaire. Autrement dit, nous ne trouverons dans ce dictionnaire que les « matériaux de construction » des énoncés, sans savoir précisément comment on les assemble. Mais c'est déjà beaucoup ! Rappelons-nous tout ce que les unités lexicales⁴ peuvent nous dire d'une langue. Elles sont des signes, des moyens d'exprimer par des formes perceptibles et particulières les réalités imperceptibles et universelles que sont les concepts, qui donnent un aperçu d'une bonne partie de la vision du monde des locuteurs de la langue en question. Les dictionnaires

³ Des illustrations concrètes de l'apport du DÉRom se trouvent par exemple dans Alletsgruber 2013 (phonologie), Baiwir 2013 (morphologie constructionnelle), Benarroch 2013 (linguistique variationnelle), Buchi 2012 (sémantique) ou encore Maggiore 2014 (morphologie flexionnelle).

⁴ Nous employons *unité lexicale* pour désigner ce qu'on appelle couramment un mot ou une unité plus large qui se comporte comme un mot (par exemple *pomme de terre*).

renseignent sur le sens et la forme des mots, mais aussi (de manière plus ou moins complète) sur leur combinatoire restreinte,⁵ disons sur leur comportement « social » propre avec d'autres unités lexicales : quels mots ou quels types de mots ils admettent dans leur entourage, quels rapports hiérarchiques ils entretiennent avec eux, comment eux-mêmes sont influencés par les autres mots. Entre autres, la mention de l'appartenance à telle partie du discours de l'unité (verbe, nom, adjectif etc.) est une information de cet ordre, ainsi que tout ce qui découle de cette partie du discours. Autrement dit, la combinatoire restreinte est l'ensemble des propriétés syntaxiques d'une unité lexicale. Par exemple, en français, s'il s'agit d'un adverbe, l'on sait qu'il est en principe invariable, contrairement à un verbe, dont la forme est influencée par divers éléments, notamment par des caractéristiques du sujet (personne et nombre), que pourtant il gouverne et dont il requiert obligatoirement la présence.

2.1.2 ... étymologique...

Encore faut-il savoir ce qu'est l'étymologie ! Elle est d'une part (le fait d'établir) un rapport de filiation entre une unité lexicale ou un affixe (notamment un préfixe ou un suffixe) et ce qui est à son origine, d'autre part une discipline qui étudie ce rapport de filiation.

Qui dit étymologie, donc, dit établissement d'un lien entre une unité lexicale (ou un affixe) d'une langue A et une unité lexicale (ou un affixe) d'une langue B (A et B pouvant être confondues),⁶ dont la première tire son origine. L'unité lexicale à l'origine de l'autre est appelée *étymon*.⁷ C'est aussi la mise en lumière du procédé par lequel cette filiation se fait. Y a-t-il hérédité, c'est-à-dire transmission orale de génération en génération de l'unité lexicale depuis une langue ancestrale jusqu'à une autre langue issue de celle-ci (ce qui se traduit par une transformation de l'étymon par une lente évolution régulière) ? Ou bien est-ce un emprunt à une autre langue ? Ou encore s'agit-il d'une unité construite à l'intérieur de la langue en question, d'après ses règles propres de créa-

5 Nous nous plaçons dans le cadre conceptuel de la Théorie Sens-Texte, dont nous empruntons les termes et parfois la typographie (cf. n. 19), et à laquelle Mel'čuk 1997 et Polguère 1998 offrent une courte introduction.

6 Les langues A et B sont confondues dans le cas des créations internes : par exemple, on peut relier fr. *montagneux* à fr. *montagne* parce que le premier est le dérivé du second par ajout d'un suffixe.

7 Par exemple, l'unité française *troïka* a été empruntée au russe : l'unité russe тройка (*trojka*) est l'étymon de l'unité française *troïka* (cf. Buchi 2010a, 488–491, 610).

tion lexicale (cf. Buchi 2016, 340–344 pour ces trois classes étymologiques) ? En principe, on considère l'unité d'« arrivée » et on cherche son origine.⁸ De cela il découle que s'il est possible de consigner dans le dictionnaire étymologique quasiment tous les étymons des issues connues dans la langue (si l'on veut que notre dictionnaire soit exhaustif), il est au contraire absurde de faire figurer le lexique complet de la (ou des) langue(s) fournisseuse(s), dont certaines unités ne seraient étymon de rien du tout. Même dans le cas où l'on voudrait se concentrer sur l'étymologie des unités d'une langue (comme le français) héritées de sa langue-mère, l'on ne fournirait pas le lexique tout entier de la langue-mère : l'on ne peut pas chercher l'origine de mots qui n'ont pas été transmis et donc n'existent pas dans la langue-fille.

2.1.3 ...roman

De quoi fait-on l'étymologie ici ? Du lexique non pas d'une langue, mais d'un ensemble de langues appartenant à une même famille, parce que descendant d'une langue ancestrale commune, en l'occurrence, pour les langues romanes, le latin parlé (ou, pour utiliser un terme scientifique plus précis, le protoroman). Parmi elles, on compte toutes les variétés du roumain, celles de l'italien, celles du français, celles de l'espagnol, celles du galégo-portugais etc. Fera-t-on l'étymologie de chaque unité lexicale de chacun des parlers romans ? Le titre, qui met l'accent sur la communauté d'un héritage (on n'a pas retenu un titre comme **Dictionnaire étymologique des langues romanes*), fait comprendre plutôt que l'on s'intéresse aux unités que ces langues ont (potentiellement) en commun. Plus précisément, le DÉRom fera l'étymologie des unités qui sont partagées par les idiomes romans et en tant qu'elles sont héritées d'une langue ancestrale commune : c'est le lexique héréditaire seul qui intéresse le DÉRom.

⁸ On peut rencontrer des dictionnaires qui présentent les données dans le sens unité à expliquer → étymon (c'est le cas le plus répandu) comme des dictionnaires qui les présentent dans le sens étymon → unité à expliquer, c'est-à-dire que dans ce dernier cas, l'on part de l'étymon et l'on montre ce qu'il est devenu (généralement il s'agit de continuateurs héréditaires, appelées *issues*, et/ou d'emprunts). Le TLFi constitue un exemple classique du premier cas de figure, tandis que le second peut être exemplifié par le *Wörterbuch deutscher Entlehnungen in den Sprachen des Südpazifiks* (« Dictionnaire des emprunts à l'allemand dans les langues du Pacifique Sud »), un projet de l'IDS Mannheim (cf. Engelberg 2006), ou encore par le dictionnaire des emprunts au russe dans les langues romanes (Buchi 2010a).

2.2 Pourquoi il est : lacunes structurelles de l'étymologie romane incarnée par le REW₃

« In memoriam Wilhelm Meyer-Lübke (1861–1936) » (Buchi/Schweickard 2014a, I) constitue la dédicace du premier volume du DÉRom. Le *Romanisches Etymologisches Wörterbuch* (REW₃ : le dictionnaire étymologique des langues romanes qui remonte à leurs sources écrites), l'œuvre de ce grand romaniste des 19^e et 20^e siècles, a émerveillé plusieurs générations par l'ampleur du travail qu'il sous-entend ainsi que par la grande richesse de la documentation et des résultats qu'il présente. Que peut apporter le DÉRom à un dictionnaire extraordinaire, qui, de plus, couvre une bien plus grande partie du lexique des langues romanes que le DÉRom (qui, en plus du lexique héréditaire, traite les emprunts relativement anciens) ? C'est que le REW₃ n'est pas sans lacune ! Il pêche sur un certain nombre de points. Parmi eux, nous développerons surtout celui qui a amené la naissance du DÉRom – qui se propose une étymologisation au niveau du protoroman, dans le but de rendre compte du lexique transmis aux parlers romans – pour refonder l'étymologie romane sur une tout autre méthodologie. D'autres points seront mentionnés plus généralement.

2.2.1 Le problème de la langue-mère et le choix de la méthode

Si l'on veut faire de l'étymologie dans le domaine du lexique héréditaire, il est indispensable de retrouver la langue-mère qui constitue l'ancêtre commun des parlers de la famille linguistique en question. Nous savons que les langues romanes sont héritières du latin. Mais de quelle(s) variété(s) du latin ? Toute langue naturelle présente une variation interne, c'est-à-dire des différences d'usage selon l'époque, le lieu, la situation de communication, l'identité du locuteur et son appartenance à différentes communautés. En particulier, il peut exister une énorme différence entre la langue spontanée et la langue normée. Souvent l'on associe la langue normée à l'écrit et la langue spontanée à l'oral. Bien que l'on puisse s'efforcer de parler dans un langage châtié ou s'autoriser à écrire dans un langage peu empreint de normes, nous opterons pour cette simplification, en raison du fait que l'écrit était un art que peu de gens possédaient dans l'Empire romain, réservé à une élite, et employé surtout pour des actes officiels de justice et d'administration, et pour les sciences et les arts. Par définition, l'écrit est (assez) fixe. Au contraire, l'oral, plus libre, est plus facilement sujet à évolution. Il peut y avoir un décalage entre les deux codes d'une langue, on le voit bien lorsqu'il faut de temps en temps « mettre à jour » la graphie pour

rendre compte de l'évolution. Il en est de même pour le latin⁹ et pour les langues romanes : si ces dernières ont évolué par rapport à la langue-mère au cours des siècles, chacune de son côté (selon les circonstances, les événements, les contacts avec d'autres langues), et à ce point, c'est parce que c'est oralement que la langue s'est transmise. En domaine roman, la langue-mère correspond aux variétés orales spontanées du latin (que l'on subsume quelquefois de façon impropre sous le terme de latin vulgaire), ou plus précisément, elle est une portion de ces variétés orales, puisque l'on peut supposer que le lexique n'a pas été transmis dans son entier aux langues romanes. Bien entendu, il n'existe pas de clivage entre le latin normé et le latin spontané ; le lexique de l'un et l'autre se chevauchent en grande partie.¹⁰

Du fait de l'abondance des sources du latin écrit, la méthodologie suivie par tous les romanistes jusqu'à très récemment était la « chasse » aux attestations latines, c'est-à-dire aux occurrences d'unités lexicales dans les textes (ou plutôt dans les dictionnaires, qui s'appuient sur les textes) latins, qui nous prouvent leur existence dans la langue. Ces attestations fournissaient les étymons (cf. Chambon 2010, 62–65), et, dans le cas où l'étymologiste trouvait une forme qui ne permettait pas d'expliquer celle des unités romanes supposées la continuer, il « bidouillait » l'unité du latin écrit pour qu'elle se rapproche de la forme attendue (cf. Buchi 2010b, 2). Ce semblant de reconstruction était une béquille, un plan de secours. Ainsi, le REW₃, dans cette même optique, présente systématiquement une unité appartenant à l'écrit (latin normé), excepté lorsqu'aucune trace écrite de l'unité lexicale n'est retrouvée, ou lorsque les unités romanes ne correspondent pas à ce qui est attendu compte tenu des règles d'évolution des sons, auquel cas il propose une forme hypothétique « bricolée » en guise d'étymon (ou en plus de l'étymon attesté dans les textes), à laquelle il adjoint un astérisque pour marquer l'absence de donnée à l'écrit. Bien que les deux

9 Les fondateurs du DÉRom avancent : « nous avons toutes les raisons de croire qu'il y avait, au moins dès le temps de Cicéron et d'Auguste, un décalage entre les codes parlé et écrit du latin (donc, une sorte de diglossie) » (Buchi/Schweickard 2009, 100).

10 Le latin oral et le latin écrit ne constituant jamais que deux codes d'une seule langue, ils n'ont à eux deux qu'un seul lexique. On peut dire la même chose pour toutes les langues qui connaissent un code écrit. Mais si l'on prend l'exemple du français, personne ne peut nier qu'il existe une différence assez notable entre l'oral et l'écrit, et pas seulement au niveau de la tournure des phrases. Certaines unités lexicales sont utilisées de façon beaucoup plus fréquente à l'oral (par exemple *truc*, ou encore *on* pour la quatrième personne), certains au contraire ne sont utilisés pratiquement qu'à l'écrit (comme *car* ou *en outre*). Dans un souci de simplification de la pensée et du discours, on peut parler d'un lexique de l'oral et d'un lexique de l'écrit ; c'est dans cet esprit de simplification que nous avons opéré une distinction entre lexique du latin oral et lexique du latin écrit.

lexiques (oral et écrit) soient en grande partie les mêmes, pour cette autre partie où l'on a des différences, et aussi parce que ce « bidouillage étymologique » ne peut pas prétendre à être qualifié de scientifique, cette façon de procéder pose problème, l'étymologie est faussée.

Pourquoi faussée ? On pourrait être tenté de dire au contraire qu'elle est plus parfaite en tant qu'elle nous fait remonter plus loin dans le temps, en présentant un état du latin où l'écrit traduisait tout à fait ce qu'était l'oral, du moins en ce qui concerne le lexique. Mais non, les étymologistes comprennent de mieux en mieux que sauter les étapes en négligeant de chercher les étymons directs (c'est-à-dire les étymons les plus récents, les plus proches de l'issue à expliquer) est source d'erreurs¹¹ et ne permet pas de décrire l'étymon précisément sous toutes ses dimensions (forme, sens et combinatoire restreinte). Pour un travail scientifique, à l'étymologie lointaine, qui cherche les étymons des étymons en ignorant les étapes intermédiaires, il faut préférer l'étymologie proche, puis remonter de proche en proche jusqu'au stade antérieur voulu.

Le schéma suivant tend à mettre en évidence le décalage entre les résultats que l'on peut obtenir en optant, comme le DÉRom, pour l'étymologie proche avec la reconstruction des variétés orales, ou, comme les ouvrages d'étymologie dite « traditionnelle » représentés par le REW₃, pour l'étymologie lointaine en cherchant essentiellement des attestations dans les textes.

¹¹ Dans son ouvrage *Galerie des linguistes franc-comtois*, Jacques Bourquin dit à ce propos, après avoir présenté des extraits d'un compte rendu du dictionnaire étymologique de Charles Toubin, que « la manie de l'étymologie lointaine jointe d'ailleurs à la méconnaissance des lois d'évolution phonétique [...] aboutit le plus souvent à des rapprochements fantaisistes ou fantastiques » (Bourquin 2003, 205).

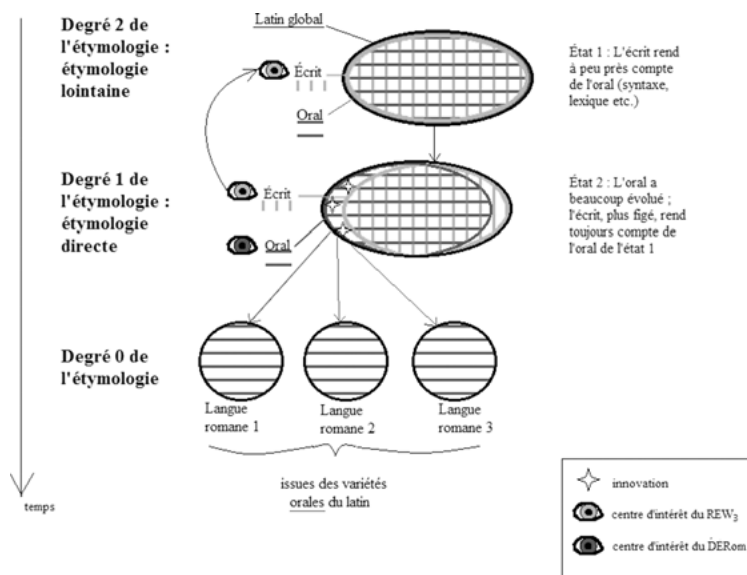


Figure 1 : Objets de focalisation du REW₃ et du DÉRom

Récapitulons. En toute rigueur scientifique, il faudrait opter pour l'étymologie proche. Pour cela, nous identifions comme langue-mère des langues romanes une portion des variétés orales spontanées du latin. C'est donc là que nous trouverons nos étymons. Mais comment ? La parole orale n'a pas été enregistrée en ce temps-là ! C'est vrai, mais elle a laissé ses traces dans les langues-filles ; on peut la reconstruire, comme on reconstruit déjà la langue ancestrale d'autres familles de langues depuis le 19^e siècle (cf. Baxter 2002). La langue-mère, la partie du latin reconstituée par la comparaison des langues romanes dans lesquelles elle laisse des traces, est appelée *protoroman* ;¹² le DÉRom est né pour le faire ressurgir. Alors que la reconstruction était une « roue de secours » pour combler des lacunes que la chasse aux attestations laissait inévitablement, elle devient première, et le chercheur n'a recours aux sources du latin écrit (littéraire et non littéraire) que lorsque la méthode de la reconstruction est à elle seule impuissante à trancher entre plusieurs hypothèses, et surtout pour dégager

¹² En réalité, il existe deux sens de *protoroman*. Dans ce chapitre, nous opérons une simplification en ne donnant que le plus prégnant dans le DÉRom. Pour le second, et pour les rapports entre l'un et l'autre, cf. Buchi 2015, 4–5.

comparativement la variation interne du latin global.¹³ Dans tous les cas, les données du latin écrit n'interviennent jamais dans le processus de la reconstruction ; elles sont consultées une fois les résultats obtenus. Elles sont tout au plus capables de confirmer ces derniers, mais n'ont aucun pouvoir légitime de les contester (cf. Chambon 2007, 69). Au contraire, le DÉRom peut de droit corriger les étymons (astérisqués ou non)¹⁴ des unités lexicales romanes héréditaires proposées par le REW₃ : il fait autorité dans le domaine de la reconstruction du lexique protoroman.

2.2.2 Autres points problématiques

De la méthodologie traditionnelle, adoptée entre autres par le REW₃, découlent un certain nombre de choix problématiques.

2.2.2.1 Notation de l'étymon

Pour retrouver les unités de la langue à l'origine des parlers romans, nous l'avons dit, il faut reconstruire. Pour reconstruire de l'oral, il faut se fonder sur de l'oral, mais aussi savoir ensuite présenter sous forme écrite¹⁵ les résultats dans le dictionnaire. Les caractères ou les groupes de caractères de l'écriture conventionnelle ne renseignent pas précisément sur la réalité phonique des

¹³ La place des données du latin écrit de l'Antiquité dans le DÉRom est décrite dans Maggiore/Buchi 2014, particulièrement dans la section « Le latin écrit comme bénédiction ».

¹⁴ En effet, « contrairement à ce que l'on aurait pu penser, la déclinaison étymologique frappe d'ailleurs même les étymons astérisqués de REW₃ : la reconstruction comparative aboutit ainsi à */m'prumut-a-/ (Maggiore 2014 in DÉRom s.v.) là où le REW₃, fidèle à sa logique de « <fiddled with> classical Latin » [...], porte *imprōmūtūare » (Buchi/Schweickard 2014b, 16).

¹⁵ À notre connaissance, il n'existe pas (encore) de dictionnaire aux entrées duquel on accéderait d'une manière tout à fait orale, c'est-à-dire en prononçant les unités lexicales cherchées. En revanche, il existe plusieurs moyens de refléter à l'écrit la phonie des unités lexicales et de se servir de cette représentation pour l'accès aux entrées du dictionnaire, soit que l'on prenne des graphèmes dont la prononciation est univoque comme représentant des sons d'une langue particulière (ainsi la saisie phonétique du TLFi), soit que l'on opte pour un système de représentation admis internationalement. Un dictionnaire aux entrées duquel on accéderait de manière orale ne pourrait sans doute pas se passer d'une transcription écrite comme seconde structure d'accès aux articles pour éviter certains inconvénients, comme celui de ne pas avoir un aperçu rapide de la couverture lexicographique (l'ensemble des unités traitées par rapport au sujet du dictionnaire), à cause de l'inscription dans le temps de la réalisation phonique du langage.

unités lexicales. Certaines, même, peuvent ne correspondre à aucun son (comme le <p>¹⁶ de <sculpture> en français), d'autres peuvent représenter des sons potentiels qui vont trouver une réalisation uniquement dans des contextes précis (comme les marques du pluriel <s> ou <x> en français, entendues dans les liaisons seulement). Quoi qu'il en soit, il est préférable de ne pas s'appuyer sur la graphie conventionnelle d'une langue, en particulier pour noter les résultats de la reconstruction.

Certains romanistes prétendent qu'il serait possible et même préférable d'employer le système graphique latin, en tirant parti de la régularité (prétendue) des correspondances entre les caractères (ou groupes de caractères) et les sons.¹⁷ Mais le choix de la graphie latine conventionnelle pose problème lorsque l'on considère entre autres le fait que certains sons du latin classique (naturellement représentés à l'écrit) disparaissent en protoroman, alors que leurs représentants graphiques, eux, subsistent. Par exemple la lettre <-m> en fin de nom commun n'est plus le reflet d'aucune réalité phonique en protoroman (Buchi/Schweickard 2011b, 631). La notation adoptée par le DÉRom veut lever toute ambiguïté sur ce plan-là. En outre, le DÉRom, qui est le fruit d'un travail scientifique, tient à être explicite dans la présentation des résultats de sa recherche. D'autre part, il veut pouvoir s'adresser aux non-romanistes et aux non-latinistes, donc à des personnes qui n'ont pas le bagage de connaissances suffisant pour extraire les traits phonologiques à la simple lecture (Buchi 2014, 404), tâche en soi ardue, d'autant plus que chez les spécialistes du domaine eux-mêmes, ces traits phonologiques sont en réalité bien méconnus (cf. Garnier 2015, 235).

Pour toutes ces raisons, les unités lexicales du protoroman sont écrites phonologiquement dans le DÉRom, c'est-à-dire dans un système d'écriture où chaque signe correspond à un « son » (ou, plus précisément, à une réalité plus abstraite appelée *phonème*, cf. ci-dessous 3.1.5) et où chacun de ces phonèmes a son signe, là où le REW₃, même pour les unités marquées par un astérisque, donne une écriture conventionnelle.

2.2.2.2 Informations superflues ou au contraire manquantes

Le REW₃ d'un côté apporte des informations peu pertinentes pour l'étymologie romane, comme la quantité (longueur) des voyelles (qui devrait être remplacée

¹⁶ Nous notons les unités graphiques, appelés *graphèmes*, ainsi que les suites d'unités graphiques, entre chevrons : <...>.

¹⁷ Ainsi, Johannes Kramer affirme : « parce que l'orthographe latine est de nature phonologique, on pourrait l'employer pour les éléments reconstruits sans perdre aucune information » (Kramer 2014, 292).

par des précisions sur leur timbre), et d'un autre côté en oublie certaines qui, elles, sont très importantes, comme l'accent tonique, ce trait qui, dans certaines langues (ainsi en allemand, en anglais ou encore en protoroman, mais pas en français, par exemple), marque une syllabe donnée dans le mot.

Le latin littéraire faisait la distinction, phonologiquement pertinente, entre les voyelles longues et les voyelles brèves. Ainsi, *venit* '(il) vient' (indicatif présent), avec /e/ bref, et *vēnit* '(il) vint' (indicatif parfait), avec /e/ long, n'étaient pas confondus par les locuteurs de l'époque classique, qui s'appuyaient sur leur perception de la durée relative des sons (quantité) pour les différencier. Au contraire, tout porte à croire qu'en protoroman, la langue orale à l'origine des parlers romans, la longueur des voyelles n'avait aucune importance distinctive. En revanche, les locuteurs faisaient une différence entre les voyelles selon qu'elles étaient produites en ouvrant plus ou moins la bouche, comme la différence entre les voyelles [ɛ] (comme dans fr. *près*) et [e] (comme dans fr. *pré*).

En ce qui concerne l'accent tonique lexical, si quasiment toutes les langues-filles (à la seule exception du français) en portent un, et au même endroit dans l'unité lexicale, c'est que le protoroman en avait un aussi. On ne peut pas omettre ce point sans risquer de perdre la possibilité d'expliquer certaines évolutions phonétiques. Pour illustrer l'importance de l'accent, voyons quelques règles d'évolution des voyelles entre le protoroman et le français.

Les voyelles portant l'accent (appelées *voyelles toniques*) en protoroman évoluent plus ou moins fortement selon leur place dans la syllabe : en fin de syllabe, elles évoluent beaucoup ; à l'intérieur de la syllabe, elles évoluent plus faiblement ; mais jamais elles ne disparaissent (Bourciez/Bourciez 1967, 31–32, 55–59). Ainsi, protorom. /'a/ en syllabe fermée (c'est-à-dire qui se termine par une consonne) se maintient en tant que /'a/ en français :¹⁸ protorom. /'part-e/ (Velasco 2011–2014 in DÉRom s.v. ; correspondant de latin classique *partem*) devient fr. /'pav/ (<part>). En revanche, lorsqu'il se trouve en syllabe ouverte (c'est-à-dire qui se termine par une voyelle), protorom. /'a/ évolue vers fr. [ɛ] : protorom. /'patr-e/ (correspondant de latin classique *patrem*) devient fr. /'pɛv/ (<père>).

La voyelle de la syllabe précédant celle qui porte l'accent ne connaît pas le même sort. Dans cette position, protorom. */a/ évolue systématiquement pour donner /ə/ en français. C'est le cas pour protorom. */orna'ment-u/ (correspondant de latin classique *ornamentum*), qui donne fr. /ɔʁnə'mɑ̃/ (<ornement>).

¹⁸ On note les unités qui sont des phonèmes entre barres obliques : /.../. Pour ce qui est de l'accent tonique, il est marqué au moyen de l'apostrophe droite (') posée avant la syllabe accentuée. Nous avons pris l'initiative ici de mettre en relief la syllabe accentuée au moyen de caractères gras et de souligner la voyelle qui nous intéresse.

Toujours dans cette position, mis à part */a/, dont nous avons déjà commenté le sort, n'importe quelle voyelle se maintient en principe à condition qu'elle se trouve devant un groupe de consonnes : ainsi protorom. */guber'nak-l-u/ (*e/ devant le groupe de consonnes */rn/ ; correspondant de latin classique *guber-naculum*) devient fr. /guber'*na*j/ (<gouvernail>). Derrière un groupe de consonnes, la voyelle évolue, comme */a/, en /ə/: protorom. */kapri-'*ḫ*oli-u/ (*i/ derrière le groupe de consonnes */pr/ ; correspondant de latin classique *capri-folium*) aboutit à fr. /ʃevʁə'fœj/ (<chèvrefeuille>). Dans tous les autres cas, la voyelle précédant directement la syllabe tonique disparaît, comme pour protorom. */kiβi'tat-e/ (correspondant de latin classique *civitatem*), qui donne en ancien français *ciptet* (puis en français moderne *cit*é).

2.2.2.3 Le statut attribué à l'étymon

Nous nous sommes tous déjà prêtés à des petits jeux dans lesquels nous devons deviner ou faire deviner des mots, ou encore trouver le plus possible de mots répondant à une consigne donnée. Nous les donnons (ou nous les attendons) toujours sous la même forme. Par exemple, nous ne disons jamais *chevaux*, ni *mangerai* ni *blanche*, non, nous disons *cheval*, *manger*, *blanc*. C'est parce que nous donnons en quelque sorte le « nom » du mot, pour englober dans notre formulation toutes ses formes, pour le donner tout entier. L'unité CHEVAL¹⁹ contient deux sous-formes (qu'on appelle *mots-formes*), correspondant au singulier et au pluriel : *cheval* et *chevaux*. L'unité MANGER en contient plusieurs dizaines, correspondant aux combinaisons de personne, nombre, temps et mode : *mange*, *manges...*, *mangerai*, *mangeras...*, *mangeai*, *mangeas...*, *ai mangé* etc. L'unité BLANC en contient quatre, correspondant à la combinaison de genre et de nombre : *blanc*, *blanche*, *blancs* et *blanches*. Lorsque nous cherchons un mot dans le dictionnaire, c'est encore sous son « nom » que nous espérons le trouver, et pas sous un de ses mots-formes, même si effectivement, c'est la forme perceptible d'un des mots-formes que l'on utilise pour désigner l'unité dans son ensemble (encore appelée *lexie*) : c'est la forme citationnelle. Conventionnellement, en français, on désigne le verbe par son infinitif, le nom, par le singulier, l'adjectif, par le masculin singulier.

Faire l'étymologie d'une lexie, c'est-à-dire d'une unité lexicale tout entière,

¹⁹ En conformité avec la Théorie Sens-Texte (cf. n. 5), nous notons en petites capitales les unités lexicales abstraites, pour les différencier des mots-formes, leurs réalisations possibles, signalés par les caractères italiques. Nous ne tenons pas compte ici de la distinction entre vocables et unités lexicales (ou lexies), et ne numérotions donc pas les unités lexicales que nous citons.

avec toutes ses formes (rassemblées sous la forme citationnelle), c'est retrouver l'étymon qui est une lexie, donc une unité lexicale tout entière, avec toutes ses formes (rassemblées sous la forme citationnelle). On ne peut pas, comme le fait par exemple le REW₃, mettre un mot-forme comme étymon d'une lexie si l'on veut prétendre à un travail scientifique. Cela ne fait aucun doute que son illustre auteur reconnaissait le problème que posait cette façon de représenter l'étymon : ce qui justifie sa pratique, ce sont de fortes contraintes lexicographiques et le poids de la tradition. On comprend que la tâche de représenter l'étymon dans son entier et non juste une de ses formes ne soit pas simple. En effet, si en français, il est de convention, comme nous l'avons déjà mentionné, que les lexies verbales soient représentées par leur forme infinitive, soit un seul mot-forme, en latin, elles sont représentées par l'ensemble de leurs temps primitifs, c'est-à-dire par plusieurs mots-formes. Ainsi le verbe latin signifiant 'donner'²⁰ doit être nommé ainsi : *dō*, *dās*, *dare*, *dedi*, *datum* (cf. Gaffiot s.v.) ; c'est sous cette forme que nous devons le chercher dans le dictionnaire. Nous constatons la même chose pour les noms, identifiés par leurs mots-formes au nominatif singulier et au génitif singulier (il en est ainsi pour *rosa*, *-ae*, 'rose') ou encore pour les adjectifs (*magnus*, *-a*, *-um*). Le DÉRom, nous le verrons plus loin, apporte une solution à ce problème en réunissant les mots-formes sous une seule forme citationnelle qui ne laisse pas de place à l'ambiguïté par rapport à son statut de lexie.

2.2.2.4 Manque de rigueur scientifique

L'approche traditionnelle telle qu'elle est représentée de façon prototypique par le REW₃ a encore d'autres imperfections que le DÉRom tâche de corriger. Dans son compte rendu du premier fascicule du REW₃, Dauzat (1911, 425) met le doigt sur le fait que les sources pour la collecte des unités dialectales ne sont pratiquement jamais indiquées, ce qui rend la critique impossible. Il déplore également le fait que l'auteur pose comme des affirmations des hypothèses dont le contenu est discutable (et discuté à l'époque même de Meyer-Lübke), sans donner de justification.

2.3 Ce qu'il dit

Le DÉRom donne sous une forme lexicographique les résultats des recherches obtenus par la méthode comparative, à savoir la forme, le sens et la combina-

²⁰ Une description du sens est donnée par une suite '...'.

toire restreinte de chaque étymon. Mais ce n'est pas tout. Idéalement, par la confrontation systématique des résultats de la reconstruction avec les corrélats latins (les unités latines qui sont les correspondants les plus proches des unités protoromanes reconstruites, autant au niveau de la forme que du sens et de la combinatoire), la méthode comparative permet de reconstituer la variation interne de la protolange (cf. Buchi/Schweickard 2013), notamment selon quatre axes : la variation dans le temps, la variation dans l'espace, la variation sociale et la variation selon la situation de communication.

2.3.1 La variation dans le temps

Si nous connaissons assez bien l'histoire de la romanisation (« latinisation ») des peuples qui parlent aujourd'hui une langue romane, le processus précis de la fragmentation de la Romania (l'ensemble géographique où ont été parlés/ sont parlés les différents idiomes romans), c'est-à-dire la manière dont s'est faite l'individuation de chaque groupe de parlers et de chaque parler en particulier, reste encore largement à établir. Néanmoins, nous en savons suffisamment (cf. Buchi 2015) pour qu'une datation relative de différents prototypes lexicaux reconstruits soit dans beaucoup de cas possible. On peut en effet établir, selon les données romanes que l'on trouve, la datation relative des unités : si tel ensemble de parlers, dont on sait qu'il s'est individué à telle époque, présente, avec les autres parlers individuels à une date postérieure, des données permettant de reconstruire l'étymon qui nous intéresse, on peut en conclure que cet étymon existait déjà lorsque l'ensemble de parlers en question n'était pas encore détaché des autres branches du protoroman. L'illustration ci-dessous (figure 2) permet, à partir de l'exemple de */'ɸak-e-/²¹ ('faire') (cf. Buchi 2009–2014 in DÉRom s.v.), de fournir un exemple concret. Deux types ont été reconstruits à partir des issues romanes, un type */'ɸak-e-/ et un type */'ɸ-a-/. Le sarde et le roumain, dont on sait qu'ils sont les premiers parlers romans à s'être détachés du tronc commun, présentent des issues du type */'ɸak-e-/ mais aucune du type */'ɸ-a-/. Pour cette raison, */'ɸak-e-/ doit être regardé comme le type le plus ancien, datant d'une époque antérieure à l'individuation des deux langues (ou ensembles de parlers), tandis que le second type est à rattacher à une époque plus récente du protoroman, qui n'englobait déjà plus ni le sarde ni le roumain.

21 Les unités citées ici sont présentées avec la notation du DÉRom, qui sera expliquée un peu plus loin (cf. ci-dessous 4.1.1). L'astérisque marque leur caractère reconstruit.

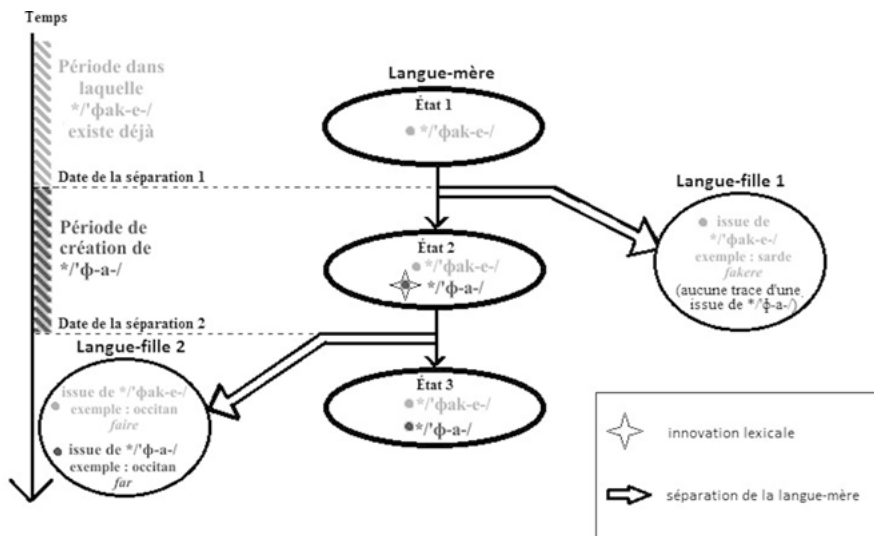


Figure 2 : Datation relative des types **'ʔak-e-/* et **'ʔa-/*

2.3.2 La variation dans l'espace

En fonction de l'identité des parlers romans (ou des groupes de parlers romans) qui offrent des données, on comprend sur quelles aires géographiques telles unités lexicales protoromanes (ou telles variantes d'unités lexicales protoromanes) étaient répandues.

Les connaissances qu'offrent les données sur la variation dans l'espace donnent souvent des indications sur le moment où l'unité lexicale et ses variantes avaient cours dans la protolangue. Par exemple, on reconstruit deux modèles de conjugaison pour le verbe protoroman **'kuer-e-/* (cf. Maggiore 2012–2015 in DÉRom s.v. **'kuer-e-/*) :²² **'kuer-e-/* (infinitif */'kuer-e-re/*) et **'kue'r-i-/* (infinitif **/kue'r-i-re/*). Il est possible de connaître quel type parmi les deux est premier grâce à l'étude de la répartition géographique des issues de chacune. Examinons la carte suivante :

²² Pour aider l'esprit à saisir de quoi il s'agit, nous pouvons nous imaginer que le verbe français signifiant 'chanter' aurait pu avoir deux modèles de conjugaison, par exemple il aurait pu se conjuguer comme un verbe du premier groupe et comme un verbe du deuxième groupe, au choix (avec ou non spécialisation de sens) : *chanter* : *je chante, tu chantes, il chante, nous chantons, vous chantez, ils chantent* ; *!chantir* : *je !chantis, tu !chantis, il !chantit, nous !chantissons, vous !chantissez, ils !chantissent* (les formes inexistantes sont marquées au moyen d'un point d'exclamation non italique antéposé).

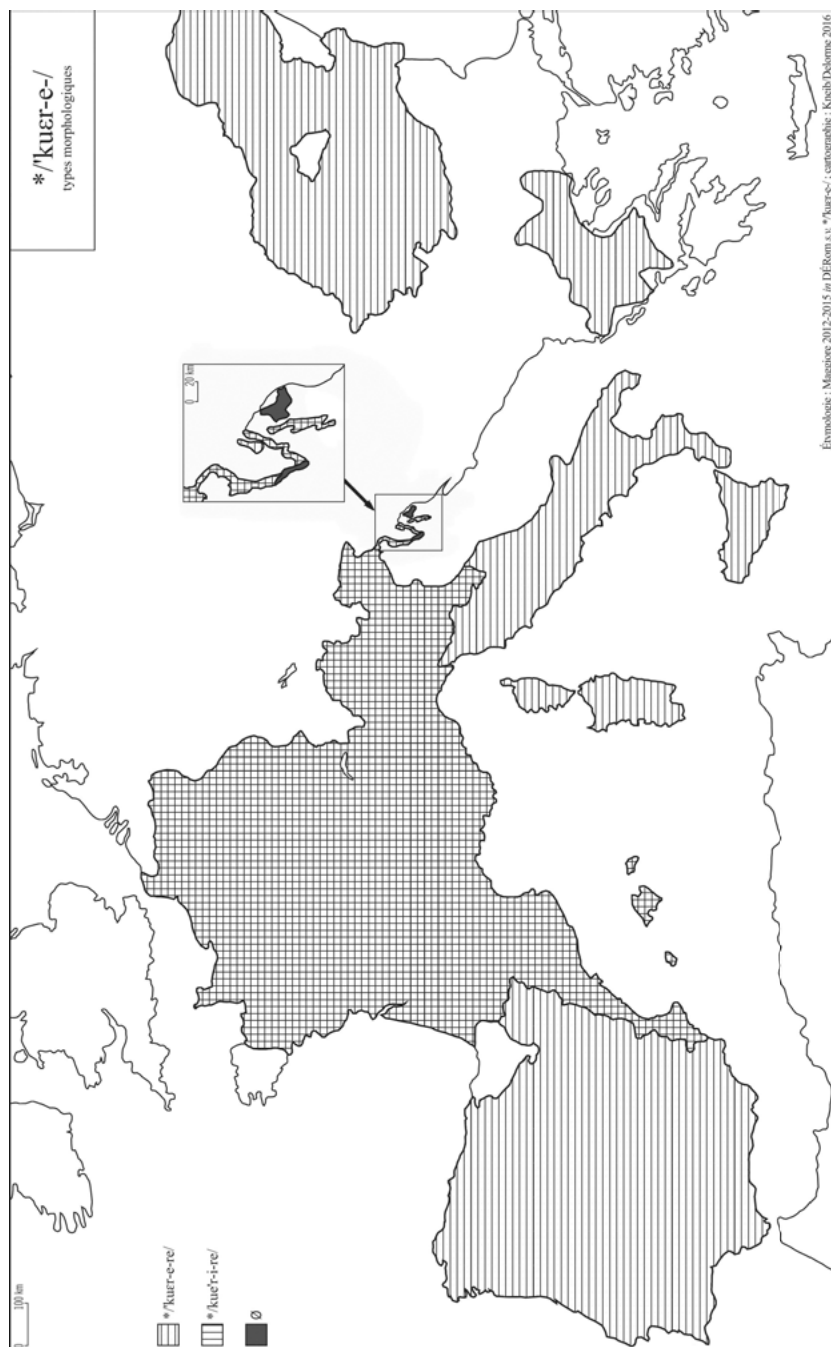


Figure 3 : Répartition géographique des types * /kuer-e- / (infinitif /kuer-e-re/) et * /kue'r-i- / (infinitif * /kue'r-i-re/)

Le premier type, */'kuɛr-e-/ (représenté dans la légende par le mot-forme de l'infinitif */'kuɛr-e-re/), a des issues dans quasiment toute la Romania.²³ Le second type, */'kue'r-i-/ (représenté par le mot-forme de l'infinitif */'kue'r-i-re/), n'a des continueurs que dans une aire compacte, comprise dans la zone où l'on trouve le premier type ; là, les deux coexistent. Qu'est-ce à dire ? Le premier type, répandu partout, était déjà présent avant les premières individualisations de langues hors du tronc commun. Le second type est une innovation plus tardive originaire du centre de la Romania (cf. Maggiore 2012–2015 in DÉRom s.v. */'kuɛr-e-/). La présence du premier type en sarde et en roumain, et l'absence du second dans ces mêmes ensembles de parlers, le confirment (cf. Maggiore 2014).

2.3.3 La variation sociale et la variation selon la situation de communication

Le principe de la reconstruction est de reconstituer les traits les plus persistants de la protolange à partir des traces qu'ils ont laissées dans les langues-filles. S'il se trouve que plusieurs variantes d'une même lexie héréditaire ont un statut différent (par exemple, une appartient à la variété haute, la plus prestigieuse, c'est-à-dire la langue soutenue, et une autre, à la variété la plus basse, la langue familière) dans un certain nombre de langues romanes, l'existence de ce trait commun doit être interprétée comme une conséquence de sa présence en protoroman. Mais de tels cas sont rarissimes. Une autre manière de reconstruire la variation interne d'une protolange selon la situation de communication consiste à s'appuyer sur des séries. Un cas très intéressant de la mise en évidence de la variation sociale du protoroman de ce type est analysé dans l'article */'laks-a-/ (Florescu 2010–2016 in DÉRom s.v.), où deux variantes phonologiques ont été reconstruites, */'laks-a-/ et */'laks-i-a/ (à côté d'un troisième type, */'daks-a-/). Le phénomène de création lexicale d'un verbe à partir d'un autre verbe par l'ajout de l'affixe */-i-/ entre le radical et le morphème (indiquant la classe) de conjugaison n'est pas rare en protoroman. Il a été observé que cet affixe imprègne la lexie d'une connotation qui la fait rattacher à une variété basse, par opposition à sa variante non affixée, qui appartient plutôt à la variété haute (cf. de Dardel 2006, 393–394). Dans le cas de */'laks-a-/ ~ */'laks-i-

²³ Dans la carte, seules les zones remplies d'un figuré font partie de la Romania. Le gris foncé figure les territoires romans dans lesquels aucune issue héréditaire n'a été trouvée. Les zones territoriales restées blanches correspondent à des aires linguistiques non romanes (pour une carte des idiomes romans, cf. ci-dessous 4.2.1).

a-/ , les issues de l'un et l'autre type sont quasiment en distribution complémentaire, c'est-à-dire que tous les parlers présentent des issues se rattachant soit à l'un soit à l'autre, mais pas aux deux (excepté l'italien central). Pour cette raison, si l'étude s'arrête à cette unité seule, elle ne donne aucune lumière sur la variation sociale. C'est en élargissant notre vision aux dimensions du lexique protoroman que l'on peut affirmer quelque chose. Ici, la reconstruction de la connotation de la lexie due à la présence de l'afixe */-i-/ se fait en s'appuyant sur les cas analogues observés antérieurement.

La variation sociale et la variation selon la situation de communication sont très souvent liées entre elles. L'être humain étant profondément social, il cherche à s'intégrer en s'exprimant de la même manière que les autres dans une situation donnée. Généralement, les personnes des classes plus basses fréquentent peu les milieux plus prestigieux qu'accompagne nécessairement un style de langue plus relevé. Une personne peu instruite n'aura pas la même aisance à adapter son discours selon la situation de communication qu'une personne ayant bénéficié d'une éducation poussée, qui maîtrisera un plus grand éventail de styles.

2.3.4 Un exemple

Les connaissances nouvelles apportées par le DÉRom sont indéniables : on peut le voir par la comparaison entre les analyses proposées par le REW₃ d'une part et celles proposées par le DÉRom de l'autre. Les résultats de l'enquête sur l'étymon de fr. FAIM et ses correspondants héréditaires dans les autres parlers romans (cf. Buchi/González Martín/Mertens/Schlienger 2012–2015 in DÉRom s.v. */'ϕamen/ ; 2015) nous serviront d'exemple. La comparaison met en évidence un gain de connaissance à plusieurs niveaux :

- pour la forme, puisqu'on reconstruit une forme un peu différente (*/'ϕamen/) de celle qui est attestée dans les textes latins (*fames*), qui correspond à un prototype plus récent (*/'ϕam-e/, cf. */'ϕamen/ II.).
- pour le sens, puisque ce ne sont pas moins de trois sens qui sont reconstruits ('faim', 'famine' et 'désir'), ce qui met en évidence la polysémie de l'étymon (c'est-à-dire le fait, pour cet étymon, de représenter plusieurs lexies, aux sens différents mais liés), tandis que le REW₃ n'en donne qu'un ('faim') : ce sont de véritables vocables²⁴ que le DÉRom reconstruit ;

²⁴ Dans le cadre de la Théorie Sens-Texte, un vocable est un groupe d'une ou de plusieurs lexies de même forme et de sens différents mais liés.

- pour la combinatoire restreinte, puisqu'on reconstruit pour ce nom commun le genre neutre, là où le corrélat en latin écrit est féminin.

Alors que le REW₃, qui incarne la méthode d'étymologie traditionnelle, se donne l'objectif de relier un étymon (lointain) déjà connu à des issues déjà connues (et par là, ce qui est découvert après l'enquête selon cette méthode est le lien entre les unités), le DÉRom, par la reconstruction, peut découvrir de nouveaux éléments du lexique protoroman. Voilà en quoi le changement est intéressant.

2.4 Limites de la méthode

Si la méthode comparative apporte des connaissances que la méthode traditionnelle ne peut pas fournir, les linguistes impliqués dans le projet DÉRom savent bien qu'à elle seule, elle n'est pas propre à donner à connaître le lexique d'un état de la langue ancestral dans toute sa densité. Les deux méthodes se complètent admirablement, chacune comblant des lacunes que l'autre laisse inmanquablement, du fait de leurs objets propres, qui ne sont pas les mêmes, et qui correspondent à des pans différents d'un même système linguistique. Aussi les déromiens ne prétendent-ils pas remplacer la méthode utilisée jusqu'ici de façon radicale et définitive, et beaucoup d'entre eux font appel à l'ancienne méthode dans des travaux qu'ils mènent en dehors du projet. Ils sont conscients que le protoroman reconstruit demeure quelque chose d'incomplet, qui reflète uniquement les traits les plus persistants du latin parlé, et que d'autres traits ont pu être perdus sans qu'ils aient laissé la moindre trace, ce qui paralyse leur reconstruction. Et encore ! Puisque l'on reconstruit les éléments du protoroman au moment où il y a des conséquences directes sur les parlers romans, il se peut que les phonèmes reconstruits n'appartiennent pas tous exactement à la même époque. En conséquence, la notation phonologique des étymons n'est pas à interpréter comme la représentation infaillible de leur réalisation (ce qui est de toute façon exclu en raison du caractère abstrait des traits phonologiques).

Nous venons de voir ce qu'est le DÉRom, pourquoi il est né. Il a pour vocation de combler des vides que la méthode traditionnelle laisse inévitablement du fait de son approche, et qu'elle ne pourrait jamais combler elle-même. Nous savons maintenant que le DÉRom s'appuie sur la comparaison des différents idiomes romans pour reconstruire la protolange. Reconstruire – mais concrètement, qu'est-ce que cela veut dire ?

3 La méthode du DÉRom : comment il fait

Comment fonctionne la reconstruction ? Le *Dictionnaire des sciences du langage* de Franck Neveu nous dit :

« On appelle *reconstruction*, en linguistique historique, la reconstitution des langues-mères (ou protolangues), dont il n'existe aucune attestation directe. On aboutit à cette reconstruction en procédant par induction à partir de formes linguistiques attestées, empruntées à des langues différentes, dont une approche comparative permet d'établir les correspondances. Si dans un groupe de langues génétiquement liées, la comparaison permet de dégager des caractéristiques phonétiques ou morphologiques communes, on peut formuler l'hypothèse selon laquelle ces caractéristiques partagent la même origine linguistique. Le regroupement des caractéristiques communes d'un groupe de langues apparentées permet d'établir une représentation, hypothétique mais plausible, de la langue-mère. La reconstruction permet donc de confirmer l'existence des familles de langues, et, partant, de décrire les principales propriétés des protolangues » (Neveu 2011, 301).

3.1 Premières étapes

3.1.1 Parenté entre les langues et hérédité

Premièrement, on constate des ressemblances, sérielles et systématiques, en nombre étendu, et appartenant au lexique de base, entre les idiomes appelés romans. À partir de là, on fait l'hypothèse de leur parenté, ce qui revient à poser également le postulat de l'existence d'une phase antérieure dans laquelle ces langues n'en formaient qu'une, la protolangue. Ce postulat repose sur la tradition, qui l'a toujours admis, et sur l'observation déjà faite en comparant ces parlers, qu'au niveau de la forme des éléments du lexique de chacune, il existe des correspondances très régulières avec les éléments du lexique des autres, si bien qu'un locuteur d'une langue romane A pourrait tout à fait se débrouiller (ou du moins survivre) dans une autre langue romane B en connaissant seulement les correspondances de forme entre les unités des lexiques de A et B, parce que la forme des unités de B est prédictible en connaissant celle des unités de A. Nous appelons *cognats* les unités prises dans différents parlers romans qui se correspondent en raison de leur caractère d'héritières de la même lexie de la même langue ancestrale ; ce sont des « unités sœurs ». Par exemple, sarde *ánnu*, dacoroumain *an*, italien *anno*, français *an*, occitan *an*, espagnol *año* et portugais *ano*, qui signifient tous 'an', sont des cognats (cf. Celac 2008–2014 in DÉRom s.v. */'ann-u/).

Avec le postulat de la parenté, on suppose que si un certain nombre de parlers romans partagent un même trait, c'est que ce trait était déjà là dans la pro-

tolangue, et par conséquent, que sa présence dans les langues-filles est le fait non pas d'une évolution qui aurait eu lieu dans chacune d'elles de façon parallèle, mais de l'hérédité.

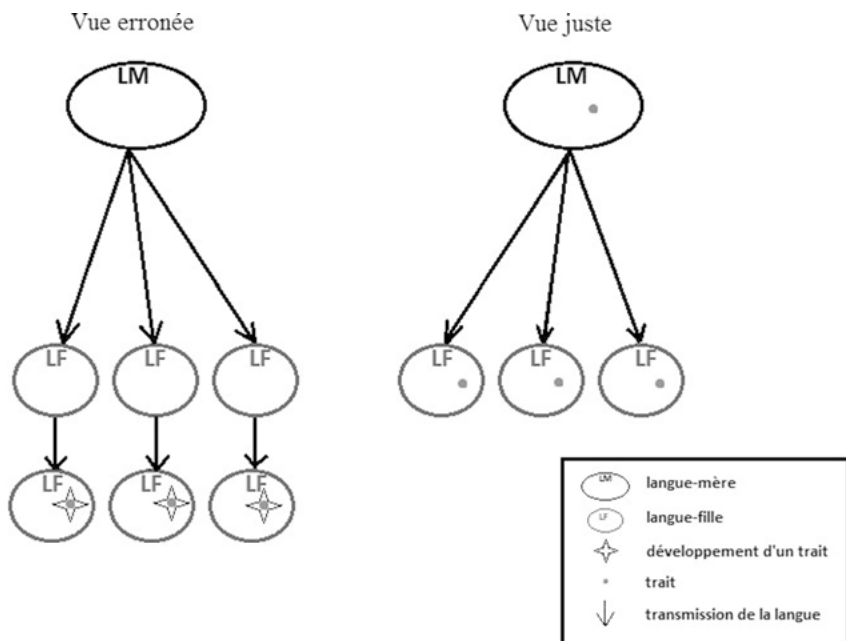


Figure 3 : Postulat de la présence d'un trait dans la protolangue

3.1.2 Rassemblement des matériaux pour la reconstruction

Puis on rassemble des séries d'unités lexicales que l'on suppose apparentées. On prend de préférence, pour établir ces séries, le vocabulaire de base (les noms des parties du corps humain, les termes de la parenté proche, les verbes dénotant les activités de base de l'être humain etc.). De cette manière, on est à peu près sûr que l'on n'aura pas d'emprunt (cf. Tadmor/Haspelmath/Taylor 2010)²⁵

²⁵ Malgré cette précaution, il existe des exceptions, c'est-à-dire des séries de cognats de vocabulaire de base incomplètes, parce qu'une des langues considérées ne présente pas de continuateur héréditaire, mais un emprunt. C'est le cas de la série de *répondre* et ses cognats, où l'istroroumain n'est pas représenté, étant donné que l'unité candidate *respundi* est en fait un

ni de création interne, qui doivent être systématiquement exclus, les lexies héréditaires étant les seuls matériaux légitimes pour la reconstruction.

3.1.3 Établissement des séries de correspondances phoniques

On dispose les cognats des différentes séries constituées, de façon à pouvoir les comparer. Il faudra *plusieurs* séries de cognats, et même *de nombreuses* séries de cognats pour pouvoir reconstruire les sons de la protolangue, et même avant cela, pour savoir précisément quels segments comparer en vue de la reconstruction. Nous tenterons de le faire comprendre grâce à un exemple (extrêmement simplifié).

Soit la série (non complète) de cognats suivante : français (abrégé *fr.*) *boire*, italien (en l'occurrence dialectal ; abrégé *it.*) *bevere*,²⁶ espagnol (abrégé *esp.*) *beber* et portugais (abrégé *port.*) *beber*.²⁷ Nous l'avons dit, c'est sur l'oral que nous devons travailler, aussi nous allons transcrire phonétiquement chacune de ces unités. L'objectif est de déterminer un certain nombre de correspondances phoniques entre ces langues, c'est-à-dire de comprendre quel segment du verbe français correspond à quel segment du verbe italien, espagnol et portugais.

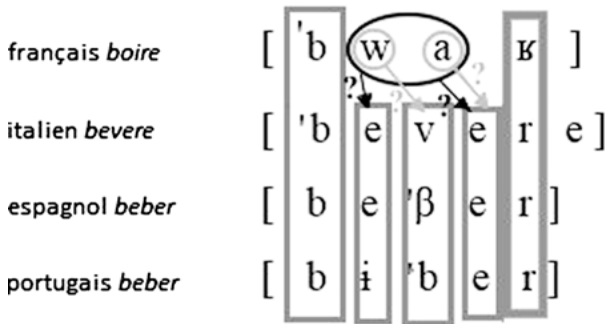


Figure 4 : Détermination des correspondances phoniques : fr. *boire* et ses cognats

emprunt à l'italien septentrional (au vénitien, plus précisément), cf. Videsott 2014–2015 in DÉRom s.v. **/res'pɔnd-e-/* n. 2.

²⁶ La forme de l'italien standard, *bere*, présente une syncope (ici, chute du segment /-ve-/) non régulière et ne peut pas être utilisée pour la reconstruction.

²⁷ La place de l'accent d'esp. [be'ber] et de port. [bi'ber] n'est pas héréditaire (cf. Victor Celac, « Normes rédactionnelles », § 8.1.8 « Verbes espagnols, asturiens, galiciens et portugais », ici 314–315).

On se rend vite compte que certaines associations de sons sont faciles à réaliser. Il en est ainsi du premier son de chaque cognat : fr. [b], it. [b], esp. [b] et port. [b] se correspondent. On peut encore associer fr. [ʁ], it. [r], esp. [r] et port. [r]. On voit aussi assez bien qu'it. [v], esp. [β] et port. [b] sont des sons plutôt proches, parce qu'ils sont articulés au niveau des lèvres. Mais d'autres liaisons sont nettement moins faciles à établir. Par exemple, à quoi correspond la séquence [wa] du mot français ? La réponse à cette question réside dans l'observation de plusieurs autres séries de cognats comparables.

Première hypothèse : puisque [w] est une semi-consonne (ou semi-voyelle) qui s'articule au niveau des lèvres, comme le [v], le [β] et le [b], il faudrait le mettre en relation avec ces trois consonnes, et [a] avec les voyelles qui les suivent.

Soit la série suivante : fr. *croire*, it. *credere*, esp. *creer*, port. *crer*. Alignons les comparats.

français <i>croire</i>	[	'k	ʁ	w	a	ʔ	ʁ	]	
italien <i>credere</i>	[	'k	r	e	d	e	r	e	]
espagnol <i>creer</i>	[	k	r	e	-	'e	r	]	
portugais <i>crer</i>	[	'k	r	e	-	-	r	]	

Figure 5 : Détermination des correspondances phoniques : fr. *croire* et ses cognats

Encore une fois, il est difficile de savoir à quoi la suite de sons fr. [wa] correspond. En tout cas, fr. [w] n'est pas à mettre en lien avec it. [v], esp. [β] et port. [b], puisqu'ici, il n'y a rien de semblable. La première hypothèse doit donc être exclue.

Deuxième hypothèse : le son [w] étant assez proche du son [u], voyelle, [w] serait à lier à la première des deux voyelles du verbe, et [a] à la deuxième.

Voyons avec la série de cognats fr. *trois*, it. *tre*, esp. *tres*, port. *três*.

français <i>trois</i>	[	't	ʁ	w a	?	]
italien <i>tre</i>	[	't	r	e		]
espagnol <i>tres</i>	[	't	r	e	s	]
portugais <i>três</i>	[	't	r	e	ʃ	]

Figure 6 : Détermination des correspondances phoniques : fr. *trois* et ses cognats

On voit par cet exemple, dans lequel les unités lexicales des quatre langues sont monosyllabiques, que cette deuxième hypothèse est également à rejeter. À première vue, cette série de cognats ne nous permet toujours pas de répondre à notre interrogation de départ. En réalité, fr. *trois* connaît une forme étendue dont nous aurions dû tenir compte dès le début, celle qui contient le son que l'on n'entend que dans la réalisation d'une liaison (comme dans *trois enfants*), à savoir [-z].²⁸

français <i>trois</i>	[	't	ʁ	w a	(z)	]
italien <i>tre</i>	[	't	r	e	—	]
espagnol <i>tres</i>	[	't	r	e	s	]
portugais <i>três</i>	[	't	r	e	ʃ	]

Figure 8 : Détermination des correspondances phoniques : fr. *trois* et ses cognats : solution

²⁸ On peut remarquer qu'en portugais, *três* n'a pas la même réalisation phonique selon qu'il est ou non suivi d'une unité lexicale commençant par une voyelle orale ; c'est la finale qui diffère. Ainsi *três* se réalise ['treʃ] dans *três casas* ('trois maisons'), mais ['trez] dans *três anos* ('trois ans') ; ['treʃ] est aussi la réalisation en mention : c'est la forme citationnelle orale.

Nous arrivons à la conclusion que fr. [ʷa] forme un tout correspondant à une seule voyelle des cognats italien, espagnol et portugais : sous l'accent, it./esp./port. [e] correspond à fr. [ʷa].

En principe, nous devrions poursuivre la recherche dans de multiples séries de cognats pour déterminer précisément avec laquelle des voyelles de chacun des cognats fr. [ʷa] est en relation. En outre, dans la série de *croire*, nous devrions justifier pourquoi nous rattachons port. [e] dans *crer* ('croire') à la première voyelle de l'italien et de l'espagnol plutôt qu'à la deuxième. Nous ne le ferons pas faute de temps, mais il est important de préciser que toutes ces étapes du raisonnement ont été parcourues par les chercheurs du DÉRom.

La détermination de l'ensemble complet des correspondances phoniques entre parlers est extrêmement complexe. En toute rigueur scientifique, même lorsqu'une correspondance paraît évidente (comme celle entre fr. [b], it. [b], esp. [b] et port. [b] initiaux dans la série de *boire* et ses cognats), il ne faut rien affirmer tant que tout n'a pas été consciencieusement vérifié par la confrontation à d'autres séries de cognats.

Les langues nous laissent parfois dans leur graphie – témoin de l'histoire –, qu'elle soit récente ou plus ancienne, de précieux indices. Le correspondant de *boire* dans un état ancien du français, *boivre*, confirme la relation que nous avons trouvée entre le tout que forme la suite de sons [ʷa] et une seule voyelle des trois autres langues. Nous aurions dû tout de suite nous fonder sur cette forme pour la confronter avec les cognats italien, espagnol et portugais, d'autant plus que, les connaissances avançant à mesure que l'on compare les unités et que l'on reconstruit le système du protoroman, on se rend compte que *boire* n'est pas le fruit d'une évolution tout à fait régulière (cf. Groß 2010–2014 in DÉRom s.v. */bɪβ-e-/ n. 5). D'une manière générale, le linguiste comparatiste ne garde, pour la reconstruction, que les formes régulières (ou, à défaut, il explique l'irrégularité de formes non régulières sur lesquelles il est néanmoins obligé de s'appuyer). Étant donné qu'il ne peut pas connaître dès le début de manière certaine si les formes des unités qu'il utilise comme matériaux de reconstruction le sont, il est contraint d'effectuer un va-et-vient continu entre ces matériaux et les règles qu'il induit à mesure qu'il compare.

3.1.4 Reconstruction des sons protoromans

Ayant en tête les indispensables connaissances sur la « mécanique » articulatoire par laquelle l'homme produit les unités sonores, il ne reste plus qu'à comparer les sons qui se correspondent d'une langue à l'autre pour reconstruire les

sons du protoroman et en constituer l'inventaire. Pour reprendre notre exemple, fr. ['wa] doit être comparé à it. ['e], esp. ['e], port. ['e] entre autres. La comparaison permet de dégager les traits phonétiques qu'ils partagent et de reconstruire les proto-sons (en l'occurrence *['e]). De bonnes connaissances sur le sens typique de l'évolution des sons dans les langues du monde au cours du temps sont précieuses pour cette phase de la reconstruction. Par exemple, dans le cas où des cognats d'une série présentent le son [k], d'autres le son correspondant [ʃ] (comme dans la série de fr. *chaîne* et ses cognats it. *catena*, esp. *cadena* et port. *cadeia*), pour déterminer lequel des deux est le plus proche du son ancestral, on doit se rappeler que les consonnes vélaires, dont l'articulation se fait en arrière de la bouche (comme précisément notre son [k]) évoluent très souvent pour laisser la place à des consonnes articulées plus en avant dans la bouche, au niveau du palais (comme notre son [ʃ]).²⁹ L'inverse n'est pour ainsi dire pas attesté (cf. Campbell 2013, 113–114 [« Directionality »]), de sorte que l'on peut identifier [k] comme plus ancestral et donc reconstruire protorom. *[k].

3.1.5 Reconstruction des phonèmes protoromans

L'inventaire des sons une fois fait, on reconstitue provisoirement les étymons des séries de cognats du vocabulaire de base. Dès lors, on peut étudier les contextes dans lesquels chaque son apparaît à l'intérieur des lexies protoromanes, pour déterminer lesquels sont une seule et même chose réalisée différemment selon les contextes, et lesquels sont bien distincts, c'est-à-dire que l'emploi de l'un ou de l'autre correspond à des unités lexicales distinctes. Les unités phoniques abstraites que l'on dégage de ces observations sont appelées *phonèmes*.³⁰ On est passé du plan phonétique (le plan de la réalisation concrète des sons) au plan phonologique (un plan plus abstrait, celui de leur statut discriminant dans le système d'une langue). L'exemple suivant permet de mieux comprendre de quoi il s'agit.

Dans la prononciation de <r> en français, que la pointe de la langue fasse contact avec la région alvéolaire (« r roulé », son représenté par [r]), qu'il y ait un « roulement » au fond de la gorge (« r grasseyé », son représenté par [ʀ]) ou que le dos de la langue touche le palais mou (« r parisien », son représenté par

²⁹ Cette évolution porte le nom de palatalisation.

³⁰ Pour rappel, nous notons les phonèmes, unités phoniques pertinentes d'une langue donnée, entre barres obliques (/.../) pour les différencier des sons tels qu'ils sont réalisés, que nous notons entre crochets ([...]).

[ʁ]) ne fait aucune différence essentielle. On considère que ce sont trois réalisations possibles d'un même phonème /r/. En portugais, au contraire, [ʀ] (ou [ʁ]) et [r] permettent de distinguer des unités lexicales comme *carro* ('voiture', prononcé ['karu] ou ['kaʁu]) et *caro* ('cher', prononcé ['karu]). On considère que l'on est en présence de deux phonèmes : /r/ réalisé [r], et /ʀ/ réalisé soit [ʀ] soit [ʁ] (cf. Barbosa 1994, 139–140).

À cette étape, on dresse l'inventaire des phonèmes, et on vérifie s'ils sont typologiquement plausibles. Ce sont ces phonèmes qui seront choisis pour la reconstruction proprement dite et définitive de la forme des unités lexicales protoromanes particulières.

Pour reconstruire chaque étymon individuel, il est hors de question de perdre du temps à répéter ces étapes à chaque fois. L'inventaire phonématique a été établi une fois pour toutes, sur la base des règles d'évolution des phonèmes protoromans en phonèmes romans que l'on a pu observer. Ces règles aident le linguiste à reconstruire rapidement les étymons. Cela suppose que l'on ait également reconstruit l'accent tonique en protoroman, très important, nous l'avons déjà dit, pour expliquer certaines évolutions, d'après la place qu'il occupe dans les unités lexicales des différentes langues romanes.

3.1.6 Identification des morphèmes

Le travail n'est pas encore terminé. Au cours des reconstructions particulières des unités lexicales, on s'interroge sur leur composition en morphèmes, qui sont les plus petites unités porteuses de sens, pour dégager les racines, les préfixes et les suffixes, et en faire également un inventaire. La décomposition en morphèmes nous dit quelque chose de plus sur les lexies, leur comportement (avec entre autres les suffixes de conjugaison pour les verbes), leur origine (s'ils sont construits au moyen d'ajout(s) de morphèmes de création lexicale, comme le suffixe *-et* en français,³¹ qui se greffe sur les noms communs désignant des réalités concrètes pour les diminuer dans les esprits : *garçonnet*, *jardinnet* etc.).

La reconstruction des morphèmes de conjugaison et de déclinaison se fait en observant plus particulièrement les morphèmes de conjugaison et de déclinaison.

³¹ Le trait d'union marque l'endroit où un morphème admet de « s'accrocher » aux autres dans la lexie. Par exemple, un préfixe sera noté avec un trait d'union à droite (comme le préfixe *in-* ('non') qui sert à former des unités telles que *inacceptable*, 'non acceptable'), tandis qu'un suffixe sera noté avec un trait d'union à gauche (comme le suffixe péjoratif *-asse*, qui sert à former des unités telles que *vinasse* s.f. 'mauvais vin').

naison à l'intérieur des mots-formes des unités lexicales romanes particulières. La conjugaison et la déclinaison se rapportent à la grammaire (donc à la partie régulière) d'une langue. Le chercheur mesure la plausibilité de la forme qu'il reconstruit au degré de récurrence de cette forme dans le lexique protoroman entier (pour des lexies comparables sur le plan de la nature syntaxique).

Reprenons l'exemple de */'kuer-e-/ (cf. Maggiore 2012–2015 in DÉRom s.v.). Comment le linguiste reconstruit-il deux sous-types présentant des morphèmes de conjugaison différents ?

Première étape. – Le chercheur relève des formes romanes qui l'incitent à reconstituer un morphème de conjugaison */'-ere/, ou, plus précisément, un morphème marquant la classe de flexion suivi d'un morphème marquant l'infinitif (*/'-e-re/) (ancien italien *cherere* v.tr. 'chercher', français dialectal *querre*, francoprovençal *querre*, occitan *querre*, catalan *querre*, entre autres cognats).

Deuxième étape. – Il observe que d'autres séries de cognats réunissant des verbes romans mènent à la reconstruction de la même suite de morphèmes hypothétique */'-e-re/ (cf. par exemple Blanco Escoda 2011–2016 in DÉRom s.v. */'batt-e-/ ou Groß 2014 in DÉRom s.v. */'βend-e-/). Cette observation lui vaut la validation de la forme */'-e-re/ jusque là candidate.

Troisième étape. –³² Par ailleurs, le linguiste recense des variantes formelles des cognats précédents, ayant en commun avec eux au moins le sens 'chercher', et formant (après enquête, cf. ci-dessus) des cognats tout aussi légitimes, le mettant sur la piste de l'existence d'une variante protoromane : italien septentrional *cherire*, français *querir*, francoprovençal *querir*, occitan *querir*, catalan *querir*, entre autres. La variation se fait au niveau du morphème flexionnel des verbes romans. De cela, il déduit que la variation porte sur le même point en protoroman, et il reconstruit une suite de morphèmes candidate de conjugaison */'-i-re/.

Quatrième étape. – Il observe d'autres séries de cognats conduisant à reconstruire le même morphème de conjugaison, ce qui confirme qu'il a pu exister une variante de son étymon en */'-i-re/ (cf. par exemple Groß/Schweickard 2010–2015 in DÉRom s.v. */'aud-i-/ ou Groß/Schweickard 2011–2014 in DÉRom s.v. */'dorm-i-/).³³ Enfin, il constate que ce phénomène de réalisation alternante d'un même étymon verbal, l'une en */'-e-re/ et l'autre en */'-i-re/, n'est pas isolé

³² Cette étape n'apparaît pas nécessairement après les deux premières, et même, plus vraisemblablement, elle se fait en même temps que la première, qui est une phase de récolte des matériaux.

³³ À l'infinitif, c'est bien le morphème flexionnel qui porte l'accent pour les deux verbes cités ici en exemple.

en protoroman (cf. Jatteau 2012–2014 in DÉRom s.v. */'ϕug-e-/ ou Gruner 2014 in DÉRom s.v. */'luk-e-/), d'où il obtient la certitude que cette variation est récurrente (cf. Benarroch/Baiwir 2014, 156–160) : la reconstruction particulière des deux sous-types */'kuɛr-e-re/ et */'kue'r-i-re/ est donc tout à fait plausible.

Parfois, il arrive que dans des unités lexicales reconstruites, on trouve ce qui pourrait ressembler à un morphème de création lexicale protoroman (c'est-à-dire que l'on serait tenté d'analyser la lexie comme un dérivé constitué d'un radical et d'un affixe dérivationnel), mais qu'après recherche, on se rend compte qu'il n'en est rien. Par exemple, après avoir reconstruit la partie phonique des unités */'ϕe'kundu/ adj. 'fécond', */'re'tundu/ adj. 'rond' et */'se'kundu/ ('second'),³⁴ on rattache chacun de ces adjectifs à sa classe flexionnelle, c'est-à-dire son modèle de déclinaison (premier découpage) : */'ϕe'kund-u/, */'re'tund-u/ et */'se'kund-u/. À ce stade, on pourrait penser que l'on peut les segmenter encore en isolant un morphème de création lexicale */-und-/ (second découpage). Si c'était le cas, il faudrait noter les unités comme suit : */'ϕe'k-und-u/, */'re't-und-u/ et */'se'k-und-u/. Mais, comme le montre Heidemeier 2014, pour la première unité, l'analyse en radical + suffixe n'est pas possible, car le radical présumé n'existe pas seul. Pour les deux autres, Ulrike Heidemeier montre qu'il y a bien eu dérivation, mais dans une langue ancestrale du protoroman, si bien que les unités sont arrivées telles quelles, figées, en protoroman. Dans ce cas, il faut conclure que */'und/ n'est pas un suffixe de création lexicale, du moins pas en protoroman (cf. Heidemeier 2014, 236–241 et passim). Il faut donc segmenter ces unités comme suit : /ϕe'kund-u/, /re'tund-u/ et /se'kund-u/.

Voilà pour ce qui est de la forme et de l'identification des morphèmes. Ce sur quoi la définition de la reconstruction du *Dictionnaire des sciences du langage* (cf. ci-dessus 3) n'insiste pas, c'est que le linguiste s'occupe de la reconstruction du sens et de la combinatoire restreinte, autant que possible, en plus de celle de la forme, et ce, toujours sur le principe de la comparaison.

3.2 Reconstruction du sens

En ce qui concerne le sens, les critères pour la reconstruction ne sont pas aussi théorisés que pour la reconstruction formelle (cf. Chauveau 2014a). Malgré cela, il est possible d'obtenir des résultats cohérents et certains.

³⁴ Les unités citées ici sont présentées avec la notation du DÉRom, excepté en ce qui concerne les frontières de morphèmes, que nous ajusterons d'après le développement de notre propos.

La reconstruction comparative donne nécessairement lieu à un étymon présentant au minimum un sens (ou, dans le cas des unités grammaticales, une fonction). Mais souvent elle mène à la reconstitution de plusieurs étymons liés entre eux et regroupés sous le même vocable : le DÉRom met en évidence la polysémie en protoroman.

Lorsque les parlers romans présentent, pour des unités lexicales correspondantes, exactement le(s) même(s) sens, on en déduit que ce(s) sens existai(en)t déjà en protoroman. Ainsi, fr. *an* et ses cognats présentant tous le sens ‘an’, nous reconstruisons sans aucun doute le sens ‘an’ pour leur ancêtre commun protoroman */'ann-u/.

Il arrive que les cognats romans ne signifient pas la même chose, mais que les sens d’un parler à l’autre soient tout de même liés par des traits communs. On sait que les langues peuvent développer des sens à partir d’autres en les spécialisant ou en les élargissant, c’est quelque chose de tout à fait naturel. C’est ce qui fait que fr. *TRAIRE*, qui avait initialement le sens général de ‘tirer’, se retrouve depuis le 12^e siècle (cf. TLFi s.v. *traire*) avec le sens actuel de ‘faire sortir un liquide (plus spécialement le lait de femelles de mammifères domestiques en leur pressant les trayons du pis), traire’. S’il arrive alors que les unités lexicales romanes présentent des sens différents mais liés par des traits généraux communs, on reconstruit le sens protoroman en généralisant les sens qui se sont spécialisés dans les langues romanes. Ainsi, pour */'pes-u/ (cf. Morcov 2014 in DÉRom s.v.), on reconstruit (entre autres) le sens décrit par la définition ‘ce qui pèse (sur qn ou qch.)’ par généralisation des sens décrits par ‘peine’, ‘objet lourd’, ‘charge morale’, ‘asthme’ (ces quatre sens en dacoroumain), ‘chose pénible qu’il faut supporter’, ‘paquet de chanvre d’un poids déterminé’ (ces deux sens en occitan), ‘grosse pierre pendue à la vis de la presse servant à la fabrication du cidre’ (asturien), ‘charge’ (en italien, en frioulan, en ladin, en romanche, en français, en francoprovençal) des lexies romanes.

Le linguiste reconstruteur doit également avoir la connaissance de tous les autres phénomènes réguliers de création de sens, comme la métonymie (ainsi la création, à partir du sens ‘matière’, du sens ‘objet fabriqué originellement dans cette matière’)³⁵ ou la création d’un sens métaphorique par rapport au sens de base,³⁶ pour se rendre compte de la plausibilité de ce qu’il reconstruit.

³⁵ Il en est ainsi de la création en français de *VERRE* s.m. ‘récipient servant à boire’ à partir de *VERRE* s.m. ‘matière siliceuse’, de *FER* ‘épée’ à partir de *FER* ‘métal’, de *CRAIE* ‘bâtonnet utilisé pour écrire’ à partir de *CRAIE* ‘roche calcaire’.

³⁶ Par exemple, on parle de *pied* pour désigner le bas d’une montagne ou d’un bâtiment par ressemblance avec la partie du corps humain en contact avec le sol.

3.3 Reconstruction de la combinatoire restreinte

Le DÉRom précise la partie du discours de l'unité lexicale protoromane reconstruite, son genre et éventuellement son nombre s'il est contraint (s'il s'agit d'un nom), enfin son régime, c'est-à-dire les types de compléments qu'elle requiert et la manière dont elle les gouverne, le cas échéant (s'il s'agit d'un verbe).

La reconstruction de la partie du discours de la lexie se fait assez facilement, compte tenu du fait que les unités romanes comparées relèvent en général de la même partie du discours.³⁷ Les chercheurs impliqués dans le DÉRom, qui considèrent que chaque unité lexicale doit avoir son étymologie, même si plusieurs possèdent la même forme, rassemblent des unités romanes comparables sur ce plan au moment où ils établissent les séries de cognats. Par exemple, pour fr. *froid*, qui peut être soit l'adjectif FROID (comme dans la phrase : *L'eau est trop froide pour que les baignades soient permises*), soit le substantif masculin FROID (comme dans la phrase : *Le froid arrive avec l'hiver*), ils auront soin d'établir deux séries de cognats différentes, pour mener deux enquêtes distinctes, ce qui donnera naissance à deux articles séparés dans le dictionnaire (s'il s'avère – hypothèse à notre avis hautement probable – que le substantif est bien héréditaire, et ne représente donc pas une création interne du français).

La reconstruction du genre pour les substantifs peut être plus compliquée, puisque celui des cognats n'est pas toujours homogène. Lorsque les cognats sont tous du même genre, on reconstruit sans trop d'hésitation ce qu'ils partagent de façon unanime. Lorsque tel n'est pas le cas, il faut étudier la répartition géographique des cognats présentant les différents genres, s'appuyer sur les études qui ont été faites sur le sujet de l'évolution du genre protoroman (ainsi de Dardel 1965), identifier éventuellement des traces du genre neutre dans les cognats, et trancher (Delorme/Dworkin 2014, 168–183 ; cf. aussi Benarroch/Baiwir 2014, 130–147).

La reconstruction du régime des verbes ne se fait pas dans la même visée que la reconstruction des autres paramètres de la lexie, où l'on cherche à retrouver le trait unique à la base de tous ceux que l'on connaît dans les langues romanes. À chaque fois que l'on trouve un type de régime (transitif direct, intransitif, pronominal etc.) chez une lexie romane descendante d'une lexie protoromane avérée, si l'on est certain que ce trait n'est pas une création interne

37 Là encore, la règle peut souffrir des exceptions. Par exemple, parmi les cognats sur lesquels on s'appuie pour la reconstruction de */kas-a/ s.f. 'maison', le cognat français est une préposition (*chez*), contrairement aux autres, qui sont des noms communs.

dans la langue romane en question, on considère que la lexie protoromane l'admettait. Par exemple, parmi les continuateurs du verbe */'klam-a-/ (cf. Mertens/Budzinski 2012–2015 in DÉRom s.v.), certains sont des verbes qui n'admettent aucun complément, avec le sens 'crier', certains demandent un complément d'objet direct, avec le sens 'appeler (qn)', certains demandent un complément d'objet direct et un attribut de l'objet, avec le sens 'proclamer (qn) (qch.)', certains demandent un complément d'objet direct, avec le sens 'nommer (qn)', certains sont pronominaux et demandent un attribut du sujet, avec le sens 's'appeler (nom propre)'. Pour la reconstruction du morphème de conjugaison de l'étymon de fr. *quérir* et ses cognats (cf. ci-dessus 2.3.2 et 3.1.6), Marco Maggiore avait essayé de (et réussi à) ramener les deux étymons intermédiaires, */'kuɛr-e-/ et */'kue'r-i-/, à un seul étymon originel, */'kuɛr-e-/. Ici, on procède différemment. On ne cherche pas à réduire les sous-types à un seul, on considère qu'ils étaient tous présents en protoroman. On attribue donc à l'étymon protoroman les cinq régimes.

3.4 Récapitulatif des fondements de la reconstruction comparative

Pour récapituler en quelques phrases le principe de la reconstruction comparative, nous pouvons dire que l'observation de ressemblances entre des unités lexico-sémantiques appartenant à des parlers différents mène à l'établissement de séries de correspondances entre des (ensembles de) traits de ces unités, ce qui permet de démontrer leur correspondance, et ainsi, au dernier niveau, la parenté des langues auxquelles ces unités appartiennent. Cette correspondance ne peut pas être le fruit du hasard, elle est le fait de l'origine commune de tous ces parlers ; l'ancêtre commun qui les rassemble est appelé *protolangue*, et, dans le cas des langues romanes, *protoroman*. De là, la comparaison des sons/phonèmes, sens et caractéristiques de combinatoire en correspondance dans les unités en correspondance de ces langues en correspondance permet au linguiste de dégager des traits communs, traits dont on suppose l'existence en protoroman, de préférence à l'hypothèse selon laquelle les différents parlers romans auraient développé le même trait chacun de son côté. Ce sont ces traits existant en protoroman que le linguiste reconstruit.

Bien qu'elle ait ses limites, la méthode de la reconstruction comparative est particulièrement intéressante et permet de faire de vraies découvertes sur le lexique du protoroman et, d'une manière générale, sur le latin global. La forme de dictionnaire est particulièrement adaptée pour rendre visibles les fruits de la

recherche scientifique et pour les transmettre. Mais puisqu'il s'agit d'un ouvrage scientifique, le DÉRom est difficilement abordable pour les personnes qui n'ont pas reçu une formation préalable. Après avoir eu un aperçu des richesses qu'il contient (ci-dessus 3), le lecteur a besoin d'apprendre le langage technique dans lequel le DÉRom est rédigé. Nous nous proposons de l'y aider dans cette quatrième partie.

4 « Dissection » d'un article du DÉRom

La qualité de contenu d'un dictionnaire étymologique, et même d'un dictionnaire en général, ne réside pas tant dans sa couverture lexicographique (l'étendue du lexique représenté, par rapport au lexique total concerné par le sujet du dictionnaire), aspect plutôt quantitatif (mais qui a son importance !), que dans la complétude et l'exactitude des données qu'il présente par rapport au but qu'il se donne. D'où la masse compacte d'informations que l'on trouve dans le DÉRom, car en étymologie, rien n'est évident : ni les résultats ni le sérieux des comportements de recherche mis en œuvre pour les trouver. Il faut tout expliciter pour que tout soit complet, tout justifier pour que le lecteur soit en mesure de vérifier lui-même les différentes étapes de l'argumentation.

La première chose qu'il nous est donné de voir dans les articles du DÉRom, c'est leur aspect général. Les informations sont toujours réparties de la même manière dans la microstructure du dictionnaire, l'article se découpe systématiquement en six parties (ou sept s'il y a des notes). La régularité de la présentation nous aide à nous y retrouver.

4.1 Le résultat de la reconstruction

Ce qui apparaît de la façon la plus évidente, c'est le fruit de la recherche, l'étymon, avec sa forme, son sens et sa combinatoire restreinte (ses propriétés syntaxiques).

4.1.1 Forme

La forme d'un étymon apparaît dans le lemme d'une manière propre au DÉRom.

4.1.1.1 Astérisque

Nous identifions les unités protoromanes grâce à l'astérisque (*) dont elles sont précédées. Ce signe est là pour indiquer toute unité obtenue par le procédé de la reconstruction comparative à partir d'unités romanes assurément héréditaires, qu'il s'agisse d'une unité lexicale tout entière (comme */lumen/ s.n. 'lampe ; lumière', Georgescu 2014/2015 in DÉRom s.v.), d'un morphème (ainsi le morphème dérivationnel */-ani-/ , Celac 2012–2014 in DÉRom s.v. */mon't-ani-a/, commentaire) ou d'un phonème (par exemple */e/, Reinhardt 2008–2014 in DÉRom s.v. */'φen-u/ ~ */'φεν-u/ n. 8).³⁸

4.1.1.2 Notation phonologique

La forme de l'étymon apparaît en écriture phonologique, les barres obliques en témoignent. Pour le déchiffrement de l'étymon, nous avons besoin surtout de connaître les signes que nous exposons ici, parce que ce sont ceux des seuls phonèmes reconstruits dans le système phonologique de la protolangue romane. Ils sont notés en alphabet phonétique international (API). Nous associons à chaque symbole un ou plusieurs exemples de lexies romanes ou non romanes dans lesquelles on trouve le son correspondant ; l'accent est noté par un symbole qui a la forme d'une apostrophe droite (') placé immédiatement avant la syllabe accentuée. Dans les illustrations ci-dessous, nous soulignons les phonèmes en question, ainsi que les caractères (lettres) qui les notent.

4.1.1.2.1 Voyelles

*/i/ : cf. fr. *visite* /vi'zɪt/, esp. *hija* /'ixa/ ('fille'), port. *positivo* /puzi'tivu/ ('positif'). */i/ protoroman peut être réalisé comme une voyelle, *[i], ou – plus rarement – comme la semi-consonne (disons, pour simplifier, un élément qui a un mode d'articulation entre une voyelle, sans obstruction, et une consonne, avec obstruction) *[j], comme dans */βi'n-aki-a/, réalisé *[βi'n-akj-a]. C'est la différence entre fr. *abbaye* [abe'i] et *abeille* [a'bej].

*/ɪ/ : cf. angl. *it* /ɪt/ ('cela'). Globalement, on ouvre un peu plus la bouche et on recule très légèrement la langue par rapport au précédent ; c'est un peu l'intermédiaire entre /i/ et /e/.

*/e/ : cf. fr. *établir* /ɛta'bliʁ/, esp. *eco* /'eko/ ('écho'), port. *português* /portu'gɐʃ/ ('portugais').

³⁸ Cela ne vaut donc pas uniquement pour le lemme.

***/ɛ/** : cf. fr. *misère* /mi'zɛʁ/, port. *fonético* /fu'nɛtiku/ ('phonétique').

***/a/** : cf. fr. *vache* /'vaʃ/, esp. *boca* /'boka/ ('bouche'), port. *obrigado* /ubɾi'gaðu/ ('obligé' ou 'merci').

***/ɔ/** : cf. fr. *bosse* /'bɔs/, port. *normal* /nɔɾ'maɫ/ ('normal').

***/o/** : cf. fr. *beau* /'bo/, esp. *vaso* /'baso/ ('verre'), port. *tricô* /tri'ko/ ('tricot').

***/ʊ/** : cf. angl. *put* /'pʊt/ ('poser'). Globalement, on ouvre un peu plus la bouche et on avance très légèrement la langue par rapport au suivant ; c'est un peu l'intermédiaire entre /o/ et /u/.

***/u/** : cf. fr. *louche* /'luʃ/, esp. *azul* /a'θul/ 'bleu', port. *tudo* /'tuðu/ ('tout').

***/u/** protoroman peut se réaliser comme une voyelle, [u], ou – plus rarement – comme la semi-consonne [w] (comme */'akuil-a/, qui se réalise *['akwil-a]), que nous trouvons dans fr. *boire* /'bwaʁ/.

Aucune voyelle nasale n'est reconstruite en protoroman.

4.1.1.2.2 Consonnes

***/p/** : cf. fr. *papa* /pa'pa/, esp. *sopa* /'sopa/ ('soupe'), port. *limpo* /'lipu/ ('propre').

***/b/** : cf. fr. *bateau* /ba'to/, esp. *sombra* /'sombra/ ('ombre'), port. *baixo* /'bajʃu/ ('bas').

***/t/** : cf. fr. *tableau* /ta'blo/, esp. *tocar* /to'kar/ ('toucher'), port. *postal* /puʃ'taɫ/ ('carte postale').

***/d/** : cf. fr. *dame* /'dam/, esp. *mundo* /'mundo/ ('monde'), port. *dom* /'dõ/ ('don').

***/k/** : cf. fr. *écrire* /e'kʁiʁ/, esp. *casa* /'kasa/ ('maison'), port. *choçar* /ʃu'kaʁ/ ('couver').

***/g/** : cf. fr. *graine* /'gʁɛn/, esp. *gato* /'gato/ ('chat'), port. *agora* /e'gɔɾɐ/ ('maintenant').

***/ʔ/** : très peu de langues le possèdent. Le son, bilabial, est intermédiaire entre /f/ et /h/ ; à l'oreille, c'est plus doux que /f/ (cf. le souffle par lequel on éteint une bougie). Les Japonais le prononcent dans des emprunts comme *Furansu* /ʔuransu/ ('France').

***/β/** : cf. esp. *lavar* /la'bar/ ('laver'). C'est un peu un intermédiaire entre /b/ et /v/, cela se rapproche d'un /w/ très bref.

***/s/** : cf. fr. *cité* /si'te/, esp. *mesa* /'meʃa/ ('table'), port. *abraço* /e'braʒu/ ('étreinte').

***/m/** : cf. fr. *permis* /pɛʁ'mi/, esp. *mano* /'mano/ ('main'), port. *amarelo* /ɐ'mɐ'relu/ ('jaune').

***/n/** : cf. fr. *nager* /nɑ'ʒe/, esp. *donde* /'donde/ ('où'), port. *banana* /be'nene/ ('banane').

***/r/** : cf. esp. *caro* /'kaɾo/ ('cher'), port. *abrasivo* /ɐbre'zivu/ ('caustique').

***/l/** : cf. fr. *blle* /'bal/, espagnol *falda* /'falda/ ('jupe'), port. *louro* /'loru/ ('blond').

4.1.1.3 Frontières de morphèmes

Les frontières entre morphèmes sont visualisées au moyen de traits d'union (-) dans les unités protoromanes reconstruites. La séparation des morphèmes est là pour marquer la façon dont l'unité est structurée. Dans la plupart des unités traitées, le dernier élément est un morphème indiquant la classe flexionnelle (morphème de conjugaison ou de déclinaison). Ce n'est pas (ou pas uniquement) une désinence (terminaison) en soi, c'est quelque chose de plus abstrait qui renvoie à un patron flexionnel. De ce fait, le lemme ne représente pas un mot-forme particulier, mais l'unité lexicale tout entière. Nous avons bien ici une unité protoromane entière comme étymon d'unités romanes entières (cf. ci-dessus 2.2.2.3).

Les lemmes */agr-u/ s.n. 'champ ; territoire rural' (cf. Alletsgruber 2014/2015 in DÉRom s.v.) et */dis-ka'ball-ik-a-/ v.intr./tr. 'descendre de selle ; faire descendre de selle ; faire cesser d'être en position de chevauchement' (cf. Hütsch/Buchi 2014 in DÉRom s.v.) nous serviront d'exemples pour visualiser cette particularité de la notation des étymons dans le DÉRom (cf. ci-dessous figure 9).

Protorom. */agr-u/ se subdivise en deux morphèmes, un morphème lexical (le radical) et un morphème indiquant la classe flexionnelle, le morphème flexionnel */-u/. Le cas de */dis-ka'ball-ik-a-/ est un peu plus complexe. Nous avons un morphème lexical central, /-ka'ball-/ ('cheval'), auquel viennent s'adjoindre */-ik-/ , un suffixe dérivationnel formateur de verbes à partir de noms³⁹ et */dis-/ , dont les auteurs de l'article disent qu'il est un « préfixe formateur de verbes qui véhicule notamment les sens de changement d'état et de négation » (Hütsch/Buchi 2014 in DÉRom s.v. */dis-ka'ball-ik-a-/ , commentaire), qui correspond à notre préfixe *dé-* en français, comme dans *décommander*, *défaire*, *démonter* ou encore *dénouer*. Et bien sûr à cela s'ajoute le morphème indiquant la classe flexionnelle */-a-/.

³⁹ Cf. Jactel/Buchi 2014/2015 in DÉRom s.v. */ka'ball-ik-a-/ , verbe protoroman signifiant 'être en selle ; chevaucher ; être à califourchon (sur) ; s'accoupler (avec une femelle)' et constitué du morphème lexical */-ka'ball-/ ('cheval') et de ce suffixe dérivationnel */-ik-/.

Les morphèmes flexionnels ne se présentent pas tout à fait de la même façon pour les substantifs (ainsi que les pronoms et les adjectifs) ou les verbes. Ceux des substantifs, des pronoms et des adjectifs sont notés par un signe *à la suite d'un tiret*, ceux des verbes sont notés par un signe *entre deux tirets*. C'est que les verbes protoromans conjugués, comme les verbes romans, peuvent admettre plusieurs morphèmes grammaticaux à la suite les uns des autres, pour marquer tel temps, tel mode etc. En français, le verbe MARCHER, conjugué à la première personne du singulier au présent de l'indicatif, se décompose en *march-e*, avec *march-* comme morphème lexical (la racine) et *-e* comme morphème flexionnel. Conjugué, cette fois-ci à la première personne du singulier au futur simple, il se décompose en *march-e-r-ai*, toujours avec *march-* comme morphème lexical, avec cette fois-ci *-e-*, la marque de la classe de conjugaison, *-r-*, la marque du futur, et *-ai*, la désinence de la première personne du singulier (pour le futur simple).

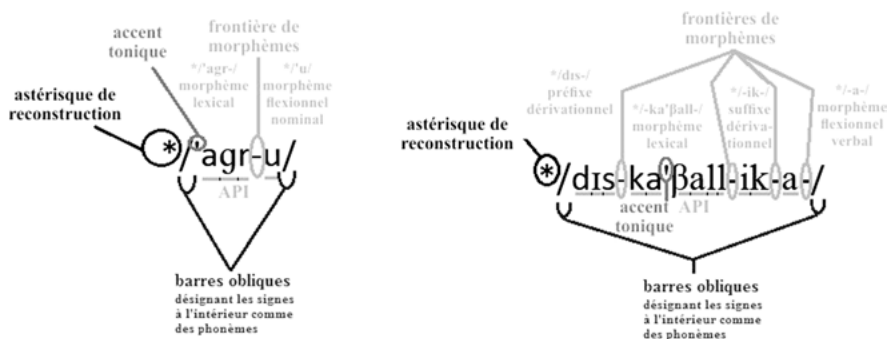


Figure 9 : Analyse des lemmes */'agr-u/ et */dis-ka'βall-ik-a-/

4.1.2 Sens

Le sens des unités lexicales du vocable protoroman est donné sous la forme non d'un simple mot équivalent en français, mais sous celle d'une véritable définition componentielle. Cela signifie que le sens est décomposé en traits sémantiques qui peuvent aider à différencier telle lexie d'une autre au sens proche. L'avantage de cette manière de procéder est qu'elle évite les ambiguïtés liées au caractère polysémique de la plupart des vocables.

Prenons fr. VIANDE, dont le sens peut être paraphrasé par 'aliment tiré des organes/muscles des mammifères et des oiseaux'. Si nous devons en trouver un

équivalent en portugais, nous avons à notre disposition CARNE. Seulement, alors que VIANDE n'admet aujourd'hui qu'un seul sens, CARNE en présente au moins deux : celui-ci même, et le sens 'composante du corps humain ou animal, constituée essentiellement des tissus musculaires'. Les sens sont proches, la différence entre les deux réside principalement dans le fait que le premier possède un trait sémantique essentiel supplémentaire : le fait que la chose soit considérée du point de vue de son caractère comestible. Le meilleur moyen d'éviter l'ambiguïté est de donner une définition qui se décompose en traits sémantiques essentiels.

Cela dit, pour éviter de répéter la définition à chaque fois que l'on a besoin de parler de la lexie concernée, on crée un « raccourci » en donnant en plus de la définition componentielle une glose, un mot qui résume l'idée en français. Elle apparaît au moins au début du commentaire. Nous verrons un peu plus loin que la glose ne sert pas uniquement à identifier une lexie particulière.

4.1.3 Combinatoire restreinte

L'information concernant la partie du discours de l'étymon protoroman suit immédiatement la forme de l'étymon. Lorsqu'il s'agit d'un substantif, le genre est précisé : masculin, féminin ou neutre.

Pour les verbes, les informations sur le régime se trouvent réparties dans deux endroits : le lemme étymologique et l'énoncé de l'étymologie au début du commentaire (cf. ci-dessous figure 10). La structure argumentale (l'ensemble des types de compléments exigés par le verbe) est résumée juste après la mention de la partie du discours (« v. »). Un verbe peut en effet admettre plusieurs constructions syntaxiques :

- impersonnelle (« v.impers. ») : une construction n'admettant qu'un sujet grammatical fixe, sans véritable référent, comme dans */'plɔβ-e-/ ('pleuvoir', cf. Hinzelin en préparation in DÉRom s.v.), ou encore les verbes français FALLOIR (*il faut*), Y AVOIR (*il y a*) ou NEIGER (*il neige*) ;
- intransitive (« v.intr. ») : une construction n'admettant aucun complément mis à part le sujet, comme dans */'dɔrm-i-/ ('dormir', cf. Groß/Schweickard 2011–2014 in DÉRom s.v.) ;
- transitive (« v.tr. ») : une construction admettant, en plus du sujet, un complément d'objet, le plus souvent un complément d'objet direct, comme dans */'batt-e-/ ('battre', cf. Blanco Escoda 2011–2014 in DÉRom s.v.) ;

- ditransitive (« v.ditr. ») : une construction admettant, en plus du sujet, deux compléments d'objet, l'un direct, l'autre indirect, comme dans */im'prumut-a-/ ('prêter', cf. Maggiore 2014/2015 in DÉRom s.v.) ;
- pronominale (« v.pron. ») : une construction au moyen du pronom réfléchi, comme dans une des lexies de */'klam-a-/ ('s'appeler', cf. Mertens/Budzinski 2012–2015 in DÉRom s.v.).

Nous venons d'utiliser les gloses des verbes protoromans non seulement pour indiquer leur sens, mais aussi pour donner une meilleure idée des constructions syntaxiques qu'ils requièrent, par le parallèle avec la construction des verbes de la métalangue du DÉRom, le français, qui leur correspondent. C'est exactement là le second rôle des gloses dans le DÉRom : elles sont choisies pour avoir le même régime que les lexies protoromanes qu'elles sont censées définir (cf. Delorme/Dworkin 2014, 190–193). Elles ont donc une fonction importante au sein de l'économie lexicographique du DÉRom, que les définitions componentielles ne sont pas toujours à même de remplir.

Alors que la glose nous indique de façon implicite quels types de compléments syntaxiques le verbe prend, les éléments entre parenthèses dans la définition donnent à voir de quels types sémantiques ces compléments doivent être. Les parenthèses ne signifient pas qu'ils sont facultatifs. Elles sont là pour montrer qu'ils sont libres, puisqu'ils n'appartiennent pas à proprement parler à la définition, donc au verbe que la formulation de celle-ci doit pouvoir remplacer. Autrement dit, grâce aux parenthèses, au moment où l'on insère le verbe dans un énoncé, on peut substituer un complément à la formule exprimant la contrainte sémantique du complément, selon le sens que l'on veut exprimer, à condition que cette contrainte soit respectée.

Le vocable */'klam-a-/ (cf. Mertens/Budzinski 2012–2015 in DÉRom s.v.) nous aidera à illustrer ce que nous venons d'énoncer sur la structure des informations liées à la combinatoire restreinte (et accessoirement au sens) dans un article du DÉRom.

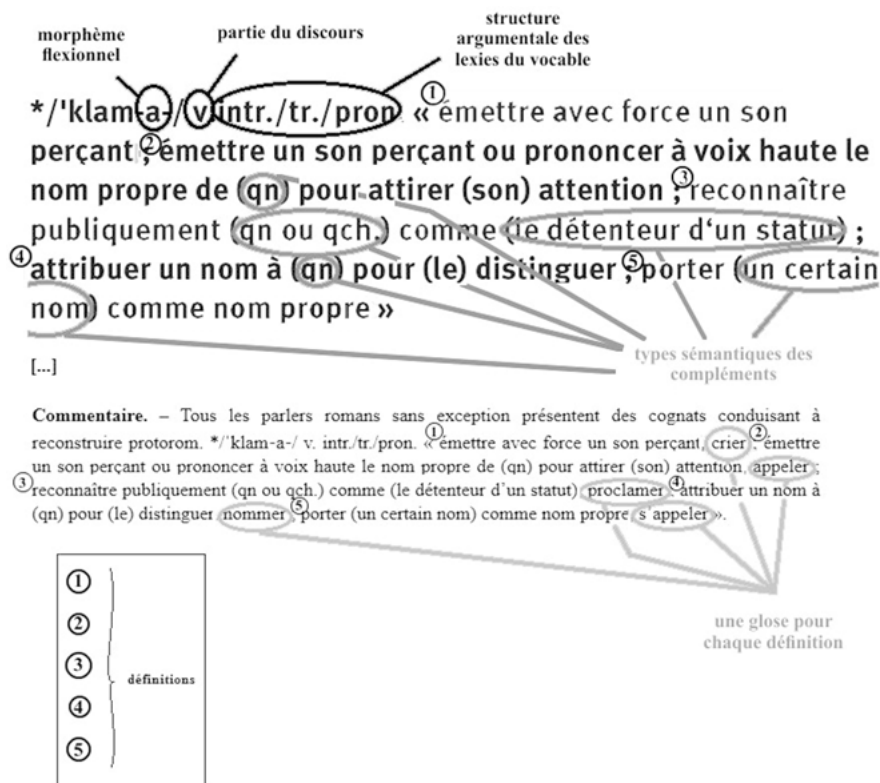


Figure 10 : Combinatoire restreinte de **/'klam-a-/* (1/2)

Les informations réparties dans les deux zones concernées (glose et indication entre parenthèses dans la définition componentielle) de l'énoncé de l'étymologie au début du commentaire se complètent, ainsi que l'illustre le tableau suivant :

	Partie du discours	Structure argumentale	Glose	Indications entre parenthèses dans la définition
1 (I.)	Verbe	Intransitif	‘crier’ (aucun complément)	–
2 (II.)		Transitif	‘appeler’ (+ COD)	‘[...] prononcer [...] le nom propre de (qn) [...]’ (COD = personne)
3 (III.)		Prédicatif transitif	‘proclamer’ (+ COD + attribut du COD)	‘reconnaître [...] (qn ou qch.) comme (le détenteur d’un statut)’ (COD = libre ; attribut du COD = statut)
4 (IV.1.)		Transitif	‘nommer’ (+ COD)	‘attribuer un nom à (qn) [...]’ (COD = personne)
5 (IV.2.)		Prédicatif pronominal	‘s’appeler’ (+ attribut du sujet)	‘porter (un certain nom) comme nom propre’ (attribut du sujet = nom propre)

Figure 11 : Combinatoire restreinte de */'klam-a-/ (2/2)

Dans la description du deuxième sens, correspondant à la subdivision II. de l'article du DÉRom, ‘[...] prononcer [...] le nom propre de (qn) [...]’, la définition pourrait nous induire en erreur par rapport au type syntaxique du complément, l'élément entre parenthèses remplissant dans cette formulation la fonction de complément du nom. C'est surtout à la glose qu'il faut se fier pour détecter qu'il s'agit en réalité d'un complément d'objet direct (COD), même si la définition telle qu'elle est présentée laisse apparaître un précieux (mais très discret) indice : elle place la préposition *de* hors des parenthèses, pour signifier que le complément n'est pas introduit par une préposition.

4.2 Les matériaux de la reconstruction et les sources

En tant qu'ouvrage scientifique, le DÉRom rend compte des matériaux (les cognats) qui lui ont permis d'obtenir les résultats affichés dans les lemmes des articles, de manière à laisser la possibilité aux experts de refaire le cheminement, de corriger, de critiquer.

4.2.1 Choix des idiomes considérés pour la reconstruction

Il serait impossible de rassembler les matériaux de tous les parlers romans : jusqu'à une époque relativement récente, chaque village avait pour ainsi dire sa variété.⁴⁰ Il faut simplifier, mais le découpage de cette réalité linguistique plurielle n'est pas facile à faire pour qu'il soit satisfaisant au niveau de la collecte des données : par exemple, est-il suffisant de chercher des matériaux en italien standard (d'origine florentine) seulement, ou bien faut-il enquêter dans les dialectes régionaux ou locaux ?

Le DÉRom distingue entre idiomes (ou parlers ou systèmes linguistiques ou langues) obligatoires, c'est-à-dire des idiomes dans lesquels on cherchera en priorité les matériaux, et idiomes facultatifs (considérés comme des subdivisions des parlers obligatoires), dans lesquels on puisera uniquement dans le cas où les parlers obligatoires ne présentent pas de donnée.⁴¹ Les parlers qui ont été retenus comme idiomes obligatoires sont les suivants (pour les idiomes facultatifs, cf. Victor Celac, « Normes rédactionnelles », ici 257–317 § 3.4) :

- le sarde (« sard. ») ;
- le dacoroumain (« dacoroum. ») ;
- l'istroroumain (« istroroum. ») ;
- le méglénoroumain (« méglénoroum. ») ;
- l'aroumain (« aroum. ») ;
- le dalmate (« dalm. ») ou plutôt le végliote (« végl. ») ;⁴²
- l'istriote (« istriot. ») ;
- l'italien (« it. ») ;
- le frioulan (« frioul. ») ;
- le ladin (« lad. ») ;
- le romanche (« romanch. ») ;
- le français (« fr. ») ;
- le francoprovençal (« frpr. ») ;
- l'occitan (« occit. ») ;
- le gascon (« gasc. ») ;

40 Pour le cas de la France, cf. le rapport de l'Abbé Grégoire, une des figures emblématiques de la Révolution (Grégoire 1794), fait dans le but de supprimer les parlers dialectaux au profit de la langue française, même si c'est sous la III^e République qu'un programme d'éradication des dialectes sera vraiment appliqué.

41 Le découpage entre idiomes obligatoires et facultatifs est un point qui est toujours en discussion au sein du projet (cf. Buchi/Schweickard 2014b, 30).

42 Suite aux réflexions de Vuletić 2013 et de Chambon 2014, cette question est sur le point d'être résolue au sein de l'équipe (cf. Buchi/Schweickard 2014, 30).

- le catalan (« cat. ») ;
- l'espagnol (« esp. ») ;
- l'asturien (« ast. ») ;
- le galicien (« gal. ») ;
- le portugais (« port. »).

Les matériaux sont toujours exposés dans le même ordre. Nous avons donné ci-dessus la liste des idiomes obligatoires en nous y conformant : d'abord le sarde, puis les parlers d'est en ouest. La carte de la page suivante situe les différents idiomes concernés dans l'espace.

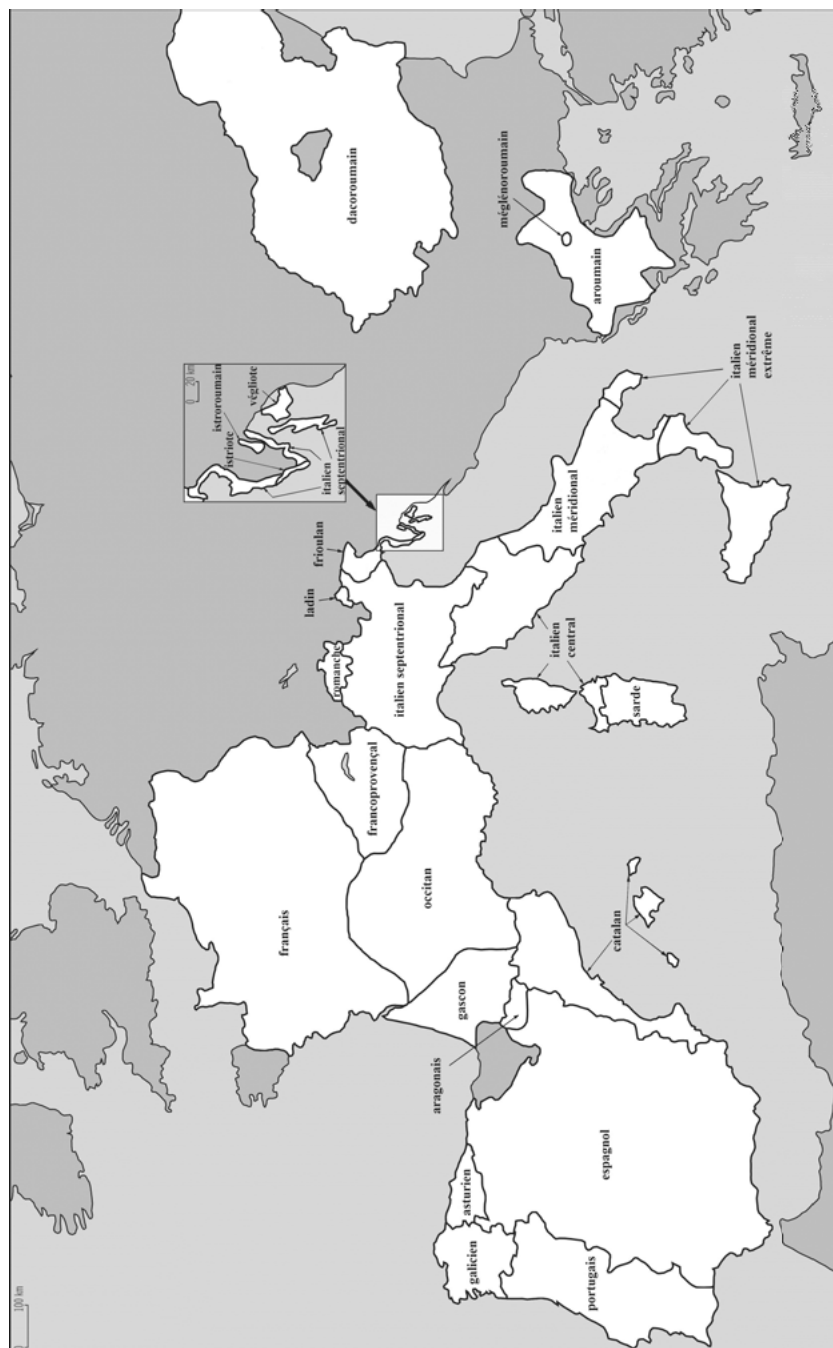


Figure 12 : Principaux idiomes convoqués dans le DÉRom

4.2.2 Cognats

Les cognats sont notés à la suite du nom du parler dans lequel on les trouve, à commencer par leur signifiant (forme). Bien que les chercheurs se basent sur la forme sonore des unités lexicales romanes, le DÉRom a fait le choix de les citer sous leur forme écrite. D'abord il estime que leur phonie peut être trouvée relativement facilement dans la majeure partie des cas, parce que les langues concernées sont très majoritairement encore en usage, et que si elles ne sont pas toujours bien décrites, elles le sont de mieux en mieux (cf. Buchi/Schweickard 2011a, 307–308). De plus, afin de faciliter la vérification et la critique, l'écriture conventionnelle est la plus appropriée, puisque c'est par la graphie que l'on accède aux entrées des dictionnaires et autres sources de référence. Si les cognats étaient notés sous leur forme phonologique, comment les retrouver dans les ouvrages lexicographiques, surtout pour les langues à l'orthographe complexe (comme le français) ?⁴³

Lorsque les parlers offrent plusieurs variantes formelles d'une lexie, on ne cite qu'une seule forme (la plus typique) dans la graphie conventionnelle contemporaine pour les idiomes standardisés, ou, dans le cas de parlers non standardisés, on choisit la plus représentative, que l'on note par des taquets de typisation : '...'. Par exemple, dans l'article */*ɸe'βrari-u/* s.m. 'février' (Celac 2009–2014 in DÉRom s.v.), on a pour le sarde '*frebàriu*'.

4.2.3 Données autour des cognats : sources et premières attestations

La date de la première attestation est précieuse en étymologie en général et pour la reconstruction en particulier : une attestation ancienne, remontant aux premiers témoignages écrits connus de l'idiome en question, est un indice pour déterminer qu'une lexie est héréditaire. La date de la première attestation est mentionnée pour les cognats des langues qui ont une tradition lexicographique importante. Elle figure comme premier élément dans la parenthèse, avant le renvoi aux sources. Par exemple, dans l'article */*'kaput/* s.n. 'tête ; extrémité' (Schmidt/Schweickard 2015/2016 in DÉRom s.v.), on a dans le type II.1., glosé par 'tête', pour le gascon : « gasc. *cap* (dp. 1495 ['maison principale d'un ordre

⁴³ Certains dictionnaires informatisés permettent un tel mode de recherche. C'est le cas du TLFi (cf. ci-dessus n. 15), mais ce n'est pas un fait répandu pour toutes les langues, loin de là.

religieux’], MillardetRecueil 91 ; FEW 2, 334a, CorominesAran 384).⁴⁴ Cette première attestation, relevée par Jean-Paul Chauveau, le réviseur des domaines français, francoprovençal, occitan et gascon, est très précieuse.

Lorsque la forme graphique du cognat utilisé pour la reconstruction ne correspond pas à celle de la première attestation, on indique cette dernière entre crochets carrés, en précisant le cas échéant de quel mot-forme de la lexie il s’agit, son sens s’il est différent du sens du cognat cité, et enfin les sources. Par exemple, dans l’article */‘laud-a-/ v.tr. ‘louer’ (Videsott 2015 in DÉRom s.v.), on a pour le gascon : « gasc. *laudà* (dp. 1261 [*lausar* ‘annoncer (une marchandise) avec éloge’], FEW 5, 206b ; Palay) ».

Le DÉRom répond aux exigences de traçabilité scientifique en donnant systématiquement les références des ouvrages dans lesquels les matériaux ont été trouvés : la base de la reconstruction doit être solide, bien établie ; les cognats choisis sont assurément attestés. Pour ne pas surcharger les articles, les références sont abrégées. Sur le site du DÉRom, où ils sont publiés au fur et à mesure de leur achèvement, on trouve les références complètes à partir du sigle de l’ouvrage en cliquant sur ce dernier.

4.2.4 Mot-forme reconstruit

La reconstruction d’un étymon complet, qui a le statut de lexie, passe par la reconstruction préalable de ses mots-formes (formes fléchies) :⁴⁵ le chercheur, avant de retrouver les unités abstraites, est contraint de travailler sur des entités concrètes. Cette étape intermédiaire est marquée dans les matériaux par la notation, au début de chaque série de cognats, d’un mot-forme protoroman, celui qui est le fruit de la reconstruction comparative sur la base des mots-formes romans utilisés comme formes citationnelles des cognats (par exemple, pour les verbes, l’infinitif). Ainsi le début des matériaux de l’article */‘dorm-i-/ v.intr. ‘dormir’ (Groß/Schweickard 2011–2014 in DÉRom s.v.) se présente comme suit :

⁴⁴ Nous ne donnons pas dans ce chapitre la liste des abréviations utilisées dans le DÉRom, que l’on trouvera ici 519–527. La version à jour de cette liste est consultable sur le site du DÉRom (<www.atilf.fr/DERom>), sous « Consultation du dictionnaire », « Avis au lecteur », « Télécharger la liste des abréviations ». Pour ce qui est de la liste des sigles bibliographiques, le lecteur la trouvera ici 529–617 ainsi que, dans une version régulièrement mise à jour, sur le site du DÉRom, sous « Bibliographie », « Télécharger la bibliographie générale ».

⁴⁵ Cela vaut, à une autre échelle, pour la reconstruction du vocable, qui passe par la reconstruction des lexies qui le composent.

« */dor'm-i-re/ > sard. *dormire* v.intr. ‘être dans un état de sommeil, dormir’ (DES ; PittauDizionario 1 ; AIS 647), [...] ».

4.2.5 Classement des matériaux

Il peut arriver que le chercheur soit amené à reconstruire, à l’intérieur d’un seul article, plusieurs types, caractérisés par plusieurs formes, plusieurs sens, plusieurs combinatoires restreintes et/ou plusieurs réalisations flexionnelles. Les cognats sont triés en fonction du type qu’ils permettent de reconstituer. Dans le cas de */'klam-a-/ (cf. ci-dessus 4.1.3), le classement a ainsi été établi en fonction des sens (liés à la combinatoire) reconstruits. La structure en est la suivante :

- I. Emploi intransitif : ‘crier’
- II. Emploi transitif : ‘appeler’
- III. Emploi prédicatif transitif : ‘proclamer’
- IV.1. Emploi transitif : ‘nommer’
- IV.2. Emploi prédicatif pronominal : ‘s’appeler’

4.2.6 Carte(s)

À partir du deuxième volume du DÉRom, un certain nombre d’articles se voit complété, suite à une suggestion de Bernard Pottier (cf. Buchi/Schweickard 2014b, 31), d’une carte (ou de plusieurs, pour les articles complexes), qui doit montrer la répartition des cognats selon le type qu’ils permettent de reconstruire. Les cartes recensent également les emprunts au protoroman par des langues de familles non romanes. Les cartes, synthétiques, peuvent aider le linguiste à interpréter rapidement les données pour dégager la chronologie des types reconstruits les uns par rapport aux autres. Comme illustration, nous prendrons la carte de l’article */'ϕak-e-/ v.tr. ‘faire’ (cf. ci-dessous figure 13).

Sans même parcourir les matériaux de l’article, la carte nous apprend beaucoup en un temps très restreint. Nous y voyons que l’auteur de l’article (cf. Buchi 2009–2014 in DÉRom s.v. */'ϕak-e-/) a reconstruit deux types qui diffèrent par leur forme, représentés par les mots-formes infinitifs */'ϕak-e-re/ et */'ϕ-a-re/. Alors que */'ϕak-e-re/ s’est maintenu dans tous les parlers romans (au cœur de la Romania), exception faite du végliote, du frioulan, du ladin et du romanche, */'ϕ-a-re/ a des continuateurs dans les idiomes du centre et de l’ouest de la Romania, mais pas en sarde ni en roumain, c’est-à-dire les deux

idiomes qui se sont distingués les premiers du protoroman commun. La présence d'issues de */'ϕak-e-re/ partout, y compris en sarde et en roumain, mène à la conclusion que ce type peut être rattaché à une période du protoroman antérieure à la séparation de ces deux idiomes. Au contraire, l'auteur déduit de l'absence de continueurs du type représenté par */'ϕ-a-re/ dans ces mêmes ensembles de parlers que ce dernier est à rattacher à une période plus tardive du protoroman, postérieure à la séparation du roumain, qui a lieu après celle du sarde. Bien sûr, la carte ne dit pas tout et ne remplace pas le commentaire, mais sa force illustrative est puissante.

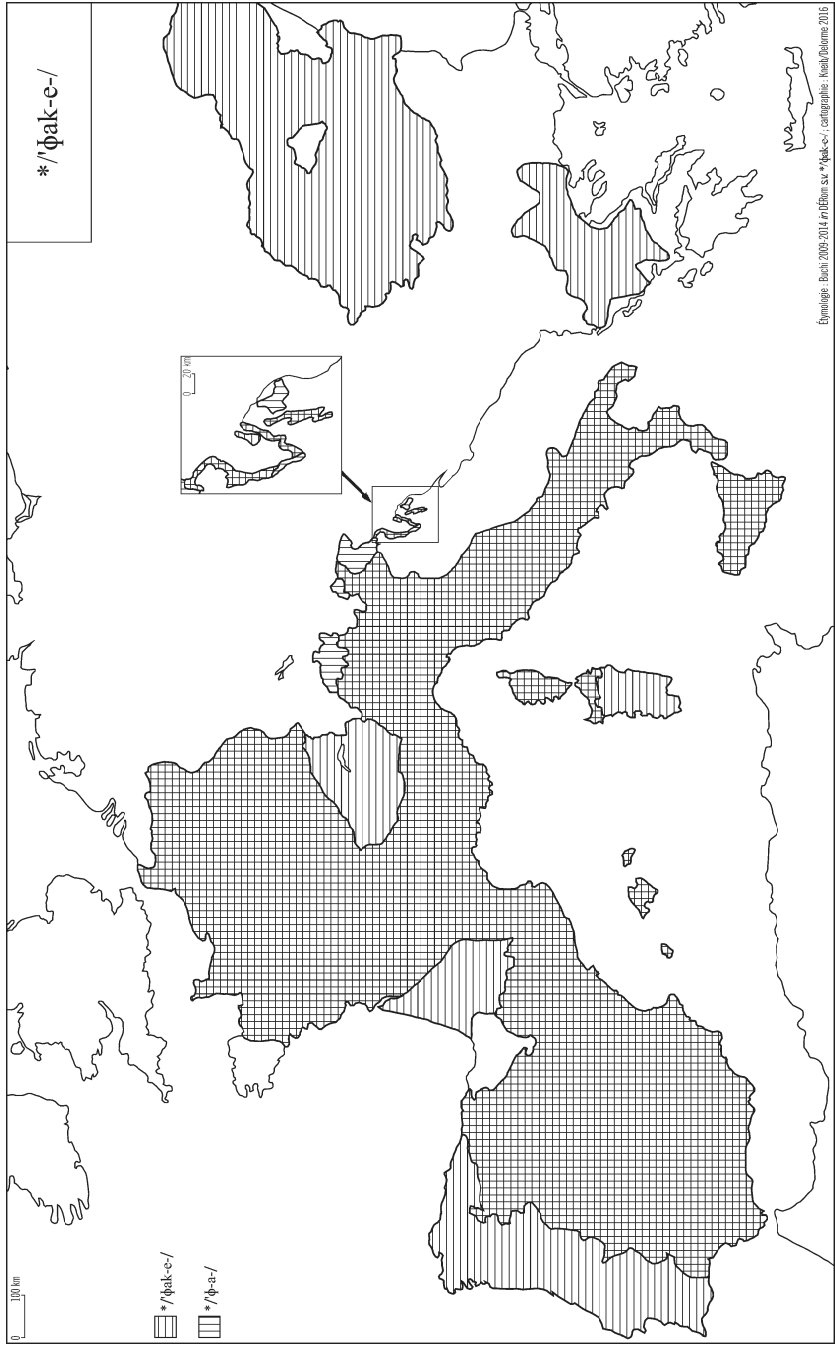


Figure 13 : Répartition des issues des types */ʰak-e-/ et */ʰa-a-/

4.3 Le commentaire

Le commentaire est là pour expliciter l'analyse qui mène à reconstruire, à partir des matériaux romans, l'étymon posé comme lemme. Son rôle, plutôt que d'apporter de nouveaux matériaux, est de montrer la cohérence des données à tous les niveaux : géolinguistique, historique, culturel etc. (cf. Chauveau 2014b). Les évolutions phonétiques, sémantiques et syntaxiques (liées à la combinatoire) qui ne sont pas évidentes et méritent pour cette raison que l'on s'y attarde, ont leur place de droit dans le commentaire. C'est ici aussi que sont expliqués les liens entre les différents types reconstruits : changements de flexion ou de genre, évolutions sémantiques répondant à un mécanisme connu (méta-phore, métonymie) ou d'autres phénomènes encore.

Dans les premiers paragraphes, le commentaire donne obligatoirement les parlers dans lesquels on trouve des cognats permettant la reconstruction. Il indique la forme, la partie du discours et le sens (avec la définition et la glose) de l'étymon et explicite la structure de l'article si les matériaux sont subdivisés en plusieurs paragraphes. Le dernier paragraphe est réservé à la comparaison entre l'étymon protoroman et le corrélat latin écrit de l'Antiquité lorsqu'il en existe un.

4.4 La bibliographie

Le paragraphe consacré aux matériaux offre déjà une littérature abondante, mais il s'agit d'ouvrages d'idiomes particuliers, et non d'ouvrages qui traitent des langues romanes en général. La bibliographie recense les œuvres de ce type qui ont été pertinentes pour la rédaction de l'article.

4.5 Les signatures

C'est à travers cette partie que l'on voit l'énorme travail que représente la rédaction d'un article, avec la mobilisation de compétences scientifiques diverses, concentrées autour d'un même objet, mais chacune avec son rôle propre, sa spécialisation. Le DÉRom est vraiment le fruit d'un travail d'équipe, le bien commun d'une (petite) communauté de chercheurs.

4.6 La date de publication

À travers le champ intitulé « Date de mise en ligne de cet article », le lecteur apprend la date de publication de la première version ainsi que celle de la version la plus récente de l'article en question.

4.7 Les notes

Le dernier champ d'un article du DÉRom, facultatif, concerne les notes en bas de page. Les notes allègent l'article (et notamment le commentaire) en fournissant le détail d'explications que celui-ci ne donne qu'en passant pour rendre sa lecture plus fluide. Elles justifient l'absence de matériaux dans certaines branches romanes si le cas se présente, le choix de tel cognat plutôt que de tel autre pour la liste des matériaux (par exemple le choix d'une forme d'un état ancestral d'une langue romane, pour la raison que l'unité la plus récente constitue un emprunt). Elles font état des difficultés que posent certaines formes pas tout à fait régulières, car l'honnêteté scientifique en toutes choses doit être un grand principe de l'étymologiste. Les notes peuvent également corriger une erreur contenue dans les ouvrages de référence. Enfin, elles précisent les emprunts faits par les langues non romanes au protoroman.

5 Conclusion

Nous avons présenté, tout au long de ce chapitre, un dictionnaire unique dans le domaine de l'étymologie romane : le DÉRom. Nous avons appris à connaître ce qu'il est, sa raison d'être, ce qui le rend si particulier. Puis nous avons cherché à saisir comment il fonctionne. Enfin, nous avons appris à « écouter » son langage pour comprendre ce qu'il a à nous dire. Il reste des choses à expliquer, notamment une grande partie des termes techniques qu'il utilise, mais ce dictionnaire scientifique ne doit plus apparaître comme quelque chose d'inaccessible.

Le DÉRom, dont les résultats de recherche reposent sur la méthode de la reconstruction comparative, apporte des connaissances uniques sur une langue dont nous avons dit au début de ce chapitre qu'elle est un peu la langue de tous. Elle est particulièrement le point de rassemblement de tous les parlers romans, et la reconstruire, forcément dans une perspective panromane, revient à partir sur les traces de l'esprit commun dont sont profondément imprégnés les locu-

teurs de tous ces parlers, là où l'étymologie d'un seul idiome donne à voir une partie de la culture commune de ses locuteurs seulement.

Dans ce dictionnaire, les dialectes sont tout autant considérés (voire davantage valorisés) que les langues standard. C'est l'occasion, avec cette recherche du protoroman, de redécouvrir, pour les uns, la richesse linguistique de leurs ancêtres, et pour les autres, d'être fiers de parler encore un patois.

Nous voulions retrouver un peu de nos racines. La connaissance de la langue ancestrale ne donne qu'une partie du trésor dont nous avons hérité, et cette langue même ne peut pas être connue dans son entier. Cependant, nous pouvons affirmer que le DÉRom remplit sa mission en nous faisant découvrir de nouveaux aspects du latin dit « global ».

6 Bibliographie

- Allettsgruber, Julia, *À la recherche d'une étymologie panromane : lexique héréditaire roman et influence du superstrat germanique dans le DÉRom* (« Dictionnaire Étymologique Roman ») : le cas de */'βad-u/ ~ */'uad-u/ 'gué', in : Anne-Marie Chabrolle-Cerretini (ed.), *Romania : réalité(s) et concepts. Actes du colloque international des 6 et 7 octobre 2011, Université Nancy 2*, Limoges, Lambert et Lucas, 2013, 123–131.
- Andronache, Marta, *Le statut des langues romanes standardisées contemporaines dans le « Dictionnaire Étymologique Roman » (DÉRom)*, in : Emili Casanova Herrero/Cesáreo Calvo Rigual (edd.), *Actas del XXVI Congreso Internacional de Lingüística y de Filología Románicas (Valencia 2010)*, vol. 4, Berlin/New York, De Gruyter, 2013, 449–458.
- Baiwir, Esther, *Un cas d'allomorphie en protoroman examiné à l'aulne du dictionnaire DÉRom*, *Bulletin de la Commission Royale de Toponymie et de Dialectologie* 85 (2013), 79–88.
- Barbosa, Jorge Morais, *Português : Fonética e fonologia*, in : Günter Holtus/Michael Metzeltin/Christian Schmitt (edd.), *Lexikon der Romanistischen Linguistik (LRL)*, vol. 6/2, Tübingen, Niemeyer, 1994, 130–142.
- Baxter, William H., *Where does the « comparative method » come from ?*, in : Fabrice Cavoto (ed.), *The Linguist's Linguist. A Collection of Papers in Honour of Alexis Manaster Ramer*, vol. 1, Munich, LINCOM EUROPA, 2002, 33–52.
- Benarroch, Myriam, *Latin oral et latin écrit en étymologie romane : l'exemple du DÉRom* (« Dictionnaire Étymologique Roman »), in : Maria Helena Araújo Carreira (ed.), *Les Rapports entre l'oral et l'écrit dans les langues romanes*, Saint-Denis, Université Paris 8, 2013, 127–158.
- Benarroch, Myriam/Baiwir, Esther, *Reconstruction flexionnelle*, in : Éva Buchi/Wolfgang Schweickard (edd.), *Dictionnaire Étymologique Roman (DÉRom). Genèse, méthodes et résultats*, Berlin, De Gruyter, 2014, 129–165.
- Bourciez, Édouard/Bourciez, Jean, *Phonétique française. Étude historique*, Paris, Klincksieck, 1967.
- Bourquin, Jacques, *Galerie des linguistes franc-comtois*, Besançon, Presses Universitaires Franc-comtoises, 2003.

- Buchi, Éva, « *Bolchevik* », « *mazout* », « *toundra* » et les autres. *Dictionnaires des emprunts au russe dans les langues romanes. Inventaire – Histoire – Intégration*, Paris, CNRS Éditions, 2010 (= 2010a).
- Buchi, Éva, *Where Caesar's Latin does not belong : a comparative grammar based approach to Romance etymology*, in : Charlotte Brewer (ed.), *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 <<http://ora.ox.ac.uk/objects/uuid%3A237856e6-a327-448b-898c-cb1860766e59>> (= 2010b).
- Buchi, Éva, *Des bienfaits de l'application de la méthode comparative à la matière romane : l'exemple de la reconstruction sémantique*, in : Bohumil Vykypěl/Vít Boček (edd.), *Methods of Etymological Practice*, Prague, Nakladatelství Lidové noviny, 2012, 105–117.
- Buchi, Éva, *Grammaire comparée et langues romanes : la discussion méthodologique autour du « Dictionnaire Étymologique Roman » (DÉRom)*, *Comptes rendus de l'Académie des Inscriptions et Belles-Lettres* 2014/1, 2014, 397–421.
- Buchi, Éva, *Pour une stratification du protoroman*, conférence présentée devant la Société de Linguistique de Paris, séance du 11 avril 2015.
- Buchi, Éva, *Etymological dictionaries*, in : Philip Durkin (ed.), *The Oxford Handbook of Lexicography*, Oxford, Oxford University Press, 2016, 338–349.
- Buchi, Éva/González Martín, Carmen/Mertens, Bianca/Schlienger, Claire, *L'étymologie de FAIM et de FAMINE revue dans le cadre du DÉRom*, *Le français moderne* 83 (2015), 248–263.
- Buchi, Éva/Schweickard, Wolfgang, *Romanistique et étymologie du fonds lexical héréditaire : du REW au DÉRom (« Dictionnaire Étymologique Roman »)*, in : Carmen Alén Garabato/Teddy Arnavielle/Christian Camps (edd.), *La romanistique dans tous ses états*, Paris, L'Harmattan, 2009, 97–110.
- Buchi, Éva/Schweickard, Wolfgang, *Sept malentendus dans la perception du DÉRom par Alberto Várvaro*, *Revue de linguistique romane* 75 (2011), 305–312 (= 2011a).
- Buchi, Éva/Schweickard, Wolfgang, *Ce qui oppose vraiment deux conceptions de l'étymologie romane. Réponse à Alberto Várvaro et contribution à un débat méthodologique en cours*, *Revue de linguistique romane* 75 (2011), 628–635 (= 2011b).
- Buchi, Éva/Schweickard, Wolfgang, *Per un'etimologia romanza saldamente ancorata alla linguistica variazionale : riflessioni fondate sull'esperienza del DÉRom (« Dictionnaire Étymologique Roman »)*, in : Marie-Guy Boutier/Pascalle Hadernann/Marieke Van Acker (edd.), *La variation et le changement en langue (langues romanes)*, Helsinki, Société Néophilologique, 2013, 47–60.
- Buchi, Éva/Schweickard, Wolfgang (edd.), *Dictionnaire Étymologique Roman (DÉRom). Genèse, méthodes et résultats*, Berlin, De Gruyter, 2014 (= 2014a).
- 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 (= 2014b).
- Campbell, Lyle, *Historical Linguistics. An Introduction*, Cambridge, MIT Press, ³2013 [¹1998].
- Chambon, Jean-Pierre, *Remarques sur la grammaire comparée-reconstruction en linguistique romane (situation, perspectives)*, *Mémoires de la Société Linguistique de Paris* 15 (2007), 57–72.
- Chambon, Jean-Pierre, *Pratique étymologique en domaine (gallo)roman et grammaire comparée-reconstruction. À propos du traitement des mots héréditaires dans le « TLF » et le « FEW »*, in : Injoo Choi-Jonin/Marc Duval/Olivier Soutet (edd.), *Typologie et comparatisme. Hommages offerts à Alain Lemaréchal*, Louvain/Paris/Walpole, Peeters, 2010, 61–75.

- Chambon, Jean-Pierre, *Vers une seconde mort du dalmate ? Note critique (du point de vue de la grammaire comparée) sur « un mythe de la linguistique romane »*, *Revue de linguistique romane* 78 (2014), 5–17.
- 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 (= 2014a).
- Chauveau, Jean-Paul, *Structuration des articles et rédaction du commentaire*, conférence présentée à l'occasion de la 2^e École d'été franco-allemande en étymologie romane (ATILF, Nancy, 30 juin–4 juillet 2014), Nancy, ATILF, <<https://apps.atilf.fr/homepages/buchi/wp-content/uploads/sites/7/2015/09/DERom-Ecole-dete-Chauveau-Structuration-des-articles-et-redaction-du-commentaire.pdf>>, 2014 (= 2014b).
- Dardel, Robert de, *Les variantes lexématiques avec l'interfixe /l/ en protoroman*, *Revue de linguistique romane* 70 (2006), 377–407.
- Dardel, Robert de, *Recherches sur le genre roman des substantifs de la troisième déclinaison*, Genève, Droz, 1965.
- Dauzat, Albert, *Compte rendu REW₁ (fascicule 1)*, *Journal des Savants* 9 (1911), 424–426.
- Delorme, Jérémie/Dworkin, Steven N., *Reconstruction microsyntactique*, in : Éva Buchi/Wolfgang Schweickard (edd.), *Dictionnaire Étymologique Roman (DÉRom). Genèse, méthodes et résultats*, Berlin, De Gruyter, 2014, 167–197.
- DÉRom = Buchi, Éva/Schweickard, Wolfgang (dir.), *Dictionnaire Étymologique Roman (DÉRom)*, Nancy, ATILF, <<http://www.atilf.fr/DERom>>, 2008–.
- Engelberg, Stefan, « *Kaisa* », « *Kumi* », « *Karmoból* ». *Deutsche Lehnwörter in den Sprachen des Südpazifiks*, Sprachreport. Informationen und Meinungen zur deutschen Sprache 22/4 (2006), 2–9.
- Gaffiot = Gaffiot, Félix/Flobert, Pierre, *Le Grand Gaffiot. Dictionnaire Latin-Français*, Paris, Hachette, ³2000 [¹1934].
- Garnier, Romain, *Compte-rendu DÉRom 1*, *Bulletin de la Société de Linguistique de Paris* 110/2 (2015), 233–239.
- Grégoire, Henri, *Sur la nécessité et les moyens d'anéantir les patois et d'universaliser l'usage de la langue française. Convention nationale, Instruction publique, séance du 16 prairial l'an II^e [4 juin 1794]*, Paris, Imprimerie nationale, s.d.
- Heidemeier, Ulrike, *Reconstruction dérivationnelle*, in : Éva Buchi/Wolfgang Schweickard (edd.), *Dictionnaire Étymologique Roman (DÉRom). Genèse, méthodes et résultats*, Berlin, De Gruyter, 2014, 211–246.
- Kramer, Johannes, *Contrepoint : ce que j'aurais fait différemment dans le DÉRom*, in : Éva Buchi/Wolfgang Schweickard (edd.), *Dictionnaire Étymologique Roman (DÉRom). Genèse, méthodes et résultats*, Berlin, De Gruyter, 2014, 289–297.
- Maggiore, Marco, *Note di etimologia romanza a margine dell'articolo */kuer-e-/ (quaerëre) 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)*, Bergamo, Sestante Edizioni, 2014, vol. 2, 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.

- Mel'čuk, Igor, *Vers une linguistique Sens-Texte – Leçon inaugurale*, Paris, Collège de France, 1997.
- Neveu, Franck, *Dictionnaire des sciences du langage*, Paris, Armand Colin, 2011.
- Polguère, Alain, *La Théorie Sens-Texte*, *Dialangue* 8/9 (1998), 9–30.
- REW₃ = Meyer-Lübke, Wilhelm, *Romanisches Etymologisches Wörterbuch*, Heidelberg, Winter, ³1930–1935 [¹1911–1920].
- Tadmor, Uri/Haspelmath, Martin/Taylor, Bradley, *Borrowability and the notion of basic vocabulary*, *Diachronica* 27 (2010), 226–246.
- TLFi = Imbs, Paul/Quemada, Bernard (dir.), *Trésor de la langue française informatisé*, Nancy, ATILF, <<http://atilf.atilf.fr/tlf.htm>>, 2004 [1971–1994].
- Vuletić, Nikola, *Le dalmate : panorama des idées sur un mythe de la linguistique romane*, *Histoire Epistémologie Langage* 35 (2013), 45–64.