Buchi, Éva (à paraître) : « Étymologie ». In : Marchello-Nizia, Christiane, Combettes, Bernard, Prévost, Sophie & Scheer, Tobias (éd.) : *Grande Grammaire Historique du Français (GGHF)*. Berlin : De Gruyter Mouton.

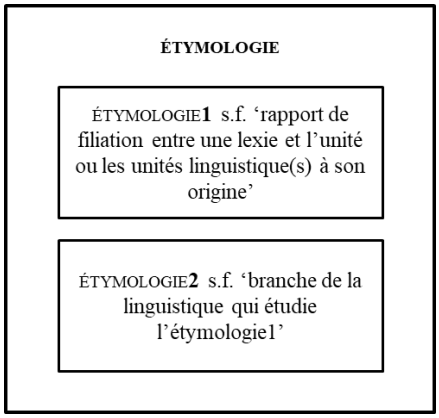
Chapitre 45 Étymologie

45.1 Introduction

Le terme technique *étymologie* connaît (au moins) deux acceptions, actualisées dans les exemples suivants :

> (a) *Sais-tu ce que veut dire le mot prolétaire ? Son étymologie est la même que celle du mot prolifique.* (M. Tournier, *Les Météores*, 1975, Frantext)
>
> (b) *Me direz-vous pour quel motif vous m'infligez ainsi ce cours d'étymologie ? – Pour vous distraire, également pour vous éclairer.* (M. Rheims, *Les Greniers de Sienne*, 1987, Frantext)

Dans la mesure où les items lexicaux polysémiques ont plusieurs étymologies (chaque « sens » mérite d'être étymologisé à part, voir Buchi 2014b), la conceptualisation de l'entité mot telle qu'elle a été réalisée dans le cadre de la Lexicologie Explicative et Combinatoire (LEC), le volet lexical de la Théorie Sens-Texte (voir Polguère 2016[3] [2003[1]] : 47-72 ; Mel'čuk 2012 : 1 : 21-44), se prête particulièrement bien à une utilisation en étymologie. Pour cette raison, ce chapitre applique les outils conceptuels, la terminologie et le marquage typographique de la LEC. Les termes *lexème* 'signe linguistique dont le sens peut être exprimé par un ensemble de mots-formes (et de syntagmes) que seule distingue la flexion' (ainsi GAI1 adj. 'qui a le goût de plaisanter', relié aux mot-formes *gai*1 'qui a le goût de plaisanter', *gaie*1 '*id.*', *gais*1 'qui ont le goût de plaisanter' et *gaies*1 '*id.*') et *vocable* 'regroupement de lexies (c'est-à-dire de lexèmes ou de locutions) associées au même signifiant et présentant un lien sémantique évident' (ainsi GAI, qui regroupe GAI1, GAI2 'qui dispose à la bonne humeur' et d'autres lexèmes) en forment le noyau central. On dira donc que le vocable ÉTYMOLOGIE comprend notamment les lexèmes ÉTYMOLOGIE1 'rapport de filiation entre une lexie et l'unité ou les unités linguistique(s) à son origine' (attesté depuis 1529, Budzinski 2015 : 39 ; voir ci-dessus exemple a) et ÉTYMOLOGIE2 'branche de la linguistique (et plus précisément de la lexicologie) qui étudie l'étymologie1' (attesté depuis 1727, Budzinski 2015 : 39 ; voir ci-dessus exemple b) :



Graphique 1 : Le vocable ÉTYMOLOGIE et ses deux principaux lexèmes

Le but de ce chapitre (voir aussi Holtus 1990, Roques 1990, Fryba-Reber 2003, Gleßgen 2008, Queffélec 2008 et Bertrand 2011[2] [2008[1]]) ne consiste à établir ni de nouvelles étymologies, ni la composition historique du lexique ([pour laquelle ZZ chap. 50]). Il n'est donc pas centré sur l'étymologie1 – dans ce domaine, les progrès se font notamment dans le cadre de projets lexicographiques comme le *Französisches Etymologisches Wörterbuch* (FEW) ou le TLF-Étym –, mais sur l'étymologie2. Dès lors, cette contribution se fixe essentiellement un triple objectif, placé au service de l'entreprise d'ensemble du manuel : dresser l'histoire de l'étymologie (ci-dessous 45.2), présenter les particularités des trois classes étymologiques qui forment le lexique français (45.3) et montrer en quoi l'étude de la disparition d'unités lexicales fait partie intégrante de l'étymologie (45.4) ; *in fine*, quelques pistes pour des recherches futures (45.5).

Tout naturellement, l'objet du chapitre englobe l'ensemble des lexies du français, non seulement les lexèmes (unités lexicales simples), mais aussi les unités polylexicales (locutions). En revanche, les unités infralexicales – dont les affixes dérivationnels, qui relèvent de ce que Cottez (1985[4] [1980[1]] : XIV) appelle la *morphétymologie* – sont traitées ailleurs [ZZ chap. 29].

45.2 Histoire de l'étymologie (française)

La conception antique et médiévale de l'étymologie y voit une sorte d'herméneutique du signe linguistique, au sein de laquelle la mission de l'étymologiste consiste à retrouver le « vrai » sens des mots : le sens du substantif grec ἐτυμολογία s.f., composé de ἔτυμος adj. 'vrai' et λόγος s.m. 'parole', se définit, selon la formule de Zumthor (1958 : 877), comme la 'recherche du vrai (dans ou par les mots)'.

À la Renaissance, l'étymologie, pratiquée abondamment dans les milieux lettrés, prend une dimension comparative (non maîtrisée), les savants cherchant à établir des correspondances entre des langues, comme l'arabe, le grec et le latin, que nous savons maintenant appartenir à des phylums différents. Ces rapprochements sont loin d'être innocents, comme le rappelle Fryba-Reber (2003 : 357) : « il s'agit de rattacher les

langues vernaculaires aux langues nobles ou sacrées que sont la triade latin, grec et hébreu ».

Cette époque voit paraître les premiers glossaires étymologiques français. Dans ses *Tabulae breves Gallicanarum vocum* (1533), Charles de Bovelles propose l'étymologie de 575 vocables, dont tout de même 362 (63%) sont confirmées par la recherche moderne (Dumont-Demaizière 1973 : 251). De son côté, Jean Le Bon publie en 1571 un *Etymologicon françois* réunissant 962 entrées (Jänicke 1991 : 10). Au 17e s. paraît, sous la plume de Gilles Ménage, le premier ouvrage portant explicitement le titre *Dictionnaire étymologique* du français, même si cet intitulé n'apparaît qu'à partir de sa deuxième édition (1694) : la première, datée de 1650, est intitulée *Origines de la langue française*. Sur la base d'un échantillon de 400 entrées – celles de la tranche alphabétique *E-* de l'édition de 1694, à l'exclusion des noms de lieux –, Popelár (1967 : 356) montre que plus de la moitié (55,75%) des étymologies proposées par Ménage résistent à un examen moderne. L'étude d'un autre échantillon, constitué des 412 étymologies exprimées dans la tranche alphabétique *B-* de la même édition du dictionnaire, aboutit à un verdict somme toute similaire : 45,5% d'étymologies justes et 25,5% d'étymologies « partiellement justes » (Leroy-Turcan 1991 : 355). Au-delà de ces statistiques, il reste que l'ouvrage de Ménage relève encore, comme ceux de Bovelles et de Le Bon, de l'étymologie préscientifique, notamment en raison de sa méconnaissance profonde des régularités de l'évolution phonétique et de son incapacité à distinguer avec rigueur lexique héréditaire et emprunts (*ibid.* : 364).

La prochaine étape marquante de l'étymologie française est constituée par l'article *étymologie* (1756) de l'*Encyclopédie* de Diderot et D'Alembert (1751–1772), attribué à Turgot. Bien avant la mise en évidence de l'arbitraire du signe par Saussure (1972 [1906–1911] : 100-102), Turgot y affirme que « les mots n'ont point avec ce qu'ils expriment un rapport nécessaire » (‹http://encyclopedie.uchicago.edu›). Il compare leur étymologie à des énigmes, qu'il convient de résoudre en faisant converger plusieurs critères :

> Quelquefois […] toutes les circonstances s'accordent parfaitement avec la supposition qu'on a faite ; l'accord de chacune en particulier forme une probabilité ; cette probabilité augmente dans une progression rapide, à mesure qu'il s'y joint de nouvelles vraissemblances ; & bientôt, par l'appui mutuel que celles-ci se prêtent, la supposition n'en est plus une, & acquiert la certitude d'un fait. La force de chaque vraissemblance en particulier, & leur réunion, sont donc l'unique principe de la certitude des *étymologies* […], & le fondement de la distinction entre les *étymologies* possibles, probables, & certaines. Il suit de-là que l'art étymologique est, comme tout art conjectural, composé de deux parties, l'art de former les conjectures ou les suppositions, & l'art de les vérifier. (*ibid.*)

Au 19e s., avec la découverte de la grammaire comparée, qui prend son origine dans la linguistique indo-européenne – voir notamment la *Vergleichende Grammatik* de Bopp (1833–1852), dont la réception dans le milieu francophone a bénéficié de la traduction par Michel Bréal (Bopp 1866–1874) –, l'étymologie se constitue véritablement en tant que science. Si l'article « De studio etymologico » de Schlegel (1823) en pose les principes méthodologiques, le *Etymologisches Wörterbuch der romanischen Sprachen* de Diez (1853[1]) en fournit une première application à grande échelle, en l'occurrence pour les langues romanes, qui sera suivie un demi-siècle plus tard par le *Lateinisch-romanisches Wörterbuch* de Körting (1907[3] [1891[1]]). La grammaire comparée et l'étymologie se trouvant dans un rapport de confirmation mutuelle (Swiggers 1997 :

250), ces deux branches de la linguistique se fortifient l'une l'autre et progressent de concert, et la sous-discipline étymologique fait un net bond en avant. En domaine français, c'est le dictionnaire de Scheler (1888[3] [1862[1]]), élaboré « d'après les résultats de la science moderne », comme le précise son titre, qui matérialise le passage à l'ère proprement scientifique de l'étymologie. La préface du dictionnaire fait écho à son titre pour revendiquer le tournant pris : « la divination a fait son temps, et l'étymologie est parvenue au rang d'une science positive, nous dirons même d'une science exacte » (Scheler 1862[1] : II), et l'auteur y rend un vibrant hommage au maître d'outre-Rhin (*ibid.* : II-III) :

> C'est, avant tout, à l'homme éminent, à qui revient la gloire d'avoir le premier fixé et méthodiquement exposé les lois qui président à la formation des langues néo-latines, au vénérable professeur Diez, de Bonn, que nous avons voulu rendre hommage, en consignant dans notre livre, pour mieux les faire valoir en dehors des frontières de sa patrie, ses heureuses découvertes, ses judicieuses démonstrations, ses habiles et prudentes conjectures.

Quant au *Dictionnaire général* (DG) de Hatzfeld, Darmesteter et Thomas (1890–1900), « curieux mélange d'étymologie, d'histoire et d'analyse synchronique à fondement logique » (Roques 1990 : 515), dont la valeur principale réside sans doute dans la qualité des définitions, il innove en accordant une place importante à l'étymologie au sein même de la lexicographie générale, préfigurant ainsi l'étymologie « intégrée » typique de la fin du 20[e] s. (Buchi et Renders 2013 : 660).

Au tournant du 20[e] s., une polémique oppose le Français Antoine Thomas (qui avait apporté son concours au DG) et l'Allemand Hugo Schuchardt : en surface, autour de l'étymologie du verbe *trouver* et de ses congénères, mais plus fondamentalement, au sujet de la primauté de la phonétique ou de la sémantique en étymologie (Pfister et Lupis 2001 : 119-124). Tandis que Thomas, qui met en avant un dérivé de *trŏpus* s.m. 'allégorie', **tropare* v.tr. 'interpréter allégoriquement' > 'inventer' > 'trouver', donne la préférence à « dame phonétique », Schuchardt, qui opte pour une hypothèse prenant son origine dans la terminologie de la pêche, *tŭrbare (aquam)* v.tr. 'agiter (l'eau, afin de guider les poissons vers les filets)' > 'trouver (les poissons)' > 'trouver', se fait le défenseur de « dame sémantique ». Que la recherche moderne ait fini par donner raison à Thomas pour cette étymologie en particulier (voir von Wartburg 1966 *in* FEW 13/2, 318b-323a, *TROPARE) importe peu. On retiendra plutôt qu'à partir de ce moment, toute hypothèse étymologique digne de ce nom se devra de satisfaire non seulement le critère phonétique, mais aussi des contraintes sémantiques (voir Chauveau 2009 pour une démonstration particulièrement convaincante).

Cette nouvelle orientation vers la sémantique se ressent clairement dans le dictionnaire étymologique de Clédat (1929[12] [1912[1]]), dont la préface affirme que « l'histoire des *sons* doit nécessairement céder le pas, dans un dictionnaire de vulgarisation, à l'histoire des *sens* » (Clédat 1920[6] : V), et dont la macrostructure même obéit à un principe sémantique, les vocables étant regroupés par familles lexicales et analogiques.

La publication du *Romanisches Etymologisches Wörterbuch* (REW) de Meyer-Lübke (1930–1935[3] [1911–1920[1]) relègue le dictionnaire de Diez (1887[5] [1853[1]]), et même celui de Körting (1907[3] [1891[1]]), dans les rangs des pionniers méritoires cités désormais exclusivement pour des raisons historiques. Conscient que la dialectologie constitue un des piliers de l'étymologie, Meyer-Lübke prend ses distances avec l'approche monolithique de ses prédécesseurs, qui réduisaient le lexique des parlers

romans à celui de leurs langues-toits standardisées, c'est-à-dire aux langues de référence qu'étaient devenues des topolectes tels le francien ou le toscan, qui s'étaient imposés au détriment d'autres variétés des différents continuums dialectaux romans. Le REW[3] compte ainsi, parmi les 558 abréviations géolinguistiques répertoriées dans sa liste des abréviations, de très nombreux items se référant à des dialectes d'oïl, et ses articles contiennent des données comme wallon *vere* s.m. 'toison ; produit du sol' (*s.v. vĕllus*), picard *raké* v. 'cracher' (*s.v. rak-*), normand *flaner* v.intr. 'flâner' (*s.v. flana*), anglo-normand *tameir* v.tr. 'craindre' (*s.v. tĭmēre*), angevin *ãvẽ* s.m. 'orvet' (*s.v. *anatŏlius*), poitevin *otẽ* s.m. 'vent du sud ; vent du sud-est' (*s.v. altānus*), saintongeais *kukü* s.f. 'ciguë' (*s.v. cĭcūta/*cucūta*), orléanais *kuvrai* s. 'époque où on sème les grains' (*s.v. coopĕrīre*), dialecte du Centre *foirole* s.f. 'mercuriale (plante utilisée comme laxatif)' (*s.v. foriŏlus*), bourguignon *riaule* s.m. 'tire-braise' (*s.v. rŭtabŭlum*), champenois *barde* s.f. 'hache' (*s.v. barda*), lorrain *gosí* s.m. 'coffin du faucheur' (*s.v. *cōtiarium*) ou encore franc-comtois *muklá* s.m. 'hameçon' (*s.v. *mŭscŭlāre*).

Dans le sillage de la controverse entre Schuchardt et Thomas, le 20e siècle voit s'accomplir une importante mutation de l'étymologie (en tout cas romane) avec l'avènement de l'étymologie « histoire du mot », au sein de laquelle la mise en évidence de l'évolution sémantique des vocables analysés joue un rôle central. Impulsée par Gilliéron, Meillet, Spitzer et von Wartburg, la nouvelle approche doit sa formulation la plus mémorable à Baldinger (1959 : 239) :

> La notion même d'étymologie est devenue ambiguë. Désormais il y aura deux sortes d'étymologies : d'une part l'étymologie au sens phonétique, traditionnel, au sens du XIXe siècle : *l'étymologie-origine*. D'autre part, l'étymologie au sens sémantique, moderne : *l'étymologie-histoire du mot*. Les travaux de Gilliéron ont ouvert la brèche. Il a déclaré la guerre à l'ancienne conception de l'étymologie en la comparant à une biographie de Balzac qui ne se composerait que de deux phrases : "Balzac, assis sur les genoux de sa nourrice, était vêtu d'une robe bleue, rayée de rouge. Il écrivit la *Comédie humaine*." L'étymologie, au sens moderne, c'est donc la *biographie* du mot.

De nos jours, l'étymologie française a ainsi définitivement abandonné l'étymologie-origine au profit de l'étymologie-histoire : l'étymon, que l'ancienne approche considérait pour ainsi dire comme autosuffisant, n'en constitue plus que le point de départ.

Le développement de l'étymologie française a particulièrement profité des avancées non seulement factuelles, mais aussi et surtout théoriques et méthodologiques, du *Französisches Etymologisches Wörterbuch* (FEW) de Walther von Wartburg, « un des plus beaux monuments des sciences du langage, […] qui ne cessera d'inspirer des prolongements ou des entreprises parallèles » (Swiggers 1990 : 347). Le sous-titre de cet *opus magnum*, *Eine darstellung des galloromanischen sprachschatzes*, rend compte, mieux que son titre, de l'envergure de son objet, même si le terme *galloromanisch* peut prêter à confusion (de notre point de vue contemporain, il ne peut faire référence qu'à une entité strictement géographique, dépourvue de toute pertinence génétique) : le lexique de l'ensemble des parlers romans qui se sont formés sur le territoire de l'ancienne Gaule. L'objet du FEW englobe donc, certes, la langue d'oïl et le français standardisé, mais aussi le francoprovençal, l'occitan et le gascon, ce dernier constituant génétiquement une langue romane à part entière (voir Chambon et Greub 2002 : 489-492) :

> Un consensus se dessine pour considérer que, tout en étant proche parent, le gascon constitue, du point de vue de sa genèse, un ensemble distinct de l'occitan proprement dit (Chambon, 2002). Mais depuis des siècles, le gascon évolue au contact de l'occitan et en symbiose avec ce dernier ; d'un point de vue synchronique (c'est-à-dire du point de vue de la langue actuelle), le gascon est donc généralement considéré comme une variété d'occitan. (Kremnitz 2007 : 7).

Dès lors, l'histoire de l'étymologie française se confond avec celle de l'étymologie galloromane. Au-delà de son gigantisme – 25 volumes, plus de 14 000 pages, près de 20 000 articles, quatre à cinq millions d'unités lexicales (Büchi 1996 : 6) –, on retiendra du FEW la place accordée à la dialectologie (d'inspiration gilliéronienne), qui dépasse de loin celle que Meyer-Lübke lui réservait dans le REW, la perspective génétique, adoptée dans le but d'expliquer l'évolution des différentes ramifications de chaque famille lexicale, ainsi que sa forte composante structuraliste, que von Wartburg a insufflée à l'œuvre (voir à ce propos aussi le titre révélateur de von Wartburg 1958[5] [1934[1]]) dans le but de toucher du doigt la dynamique et l'interconnexion des changements linguistiques. Opposant le FEW à des dictionnaires philologiques comme le *Oxford English Dictionary* (OED[1]) ou le *Thesaurus Linguae Latinae* (TLL), von Wartburg affirme ainsi avec force que « le FEW, par contre, est un dictionnaire linguistique » (Wartburg 1953 : 98). L'étendue pan-diasystémique et la méthode rigoureuse qui le caractérisent ont fait de ce dictionnaire l'ouvrage de référence en lexicographie historique du domaine considéré : le FEW fonctionne – et fonctionnera encore pendant des générations – à la fois comme lieu de synthèse et comme *tertium comparationis* au niveau galloroman. Mais on aurait tort de ne voir dans le FEW qu'un dictionnaire :

> Menée au terme de sa logique interne et à l'apogée de sa réalisation dans le FEW, l'étymologie wartburgienne se pose comme une linguistique totale du lexique. De celle d'un archéologue préhistorien, la figure de l'étymologiste s'est muée en celle d'un architecte, pour ne pas dire d'un urbaniste. (Chambon et Büchi 1995 : 946).

Les héritiers de von Wartburg ne se sont pas contentés de reproduire sa méthode à l'identique. Le développement de l'étymologie doit notamment beaucoup à Jean-Pierre Chambon, qui a dirigé l'équipe du FEW de 1983 à 1992 et y a impulsé des remaniements méthodologiques importants, porteurs bien au-delà du dictionnaire en tant que tel (voir Greub 2017). Estimant que le FEW, « ce n'est pas seulement un dictionnaire à achever, c'est aussi un esprit et une certaine façon de regarder les faits de langue » (Chambon 1989 : 328), le savant a constamment œuvré pour en faire un objet « à la fois plus philologique (en resserrant la critique des matériaux, en ayant davantage recours aux textes) et plus linguistique encore (en tenant davantage compte de la compétence native, en nous efforçant de mettre en lumière les grands facteurs qui règlent l'histoire des familles lexicales) » (*ibid.* : 336).

La publication de l'index raisonné en deux volumes (Buchi 2003), puis, en 2013, la mise en ligne de la version image du FEW (‹https://apps.atilf.fr/lecteurFEW›), ont contribué à sortir ce chef-d'œuvre de la recherche étymologique de l'état de sous-exploitation déplorable dans lequel il se trouvait notamment en raison de la complexité de ses structures lexicographiques (voir Büchi 1996). Les possibilités d'interrogation (y compris croisées) de l'ouvrage seront multipliées dans la version texte du FEW électronique, actuellement en cours d'élaboration sous la direction de Yan Greub, sur la base d'un modèle réalisé par Pascale Renders (2015).

Avec le *Dictionnaire Étymologique de l'Ancien Français* (DEAF, 1974–), initié par Kurt Baldinger, puis dirigé par Frankwalt Möhren et maintenant par Thomas Städtler, l'étymologie wartburgienne ne s'est pas seulement dotée d'un ouvrage de référence pour un état de la langue commune que le FEW aborde de façon bien trop superficielle, mais aussi d'une nouvelle norme de référence méthodologique, notamment en raison des principes philologiques extrêmement rigoureux qui sous-tendent l'œuvre. Les réflexions publiées en marge du dictionnaire (ainsi Städtler 2012) permettent d'en apprécier l'accomplissement.

Pour une consultation rapide, plusieurs dictionnaires étymologiques en un (ou plusieurs) volume(s) sont à la disposition du public : le *Etymologisches Wörterbuch der französischen Sprache* de Gamillscheg (1969[2] [1928[1]]) ; le Bloch/Wartburg (1968[5] [1932[1]]), qui a bénéficié des apports de première main du FEW ; le *Grand dictionnaire étymologique & historique du français* de Dubois, Mitterrand et Dauzat (2001 [1938[1]]) ; le *Dictionnaire étymologique du français* de Jacqueline Picoche (2008 [1971[1]]) ; le *Dictionnaire historique de la langue française* (ou « Robert historique ») (Rey 2016 [1992[1]]), un bon ouvrage de vulgarisation, pour lequel il faut toutefois tenir compte des mises en garde de Roques (1993) et de Thomassen (1997) ; enfin, le *Dictionnaire étymologique et historique de la langue française* de Baumgartner et Ménard (1996). Si tous ces ouvrages ont leurs mérites, les plus récents ne sont pas les plus fiables, comme le montrent Buchi, Chauveau, Gouvert et Greub (2010) et Buchi, González Martín, Mertens et Schlienger (2015) sur la base de l'analyse de plusieurs cas concrets. La publication d'un dictionnaire étymologique en un volume tenant compte de l'état de l'art constitue donc un *desideratum* de la linguistique française (Fryba-Reber 2003 : 362). En attendant, le *Dictionnaire du français médiéval* de Takeshi Matsumura (2015) rend de très grands services : par la précision de ses définitions, par l'attention portée à la description géolinguistique des données (ainsi « pic. » [picard] *s.v. naviron*, « surtout agn. [anglo-normand] ouest » *s.v. vïoge*, « frpr. [francoprovençal] bourg. [bourguignon] » *s.v. felor*), mais aussi par ses renvois systématiques au FEW et au DEAF.

Enfin, le *Dictionnaire des étymologies obscures* de Pierre Guiraud (1982) occupe une place à part : sur la base des cas de doute de la nomenclature du Bloch/Wartburg (1968[5]), l'auteur propose des étymologies originales pour quelque 1500 vocables français (voir Meier 1984 pour une évaluation d'ensemble de l'entreprise).

45.3 Structuration du lexique en trois classes étymologiques

Comme celui de n'importe quelle langue naturelle (orale), le lexique français se subdivise en trois grandes classes étymologiques : (1) le lexique héréditaire, c'est-à-dire le lexique transmis oralement de génération en génération depuis l'ancêtre commun que le français partage avec les autres langues romanes, que l'on appellera, en fonction du cadre théorique dans lequel on se situe, le protoroman, le latin parlé spontané ou le latin vulgaire ; (2) les emprunts, qui englobent (dans un sens large, incluant les calques) le lexique inspiré de celui de langues avec lesquelles le français a été en contact [ZZ 51.3] ; (3) les créations internes, qui comprennent le lexique formé à l'aide des propres ressources constructionnelles du français, qu'il s'agisse de dérivés, de composés ou de créations d'un type moins répandu, mais aussi les innovations sémantiques spontanées [ZZ chap. 45 et 48].

45.3.1 Lexique héréditaire

45.3.1.1 Cadre théorique

En raison de la surabondance des témoignages écrits du latin, réputé être la langue-mère des langues romanes, l'étymologie romane a longtemps considéré inutile le recours à la grammaire comparée pour l'établissement des étymons du lexique héréditaire, préférant les prélever dans les (excellents) dictionnaires latins disponibles (voir Chambon 2007 et 2010). Cette méthode « latinisante », dans le sens où elle s'appuie fortement sur la philologie latine, continue à donner de beaux résultats de recherche, ainsi dans la pratique lexicographique du FEW et du DEAF. Ce chapitre adopte *a contrario* l'approche méthodologique développée, dans le prolongement des deux articles programmatiques de Jean-Pierre Chambon précités, dans le cadre du *Dictionnaire Étymologique Roman* (DÉRom), qui a donné lieu à un débat paradigmatique animé (voir Greub 2014). Le postulat de base de cette nouvelle approche est que l'étymologie romane héréditaire a tout à gagner à appliquer la même méthode que celle pratiquée pour les autres familles linguistiques, à savoir la reconstruction comparative : il s'agit donc d'une tentative d'arrimage de l'étymologie romane aux courants majoritaires en linguistique historique. Cette option méthodologique implique que le point de départ de la recherche n'est plus constitué d'un étymon latin, mais de cognats romans, c'est-à-dire de lexèmes reliés par un faisceau de correspondances phoniques, sémantiques et combinatoires, dont il convient de reconstruire l'ancêtre commun dans la protolangue, en l'occurrence en protoroman. Le latin écrit n'est pas absent de la démarche, au contraire (voir Maggiore et Buchi 2014). En effet, la reconstruction des étymons protoromans, de nature orale, rend possible une confrontation inédite avec leurs corrélats (ou correspondants codiques) du latin écrit de l'Antiquité, ce qui permet de les situer au sein du diasystème du « latin global » (voir de Dardel 2009 pour ce concept).

La méthode latinisante et la méthode comparative génèrent des résultats de recherche proches, quoique non identiques. Par analogie avec la déclinaison magnétique, cette différence d'angle observable entre le nord magnétique et le nord géographique, on a pu parler de déclinaison étymologique pour caractériser leur écart

(Buchi 2014a) : il est évident que, orientées toutes les deux vers le même point cardinal, elles sont appelées à communiquer entre elles et à s'enrichir mutuellement.

45.3.1.2 Statut du lexique héréditaire

Quoique numériquement relativement faible – selon Wolf (1991[2] [1979[1]] : 50), cela concerne 5% environ des vocables de l'ensemble du stock lexical du 20[e] s. seulement –, le lexique héréditaire joue un rôle primordial au sein de l'architecture du lexique français, et cela au moins à trois titres. Premièrement, le lexique héréditaire en constitue le noyau le plus ancien, présent dès sa formation. S'il est impossible d'assigner une « date de naissance » au français, on peut rappeler, pour donner un ordre de grandeur, que Straka considère, sur la base de la chronologie relative des changements phonétiques, qu'au début du 5[e] s., « le Sud et le Nord de la Gaule ne formaient plus de communauté linguistique » (Straka 1956 : 261), et que leur analyse de légendes monétaires mérovingiennes amène Chambon et Greub (2000 : 174) à la conclusion que « la spécification du francoprovençal para[ît] acquise aux 6[e]/7[e] siècles ». Les lexèmes héréditaires du français remontent donc au minimum à l'époque précédant immédiatement cette période grossièrement délimitée, celle où le français et ses langues sœurs les plus proches (le francoprovençal, l'occitan, le gascon, etc.) formaient encore un seul et même système linguistique (naturellement caractérisé par une variation interne). Deuxièmement, le lexique héréditaire est omniprésent dans le vocabulaire de base (le « français fondamental », utilisé tous les jours par l'ensemble des francophones) ; Wolf (1991[2] [1979[1]] : 50) a ainsi établi, sur la base de Juilland, Brodin et Davidovitch (1970), dont l'analyse se fonde sur un corpus de 500 000 mots de textes publiés entre 1920 et 1940, que sur les cent vocables les plus fréquents du français, 92 sont d'origine héréditaire (voir aussi les statistiques présentées ZZ 47.1-47.3 et 51.3.9). Troisièmement, le lexique héréditaire relie le français à ses langues sœurs (fr. *sel* trouve des cognats dans sarde *sale*, dacoroumain *sare*, italien *sale*, francoprovençal *sal*, occitan *sal*, etc.), et lui confère donc son identité de langue romane. Inversement, il est vrai que le lexique de l'anglais fourmille d'emprunts au français (et notamment à l'anglo-normand, la variété d'ancien français parlée en Angleterre suite à la conquête du royaume anglais par Guillaume le Conquérant en 1066) : parmi les 11 624 vocables des tranches alphabétiques *A-alz* et *M-R* du OED[3] dont la première attestation remonte à la période comprise entre 1150 et 1499, Durkin (2014 : 256-257) recense ainsi 21% d'emprunts au français et 10% d'emprunts qui pourraient remonter soit au français, soit au latin, soit à ces deux langues. Si l'anglais demeure, en dépit de ces chiffres, une langue germanique, c'est en raison du caractère héréditaire non seulement de ses structures morphologiques, mais aussi du noyau originel de son lexique (ainsi anglais *salt* s.m. 'sel', qui remonte, comme ses cognats allemand *Salz*, néerlandais *zout*, etc., au protogermanique **salto*[m], voir OED[2]).

45.3.1.3 Critères permettant d'établir le caractère héréditaire d'une lexie

Afin d'établir le caractère héréditaire d'une lexie française, le chercheur peut s'appuyer sur quatre critères : le sens, la datation, la phonie et l'existence de cognats dans les autres idiomes romans. Si ces critères ne sont pas tous exploitables dans l'ensemble des

cas de figure, ils permettent, combinés entre eux dans des configurations variables, d'asseoir une étymologie de façon assez fiable.

Premièrement, une lexie héréditaire française présente nécessairement un sémantisme ayant correspondu à un besoin de communication de la civilisation romaine (pour les plus anciennes d'entre elles) ou du moins gallo-romaine (pour celles qui remontent à des strates plus récentes de la protolangue), donc d'un peuple constitué très majoritairement d'agriculteurs et d'artisans. Ce critère sémantico-référentiel exclut donc des lexies comme *suzerain* et *vassal*, apparues au Moyen Age en même temps que les réalités qu'elles désignent, *chocolat*, emprunt remontant, comme l'apparition en France de la gourmandise ainsi dénommée, à la Renaissance, ou encore *avion*, dont la formation, comme le moyen de transport que désigne ce substantif, est datée du 19e s. Typiquement, le lexique héréditaire présente un sens plutôt basique. Ainsi, la quasi-totalité des dénominations françaises possibles des « primitifs sémantiques » (équivalents d'universaux lexicaux) du *natural semantic metalanguage* de Wierzbicka (1992 : 150) sont héréditaires :

– *je* 'I' (< protoroman */'ɛgo/), *tu* 'you' < (< */'tu/), *gens* 'people' < (< */'gɛnt-e/) ;

– *ce* 'this' (< */ekke'ɔk/), *même* 'the same' (< */me'tɪps-e/), *autre* 'other' (< */'altr-u/), *un* 'one' (< */'un-u/), *deux* 'two' (< */'dʊ-i/), *tout* 'all' (< */'tɔtt-u/), *moult* 'much/many' (< */'mʊlt-u/ ; évincé dès le moyen français par *beaucoup* et *très*, voir ci-dessous 45.4) ;

– *savoir* 'know' (< */'sap-e-/), *vouloir* 'want' (< */'βɔl-e-/), AF. MF. *cuidier* 'think' (< */'kogit-a-/ ; évincé au 17e s. par le latinisme *penser*), *pouvoir* 'can' (< */'pɔt-e-/), *sentir* 'feel' (< */'sɛnt-i-/), *dire* 'say' (< */'dik-e-/), *faire* 'do' (< */'ɸak-e-/), *arriver* 'happen' (< */ar'rip-a-/) ;

– *bon* 'good' (< */'bɔn-u/), *mauvais* 'bad' (< */mal'fat-i-u/), *grand* 'big' (< */'grand-e/), *petit* 'small' (< */pet'tit-u/) ;

– *très* 'very' (< */'tras/) ;

– *si* 'if' (< */'se/), *car* 'because' (< */'kare/), *non* 'no' (< */'non/), *comme* 'like/how' (< */'komo/) ;

– *quand* 'when' (< */'kuando/), *où* 'where' (< */'ʊβe/), *après* 'after/before' (< */ap'prɛss-u/), *sous* 'under/above' (< */'sʊb-tu/) ;

– AF. *giens* '*kind of' (attesté dans le sens secondaire '[particule renforçant la négation]' ; < */'gɛner-e/ ; évincé à époque prélittéraire par le latinisme *genre*, lequel est concurrencé dès le moyen français par *sorte*, un autre latinisme), *part* 'part of' (< */'part-e/).

Ne font exception que *quelqu'un* 'someone' et *quelque chose* 'something', qui sont des créations internes du français sur la base de lexèmes eux-mêmes héréditaires.

Il est intéressant de comparer cette liste à celle proposée par Bert Peeters (2015), plus riche et basée résolument sur le lexique du français contemporain (*beaucoup* à la place de *moult*, *penser* à la place de *cuidier*, *types* à la place de *giens*) : les ajouts comme *vivre* 'live' (< */'βiβ-e-/) et *mourir* 'die' (< */'mɔr-i-/), non couverts par Wierzbicka (1992 : 150), que fournit cet auteur sont également héréditaires (voir aussi ZZ 47.2).

Appartiennent donc au lexique héréditaire du français les verbes désignant des activités humaines ordinaires (*aller*, *venir* ; *dormir* ; *manger* ; *parler* ; etc.), les noms de parenté de premier degré (*père*, *mère*, *fils*, *fille*), les substantifs désignant des parties du corps (*bouche*, *bras*, *jambe*, *main*, *nez*, *œil*, *pied*, *tête*, etc.), les adjectifs de qualité de

base (*grand*, *petit* ; *gros*, *maigre*, etc.) ou encore les noms des petits nombres (*un*, *deux*, *trois*, etc.).

Pour les langues, tel le français, dont le passage à l'écrit est ancien (en l'occurrence, il remonte au 9e s.), un deuxième critère pour déterminer le caractère héréditaire d'une lexie est constitué par l'ancienneté de ses premières attestations textuelles. Ainsi, parmi les 150 premiers articles du *Dictionnaire Étymologique Roman* (DÉRom), les 142 vocables français concernés (huit articles ne contenant pas de donnée française) sont quasiment tous attestés depuis le Moyen Age :

– 5 items depuis le 9e s. : *chef*, *faire*, *frère*, *laisser*, *part* ;

– 14 items depuis le 10e s. : *boire* (sous la forme *boivre*), *chanter*, *deuil* (sous la forme *duel*), *deux*, *écouter*, AF. *eissir* v.intr. 'sortir', *fuir*, AF. *ierre* s.m. 'lierre', le suffixe adverbial *-ment* (issu, à époque prélittéraire, d'un substantif signifiant 'manière', voir Groß 2011-2016 *in* DÉRom *s.v.* */ˈmɛnt-e/), *montrer* (sous la forme *mostrer*), *ouïr*, *roue* (sous la forme *reue*), *voler*, *votre* ;

– 32 items depuis le 11e s. : *an*, *anneau*, *barbe*, *battre*, *chair*, *choir*, *craindre* (sous la forme *criembre*), *croire*, *dix*, AF. MF. *espardre* v.tr. 'disperser', *faim*, AF. MF. *fraindre* v.tr. 'briser', *gros*, *lieu*, *loin*, *long*, AF. *lor* s.m. 'laurier', *louer* v.tr. 'déclarer (qqn ou qqch.) digne d'admiration ou d'approbation', *mil* s.m. 'millet', *pécher*, *pieu*, *prix*, *quand*, AF. MF. *querre* v.tr. 'quérir', *répondre*, *sauge*, *sourd*, *tailler*, AF. MF. *teil* s.m. 'tilleul', *tendre*, *va* (représentant l'ensemble des mots-formes du radical supplétif en /v-/ du verbe *aller*), *vent* ;

– 77 items depuis le 12e s. : *aigle*, *ail*, *aire* s.f. 'territoire d'un animal sauvage (en particulier nid d'un oiseau de proie)', AF. MF. *aisil* s.m. 'vinaigre', *août*, *arbre*, *arme*, AF. MF. *ascouter* v.tr. 'écouter', *avril*, AF. MF. *avrille* s.m. 'avril', *carême* (sous la forme *caresme* f.), *chaîne*, *charme* s.m. 'arbre de la famille des Bétulacées au bois blanc et dur (*Carpinus betulus* L.)', *châtaigne*, *chercher* (sous la forme *cerchier*), *cheval*, *chevaucher*, *clamer*, *clé*, *croître*, *cul*, *dent*, AF. MF. *deschevauchier* v.intr. 'descendre de selle', *dormir*, *écrire*, *emprunter*, AF. MF. *esvoler* v.intr. 's'envoler', *fève*, *février*, *fil*, *foin* (sous la forme *fein*), *fontaine*, *gésir*, *gras*, *gué*, *herbe*, *lait*, *laver*, *lever*, AF. *loiter* v.intr. 'lutter', *luire* (sous la forme *luisir*), AF. *lume* s.f. 'lumière', *lune*, *mai*, *main*, AF. *maistre* s.f. 'maîtresse', *maître*, *mars*, *merle* s.f. 'merle', *mont*, *montagne*, AF. MF. *moudre* v.tr. 'traire', *mûre* (sous la forme *meure*), AF. MF. *nef* s.m. 'navet', *net*, *nœud*, AF. MF. *noif* s.f. 'neige', *oindre*, *pain*, *péché*, *plaire* (sous la forme *plaisir*), *poids*, *pondre*, *pont*, *prêter*, *rave* (sous la forme *reve*), *rompre*, *rond*, AF. MF. *saiete* s.f. 'flèche', *saluer*, *sel*, *tison*, *tonner*, *trembler*, *vendre*, *venger*, *vin* ;

– 9 items depuis le 13e s. : AF. *aire* adj. 'aigre', *bave* (sous la forme *beve*), AF. MF. *enchevaucher* v.intr. 'monter en selle', *menthe*, *merle* s.m., *moût*, AF. MF. *oint* s.m. 'saindoux', *plantain*, AF. MF. *rore* v.tr. 'ronger' ;

– 2 items depuis le 14e s. : oïlique *emprêter* v.ditr. 'emprunter', *rance* ;

– 2 items depuis le 15e s. : *brume*, *lambruche* s.f. 'vigne sauvage' ;

– 1 item n'étant doté d'aucune attestation médiévale : lorrain *stremi* v.intr. 'frissonner'.

Compte tenu de la densité des textes médiévaux français, notamment des 12e et 13e siècles (voir Möhren 2016 : 897-939), on s'attend donc à ce qu'une lexie héréditaire française soit attestée en principe avant le 13e s. Les exceptions à cette règle doivent être dûment justifiées, ainsi pour *brume*, dont la première attestation remonte seulement au 15e s.. Birrer, Reinhardt et Chambon (2013–2015) *in* DÉRom *s.v.* */'brum-a/ expliquent

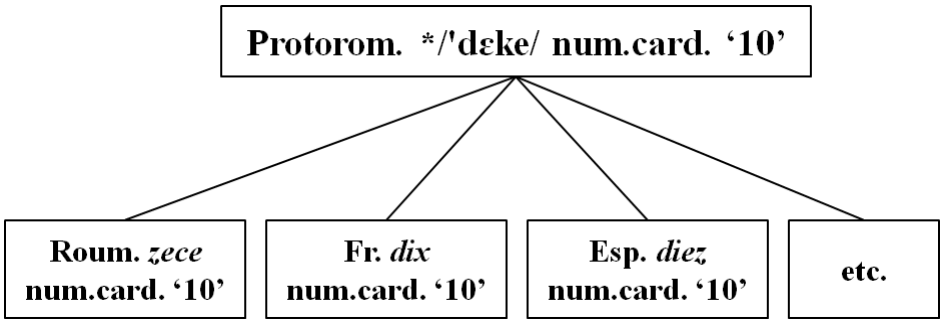
que la datation tardive des continuateurs de */'brum-a/ dans le sens ‘brouillard (surtout brouillard sur mer)’ n’est pas à considérer comme un indice de leur caractère emprunté, mais s’explique par « l’accès tardif au code écrit d’un emploi relevant du vocabulaire spécial des marins ».

En première approche, le critère phonique contribue à déterminer si une lexie donnée est héréditaire (si elle remonte par transmission ininterrompue entre générations successives au protoroman, ou au latin parlé spontané) ou empruntée au latin littéraire, notamment de l’Antiquité. Dans le premier cas, le signifiant de la lexie en question portera la marque de l’usure phonique apposée par les siècles (*aigle* < protorom. régional */'aikul-a/ < protorom. commun */'akuil-a/ s.f. ‘aigle’ [corrélat de lat. *aquila*, Greub 2014–2016 *in* DÉRom *s.v.* */'akuil-a/] ; *chef* < protorom. */'kap-u/ s.m. < */'kaput/ s.n. ‘tête’ [corrélat de lat. *caput*, Schmidt/Schweickard 2015–2016 *in* DÉRom *s.v.* */'kaput/] ; *maître* < protorom. */ma'gɪstr-u/ s.m. ‘maître’ [corrélat de lat. *magister*, Kroyer/Reinhardt 2014 *in* DÉRom *s.v.* */ma'gɪstr-u/]). En revanche, dans le second cas, son signifiant sera proche de l’étymon latin tel qu’on le trouve dans les textes. L’adaptation flexionnelle est toutefois de règle (*aquilin* < lat. *aquilinus* adj. ‘d’aigle’ ; *capital* < lat. *capitalis* adj. ‘qui concerne la tête’ ; *magistrat* < lat. *magistratus* s.m. ‘fonctionnaire public’), et, du moins pour les emprunts les plus anciens, des adaptations phonétiques subséquentes ne sont pas rares. Si l’emprunt à lat. *gloria* s.f. ‘gloire’ se présente ainsi au début – de la fin du 11^{e} s. au milieu du 13^{e} s. – sous la forme attendue *glorie* (> angl. *glory*, OED^{2}), on relève, dès la 2^{e} moitié du 12^{e} s., des attestations du type moderne *gloire* (Gier/Städtler in DEAF G 862).

Le critère phonique est notamment opérant dans le cas des doublets étymologiques, ces paires formées par un lexème héréditaire et un emprunt savant remontant au corrélat latin de l’étymon de cette lexie héréditaire (voir Python 2015) : *cheville* (< protorom. */ka'βɪkl-a/, corrélat de lat. *clavicula*), *nourrisson* (< protorom. */nʊtri-'tion-e/, corrélat de lat. *nutritio*) et *serment* (< protorom. */sakra'mɛnt-u/, corrélat de lat. *sacramentum*) sont héréditaires, tandis que *clavicule* (< lat. *clavicula*), *nutrition* (< lat. *nutritio*) et *sacrement* (< lat. *sacramentum*) sont des emprunts savants.

Plus profondément, le critère phonique – tel qu’il est exploité en linguistique française et romane, il s’agit plus précisément d’un critère phonologique, la réalisation phonétique des phonèmes dans leurs différentes positions à travers les siècles étant difficile à établir avec précision –, combiné avec le dernier critère, le plus puissant, celui de l’existence de cognats, fournit, par le biais de la régularité des évolutions phonologiques, un critère déterminant pour établir le caractère héréditaire d’une lexie.

Les cognats sont les lexies des langues sœurs du français (sarde, roumain, végliote, istriote, italien, frioulan, ladin, romanche, francoprovençal, occitan, gascon, catalan, aragonais, espagnol, asturien, galicien et portugais), dans toutes leurs variétés diasystémiques (notamment dialectales), qui sont apparentées à une lexie française donnée en vertu de leur héritage commun du protoroman. Par exemple, les cognats de fr. *dix* num.card. ‘10’ sont sard. *dèke* num.card. ‘10’, roum. *zece*, végl. *dik*, istriot. *ǧíze*, it. *diece* (> *dieci*), frioul. *dîs*, lad. *dìsc*, romanch. *desch*/*diesch*, frpr. ˹*dies*˺, occit. *detz*, gasc. *dets*, cat. *deu*, arag. *dietz*, esp. *diez*, ast. *diez* et gal./port. *dez* (voir Benarroch 2008-2014 *in* DÉRom *s.v.* */'dɛke/). Cet héritage commun de fr. *dix* et de ses cognats de protorom. */'dɛke/ (corrélat de lat. *decem*) peut être schématisé comme suit :



Graphique 2 : Fr. *dix*, ses cognats romans et leur ancêtre protoroman commun

Pour autant, une seule série de cognats, comme celle formée par roum. *zece*, fr. *dix*, esp. *diez*, etc., n'est pas en mesure de démontrer le caractère régulier des évolutions phonologiques qu'elle exhibe (par exemple, en l'occurrence, celle de protorom. */d-/ > roum. /z-/ ~ fr. /d-/ ~ esp. /d-/, etc.), car des facteurs idiosyncrasiques, comme l'analogie lexicale, pourraient expliquer certains traits phoniques. Il faut donc nécessairement aligner plusieurs séries de cognats présentant le même contexte phonologique pour que la démonstration soit concluante, comme dans le tableau ci-dessous (adapté de Campbell 2013[3] [1998[1]] : 110), qui met en correspondance les séries constituées de fr. *chèvre*, *chair* et *cher* et de leurs cognats.

Protoroman	**Sarde**	**Roumain**	**Italien**	**Français**	**Espagnol**	**Portugais**
*/'**k**a̲pr-a/ (s.f. 'chèvre')	***k**r**a̲**ba* /**k**-**a̲**-/	***c**a̲**p**ră* /**k**a̲-/	***c**a̲**p**ra* /**k**a̲-/	***ch**è̲**v**re* /ʃɛ̲-/	***c**a̲**b**ra* /**k**a̲-/	***c**a̲**b**ra* /**k**a̲-/
*/'**k**a̲rn-e/ (s.f. 'chair')	***k**a̲**r**re* /**k**a̲-/	***c**a̲**r**ne* /**k**a̲-/	***c**a̲**r**ne* /**k**a̲-/	***ch**a̲**i**r* /ʃɛ̲-/	***c**a̲**r**ne* /**k**a̲-/	***c**a̲**r**ne* /**k**a̲-/
*/'**k**a̲r-u/ (adj. 'cher')	***k**a̲**r**u* /**k**a̲-/	_	***c**a̲**r**o* /**k**a̲-/	***ch**ɛ̲**r*** /ʃɛ̲-/	***c**a̲**r**o* /**k**a̲-/	***c**a̲**r**o* /**k**a̲-/

Tableau 1 : La régularité de l'évolution phonologique de protorom. */'ka-/ > fr. /'ʃɛ-/

La comparaison de ces trois séries de cognats incite à postuler, dans un premier temps, la correspondance phonique [sard. roum. it. esp. port. /'ka/ : fr. /'ʃɛ-/], et permet, dans un second temps, d'identifier fr. *chèvre*, *chair* et *cher* comme des continuateurs réguliers de protorom. */'kapr-a/, */'karn-e/ et */'kar-u/, et ainsi d'établir leur caractère héréditaire.

45.3.1.4 Substrat gaulois et superstrat ancien francique

Un lieu commun de l'histoire de la langue française consiste à partir du principe que le lexique français contiendrait des emprunts au gaulois. Le TLF(i), qui, loin d'innover sur ce point, ne fait que refléter la *communis opinio*, analyse ainsi des lexies comme *boue* (« du gaul. **bawa* »), *if* (« du gaul. **ivos* ») ou encore *ruche* (« du gaul. *rusca* ») comme des emprunts que le français aurait faits au gaulois. Or, même s'il est difficile d'établir avec précision les dates de formation et d'extinction du gaulois (Lambert 2003[2] [1994[1]] : 14 : « 300 av. J.-C. – ? 200 apr. J.-C. »), il n'y a guère de doute que cette langue était éteinte au moment de la formation du protofrançais environ au 6e s. Du point de vue du français, *boue*, *if* et *ruche* ne sont donc pas à conceptualiser comme des emprunts, mais comme des lexèmes héréditaires : le processus d'emprunt est intervenu en protoroman (continental italo-occidental) régional de la Gaule (voir aussi ZZ chap. 7.1.1 et 51.2.1).

La situation du superstrat ancien francique (voir Pitz 2006) est plus complexe : tandis que leur diffusion romane assigne certains emprunts comme *blanc* au protoroman, d'autres, comme *jardin*, semblent pouvoir être attribués au protofrançais (Stefenelli 1981 : 83-85 ; 113-120 ; Felixberger 2003 : 600-604 ; ZZ chap. 51.2.1).

45.3.2 Emprunts

45.3.2.1 Vers une définition de l'emprunt linguistique

Le vocable EMPRUNT contient plusieurs lexèmes relevant du domaine de la linguistique. Le TLF(i) en répertorie deux : 'fait pour une langue d'incorporer une unité linguistique, en particulier un mot, d'une autre langue' (processus : « l'emprunt de *bagatelle* à l'italien remonte au 16e s. ») et 'unité de langue incorporée' (résultat : « *bagatelle* est un emprunt à l'italien »), tandis que le *Dictionnaire des sciences du langage* (Neveu 2004) se concentre sur le premier, qu'il définit de façon plus précise : 'processus selon lequel une langue acquiert une unité lexicale intégrée au lexique d'une autre langue'. Le terme technique *emprunt*, pourtant bien lexicalisé, n'est peut-être pas très bien choisi, car il semble sous-entendre qu'une lexie empruntée doit être « rendue » après usage à la langue prêteuse. Pour ce qui est des verbes *incorporer* et *acquérir* dans les définitions ci-dessus, ils prêtent aussi à confusion, dans la mesure où ils semblent sous-entendre l'idée d'une intégration telle quelle d'une unité constituée au préalable. La formule 'processus consistant à créer de nouvelles unités lexicales en imitant tant bien que mal un modèle extérieur' proposée par Thibault (2009 : 11), qui met en avant l'idée d'une création fondée sur l'imitation, est sans doute plus apte à décrire la nature de l'emprunt linguistique. Dans ce cadre conceptuel, et en faisant abstraction du lexème désignant les processus d'emprunt ('fait pour une langue d'incorporer […]' dans la citation du TLF(i) ci-dessus), on peut retenir les définitions suivantes : EMPRUNT**1** 'lexie créée par imitation à la fois formelle et sémantique d'un modèle extérieur' (qui renvoie aux emprunts *stricto sensu*) et EMPRUNT**2** 'lexie créée par imitation d'un modèle extérieur (que ce soit par imitation formelle et sémantique ou par imitation exclusivement sémantique)', catégorie qui englobe les emprunts *stricto sensu*, mais aussi les calques comme fr. *route de la soie* 'piste caravanière qui traversait l'Asie depuis la Chine vers

l'Occident' < all. *Seidenstraße* 'id.' (Rainer 2016 *in* TLF-Étym) ou *laconique* adj. 'qui est exprimé de manière concise' (dp. 1557), calqué de grec λακωνικός 'rappelant les Laconiens par son caractère concis' sur la base de fr. *laconique* 'relatif à la Laconie' (dp. 1372/1374, Städtler 2005 *in* TLF-Étym). Dans le souci d'opérer avec une terminologie univoque, le projet TLF-Étym nomme l'hyperonyme emprunt2 *transfert linguistique*, mais ce terme n'est pas très courant. En tout état de cause, les emprunts1, où l'imitation porte à la fois sur la forme et sur le concept, sont les emprunts les plus prototypiques et les mieux étudiés, ne serait-ce que parce qu'ils sont beaucoup plus faciles à détecter en tant qu'éléments non autochtones que les calques (voir aussi Winter-Froemel 2015 et [ZZ chap. 50]). Pour donner un ordre de grandeur, on peut citer Gleßgen (2008 : 2954), qui relève 6 374 emprunts1 (par ordre décroissant : à l'anglais, à l'italien, à l'allemand, à l'occitan, à l'espagnol, à l'arabe, au néerlandais, au portugais et au russe) dans le Petit Robert 2007.

45.3.2.2 Critères permettant d'établir le caractère emprunté d'une lexie

Pour déterminer qu'une lexie française est empruntée, différents critères ont été proposés par les chercheurs. Cinq d'entre eux sont particulièrement utiles, à condition d'être appliqués de façon complémentaire : l'absence de cognats dans les autres langues romanes, le phonétisme, l'existence d'un contact linguistique avec la langue prêteuse, la datation relative, enfin les indices relevés dans les premières attestations.

Le premier critère ne sert à distinguer les emprunts que du lexique héréditaire (mais pas du lexique de création interne) : constituant le pendant du quatrième critère traité ci-dessus (45.3.1.3), il stipule qu'une lexie empruntée ne connaît pas de cognats dans ses langues sœurs. Ainsi fr. *goéland* s.m. 'espèce d'oiseau palmipède des bords de mer', emprunté à breton *gwelan* s.m. 'id.' (FEW 20, 10b), se traduit par exemple par italien *gabbiano* et espagnol *gaviota* (tous les deux formés sur lat. *gavia* s.f. 'hirondelle de mer', DELI^2 et DCECH), dont le signifiant montre qu'il ne s'agit pas de cognats : en dépit d'une ressemblance formelle de surface (initiale /g/), il n'existe pas de correspondances phoniques régulières qui permettraient de rattacher fr. *goéland* s.m., it. *gabbiano* s.m. et esp. *gaviota* s.f. à un ancêtre commun.

Un deuxième critère, portant sur le signifiant des lexies à analyser, permet lui aussi seulement de différencier les emprunts du lexique héréditaire : les latinismes n'ont pas subi l'effet de l'usure phonétique, même s'ils présentent, comme tous les emprunts, des adaptations au système français (ainsi la perte des suffixes flexionnels latins). Contrairement à *cheville*, *nourrisson* et *serment*, qui appartiennent au lexique héréditaire et présentent à ce titre la marque d'évolutions phonétiques comme la syncope de la voyelle pénultième atone, l'assimilation du groupe */tr/ intervocalique ou encore la fermeture en /ɛ/ de */a-/ initial suivi de yod (voir ci-dessus 45.3.1.3), *clavicule* (< lat. *clavicula*), *nutrition* (< lat. *nutritio*) et *sacrement* (< lat. *sacramentum*) sont des emprunts savants.

C'est un critère de nature non pas linguistique, mais historique, qui sera déterminant pour l'établissement du caractère emprunté d'une lexie : l'existence d'un contact linguistique entre la langue emprunteuse, en l'occurrence le français, et la langue prêteuse à l'époque où l'emprunt présumé a eu lieu. Afin de questionner l'étymologie de fr. *sonnet* s.m. 'espèce de poème strictement codifié' (dp. 1536), que les ouvrages de référence s'accordent à analyser comme un emprunt à l'italien (von Wartburg 1963 *in*

FEW 12, 103a, SŎNUS II 2 ; Bloch/Wartburg (1968[5]); TLF[i]), il convient donc de se demander si au milieu du 16[e] s., la population francophone, ou du moins une frange de la population francophone, était en contact avec la langue italienne. Il s'avère que c'était tout à fait le cas : comme le rappelle Rickard dans son histoire de la langue française (voir déjà Brunot 1906 : 198-212), suite notamment aux premières guerres d'Italie, à partir de 1494, aux cours desquelles la noblesse française était entrée en contact avec une culture qu'elle ne pouvait qu'admirer pour ses prouesses dans le domaine de la musique, des arts ou encore de l'architecture, la culture italienne jouissait d'un très grand prestige dans la société française de la Renaissance. L'auteur juge que « the affectation of Italianisms in speech, and to a lesser extent in writing, was merely a side-effect of Italian influence in general » (Rickard 1989[2] [1974[1]] : 88). Dans son ouvrage *Lexical Borrowing in the Romance Languages*, Hope (1971 : 1 : 148) recense 462 emprunts de l'italien au français au cours du siècle, la période la plus intense se situant entre 1540 et 1560 (*ibid.* : 233), donc quelques années seulement après l'apparition de *sonnet* en 1536.

Contrairement au lexique héréditaire, pour lequel on raisonne en termes de datation absolue (*grosso modo* avant le 13[e] s. dans le cas du français, voir ci-dessus 45.3.1.3), c'est la datation relative qui est pertinente pour le lexique emprunté : on postulera que la lexie qui constitue la source d'inspiration d'une nouvelle lexie (c'est-à-dire son étymon) doit en principe être attestée avant celle-ci. Dans le cas de fr. *sonnet*, attesté depuis 1536, l'étymologiste ne se satisfera ainsi de l'hypothèse d'un emprunt à l'italien qu'après s'être assuré qu'it. *sonetto* est attesté avant cette date, en l'occurrence depuis la fin du 12[e] s. (DELI[2]). Mais la portée de ce critère est limitée : dans le meilleur des cas – le hasard des attestations textuelles fait qu'il existe des contre-exemples –, la datation relative est opératoire comme indice pour les langues prêteuses qui sont passées à l'écrit plus ou moins à la même époque que la langue emprunteuse, en l'occurrence le français. C'est le cas de l'italien, mais aussi de l'espagnol, de l'allemand, de l'anglais et de beaucoup d'autres langues européennes : *alezan* adj. '(cheval) dont la robe est de couleur fauve tirant sur le roux' (dp. 1552) < esp. *alezán* adj. 'id.' (dp. *ca* 1275, Baiwir 2010 *in* TLF-Étym) ; *fricatif* adj. '(consonne) dont l'articulation comporte un resserrement du canal vocal' (dp. 1868) < all. *frikativ* adj. 'id.' (dp. 1854, Mertens 2012 *in* TLF-Étym) ; *score* s.m. 'nombre de points obtenus dans une compétition sportive' (dp. 1896) < angl. *score* s. 'id.' (dp. 1742, Lupu 2006 *in* TLF-Étym). En revanche, le critère de la datation relative ne sera d'aucun secours dans les cas où l'inspiration vient d'une langue dont le passage à l'écrit est plus récent que celui du français, voire d'une langue non écrite. Ainsi fr. *chourin* s.m. 'couteau', emprunté à romani *čuri* s.m. 'id.' (Gossen 1964 : 263), connaît son premier témoignage écrit en 1837, alors que les textes en langue romani sont rares avant le 20[e] s. (Courthiade 2008 : 922). De la même manière, on relève le terme technique de saline marine *fare* s.m. 'œillet par où passe l'eau salée', emprunté – à travers le français technique régional – à breton vannetais *fard* s.m. 'id.', depuis 1404 en français, tandis que sa première attestation bretonne ne remonte qu'à 1744 (Buron/Baudinot 2006 *in* TLF-Étym).

La recherche, puis l'analyse minutieuse, des premières attestations textuelles constitue un des piliers de la recherche étymologique : « sul più generale piano del metodo vogliamo […] far rilevare come, molte volte, una prima attestazione già contenga in sè la chiave della parola da giudicare » (Pfister et Lupis 2001 : 95). Des indices particulièrement décisifs peuvent être glanés dans les premiers témoignages

écrits de lexèmes empruntés ; dans les cas les plus favorables, la langue d'origine est même explicitement mentionnée. Ainsi *bicyclette*, à la consonance si française, est en réalité emprunté à l'anglais (plus précisément à un nom de modèle de cycle anglais), comme le montre la première attestation, datable du 28 janvier 1886 et tirée de la revue *Le Véloce-sport : organe de la vélocipédie française* (Hütsch 2013–2014 *in* TLF-Étym) :

> Étranger. – Angleterre. – La maison 'Rudge' met en vente cette année un nouveau genre de bicycle de route qui a nom la *bicyclette*. La roue motrice, qui est la plus grande, se trouve derrière. Plusieurs amateurs, entre autres MM. Corey et Chambet, ont essayé cette nouvelle machine et en disent merveille. Les connaisseurs prononcent ce *Safety* comme imbattable pour les courses. Du reste nous donnerons sous peu tous les détails sur la construction de ce nouveau bicycle.

Le fait qu'un néologisme apparaisse d'abord dans une traduction constitue aussi un indice intéressant pour orienter vers l'hypothèse d'un emprunt. Ainsi *dialectologie* s.f. 'étude linguistique des dialectes', qui, au vu de ses propriétés morphologiques (synchroniques), aurait très bien pu être de création française (TLF[i] : « dér. de *dialecte* ; suff. *-logie* »), semble, au vu du contexte de sa première apparition, en 1810, dans la traduction du titre d'un ouvrage allemand – ainsi que de plusieurs autres attestations anciennes faisant référence à un milieu germanique –, constituer un germanisme (Deflorenne 2009 *in* TLF-Étym) :

> Essai sur la langue maltaise, considérée comme dérivant de l'ancienne langue punique ; mémoire pour servir à la dialectologie arabe (traduction de *Versuch über die Maltesische Sprache [zur Beurtheilung der neulich wiederholten Behauptung, dass die ein Überrest der altpunischen sey, und als Beytrag zur arabischen Dialektologie]*).

De son côté, *chocolat* s.m. 'aliment fait avec des fèves de cacao torréfiées et broyées' est attesté pour la première fois, sous la forme *chocolate*, dans la traduction française, datée de 1598, de la *Historia natural y moral de las Indias* (1590) du jésuite espagnol José d'Acosta, qui œuvrait en tant que missionnaire et naturaliste en Amérique latine. La nature du texte où ce premier témoignage de *chocolat* est relevé, de même que son phonétisme hispanisant, constituent de solides arguments pour rattacher le terme à esp. *chocolate* s.m. 'id.' (TLF[i]).

Sans que ce soit un véritable critère décisionnel, on peut en outre se demander quels types de lexies se prêtent plus ou moins bien à être empruntés (voir aussi ZZ 51.2.2). Sur la base d'une étude comparative des emprunts lexicaux dans 41 langues du monde, Tadmor, Haspelmath et Taylor (2010), les auteurs de la « liste Leipzig-Jakarta » du vocabulaire de base (qui renouvelle la « liste Swadesh » précédemment utilisée), mettent en évidence les tendances suivantes : les items à sens lexical s'empruntent plus facilement que ceux à sens grammatical, les substantifs plus facilement que les verbes, les adjectifs et les adverbes, enfin, les lexèmes à contenu culturel (religion, habillement, foyer, domaine juridique, relations sociales, agriculture, alimentation, etc.) plus facilement que ceux à contenu plus universel (pronoms personnels comme 'il/elle' et 'nous', verbes concrets comme 'démanger' et 'monter', adverbes et pronoms interrogatifs comme 'comment ?', 'lequel ?', 'où ?', 'pourquoi ?', etc.).

45.3.2.3 Processus d'adaptation

Au cours de leur diffusion au sein de la communauté francophone et de leur intégration au système linguistique du français, qui peuvent durer plus ou moins longtemps, les emprunts lexicaux subissent, à des degrés variables, différents types d'adaptations.

Dans le domaine phonétique, on pensera par exemple à des adaptations (segmentales et suprasegmentales) du type de celles qui font passer d'anglais [rɪ'pɔːtəʳ] (*reporter*) à français [ʁəpɔχ'tɛːʁ] (*reporter*) ou d'italien [balle'rina] (*ballerina*) à français [bal'ʁin] (*ballerine*). La situation se complique singulièrement quand l'emprunt passe (surtout) par la langue écrite.

L'adaptation graphique est évidente dans les cas où on change de système d'écriture (emprunts à l'arabe, au chinois, au grec, etc.), mais elle concerne aussi d'autres cas, comme le remplacement systématique de la majuscule initiale par une minuscule dans le cas des emprunts à des substantifs allemands (all. *Krach* s.m. 'effondrement des cours de la Bourse' > fr. *krach* s.m. 'id.') (voir aussi ZZ 51.1).

Au niveau sémantique, il faut tenir compte du fait que l'emprunt, qui est dans un premier temps un néologisme, doit « faire son nid » au sein du réseau lexical du français, dont la structure est forcément différente de celle du réseau lexical de la langue prêteuse. En empruntant angl. *building* s. 'bâtiment', par exemple, le français lui a attribué le sens 'vaste immeuble moderne à nombreux étages' : fr. *building* s.m. a ainsi acquis un sémantisme qui l'articule avec les autres membres de son champ lexical en vertu des relations lexicales – d'hyponymie par rapport à *bâtiment*, *édifice*, *immeuble* ou peut-être même *maison*, d'opposition par rapport à *cabane*, *case* ou encore *hutte*, de (para-)synonymie par rapport à *tour* – qu'il noue avec eux (voir aussi ZZ chap. 51).

Dans le domaine syntaxique (plus précisément microsyntaxique), une adaptation est nécessaire quand il y a désaccord entre les catégories syntaxiques de la langue prêteuse et celles du français. En empruntant des substantifs anglais, le français doit ainsi systématiquement les intégrer à un de ses genres, souvent, comme dans le cas de *building*, au masculin, mais on trouve aussi le féminin (*cordite*, *moleskine*, *performance*, etc.).

L'adaptation flexionnelle des emprunts, pourtant tout à fait banale, passe souvent inaperçue, car les locuteurs, mais aussi les chercheurs, ont tendance à concevoir les lexies sous leur seule forme citationnelle : infinitif pour les verbes, singulier pour les noms, masculin singulier pour les adjectifs. Or, même dans un cas relativement simple comme *building*, l'adaptation flexionnelle est bien réelle : son pluriel n'est pas marqué (sauf à l'écrit), alors que l'étymon anglais a [-z] au pluriel. Pour l'intégration flexionnelle de l'emprunt français à russe большевик (*bol'ševik*) s.m. 'membre de la fraction majoritaire du parti social-démocrate russe, favorable aux thèses de Lénine en 1903, qui arriva au pouvoir par la révolution d'octobre et forma par la suite le parti communiste russe', cas plus complexe, on peut distinguer trois stades formellement distincts qui se suivent et même s'enchevêtrent en l'espace de quelques mois (voir Buchi 2010 : 80-92) : dans un premier temps, le néologisme, introduit par la presse écrite, présente le pluriel russe en *-i* (degré zéro de l'adaptation flexionnelle : a) ; puis la désinence *-i* perd son association univoque avec le pluriel (on trouve aussi des attestations du type *bolcheviki* adj.f.sg. : b) ; enfin la marque *-i* disparaît complètement, et le pluriel se forme dorénavant de façon parfaitement régulière par l'adjonction (à l'écrit) de *-s* (c).

(a) *Bolchevik* s.m.sg. (22/11/1917) ~ *Bolchevik**i*** s.m.pl. (28/04/1918)
(b) *bolchevik* s.m.sg. (1918) ~ *bolchevik**is*** s.m.pl. (15/03/1918)
(c) *bolchevik* s.m.sg. (1918) ~ *bolchevik**s*** s.m.pl. (dp. 1918)

Pour ce qui est de la morphologie constructionnelle, l'emprunt à большевик pourra également servir d'exemple pour l'adaptation qui y est à l'œuvre, car le substantif français *bolchevik* a d'abord donné lieu, par création interne (conversion), à l'adjectif *bolchevik*, puis la finale *-ik* de ces deux lexies, qui représente le suffixe russe -ик (formateur de noms d'agent, voir Timberlake 1993 : 855), *a priori* inanalysable en français (a), a été réinterprétée en *-ique*, suffixe formateur d'adjectifs de relation comme *alcoolique* ou *électronique* (b).

(a) *Bolchev**ik*** (22/11/1917–1933), *bolschew**ik*** (23/12/1917), *bolchev**ick*** (dp. 24/07/1918), *bolchev**ik*** (dp. 1918), *bolchév**ik*** (1919–1939), *Bolchév**ik*** (19191957), *bolschev**ik*** (1934)
(b) *bolchév**ique*** (dp. 1931), *bolchev**ique*** (dp. 1933)

À travers cet ensemble d'adaptations, phonétiques, sémantiques, syntaxiques et morphologiques, l'avantage de conceptualiser les emprunts comme des imitations plutôt que comme des incorporations ou des acquisitions (voir ci-dessus 45.3.2.1) apparaît clairement.

45.3.2.4 Emprunts « internes »

La sélection lexicale du français, intervenue dans le cadre de sa standardisation progressive, a maintes fois donné la préférence à des lexèmes non franciens comme *caillou*, la variante normanno-picarde de francien *chaillou* (Fouché 1961 : 3 : 557), ou *foin*, la variante orientale (et peut-être plus particulièrement bourguignonne, la Bourgogne ayant joué le rôle de grenier pour Paris) de francien *fein* (Bourciez et Bourciez 1967 : 79).

Au sein du français standardisé moderne, on dénombre des emprunts « internes », c'est-à-dire des emprunts aux parlers pour lesquels le français fonctionne comme langue-toit : les dialectes d'oïl, le francoprovençal, l'occitan et le gascon. On a longtemps voulu croire que ces emprunts, qui ne méritent, *stricto sensu*, la qualification d'internes que dans les cas où l'inspiration vient des dialectes d'oïl, étaient très nombreux. En réalité, un ensemble de travaux réalisés par et autour de Jean-Pierre Chambon a montré qu'il s'agit là d'une sorte d'illusion d'optique (Chambon et Chauveau 2004) : à y regarder de près, les emprunts du français aux dialectes galloromans sont très rares.

Ainsi, une enquête conduite par Chambon (1997) sur les 19 emprunts présumés à l'occitan auvergnat répertoriés par Gebhardt (1974) fait apparaître que le français général ne connaît pas, à proprement parler, d'emprunts aux parlers auvergnats. Ce qu'il contient, en revanche – mais cette connaissance ne s'est livrée qu'au prix de l'élaboration d'une importante documentation historique spécifique –, c'est un ensemble de lexies dont l'origine doit être cherchée dans le français régional pratiqué en Auvergne. L'exemple de *fourme* s.f. 'type de fromage au lait de vache, à pâte non cuite, produit dans le Massif Central ; pièce de ce fromage, de forme cylindrique' est révélateur à ce propos. Considéré en général comme un occitanisme (voir par exemple TLF[i] : « terme de l'Auvergne, empr. à l'a. prov. *forma* 'meule de fromage' »), *fourme* constitue, quand il accède, au milieu du 19^{e} s., au français de référence, plus précisément un emprunt aux

variétés du français parlées dans les régions productrices de ce type de fromage (Auvergne et Rouergue), au sein desquelles l'emprunt à l'occitan s'était produit auparavant (voir Chambon 1997 : 391-392) :

> (a) Attestations en français de référence (lexicographie générale) : depuis 1803 (*forme*) ; depuis 1845 (*fourme*).
>
> (b) Attestations en français régional d'Auvergne : 1421 (*fourme*) ; 1535 (*forme*) ; 1589 (*fourme*) ; 1609 (*fourme*) ; 1736 (*fourme*) ; 1756 (*fourme*) ; 1778 (*forme*) ; 1790 (*fourme*) ; 18e s. (*fourme*) ; 1803/1804 (*fourme*) ; 1845 (*fourme*) ; 1860 (*fourme*) ; 1861 (*fourme*) ; 1863 (*fourme*) ; 1903 (*fourme*).
>
> (c) Attestations en ancien occitan auvergnat : 1380 (*forma*) ; 1533/1534 (*forma*) ; av. 1689 (*fourma*).

C'est donc le concept de dérégionalisation qui décrira le mieux le processus à l'œuvre dans ces cas de figure.

De la même manière, Gouvert (2007) montre que, à l'exception probable de *piolet*, le français de référence n'a pas emprunté de lexies au francoprovençal : le processus d'emprunt est systématiquement passé (ainsi pour *chalet*, *luge* ou encore *sérac*) par des variétés régionales du français.

À la lumière de ces travaux, il conviendrait de soumettre les listes d'emprunts internes présumés (Guiraud 1978[3] [1968[1]] : 114-126 ; Walter 1988 : 165-166 : Holtus 1990 : 522-523 ; Swiggers 1998 : 70-73 ; Trotter 2006 : 1780-1783) à un examen linguistique et philologique approfondi, dans le but de déterminer la véritable part des emprunts internes au sein du lexique français.

45.3.3 Créations internes

45.3.3.1 Typologie des créations internes

On peut distinguer deux sous-classes au sein de la classe des créations internes : celles qui donnent lieu à de nouveaux vocables (CAPITAL > CAPITALISME) et celles qui donnent lieu « seulement » à de nouveaux lexèmes au sein de vocables déjà constitués (ainsi EXPLOSION**1** 'libération violente et bruyante d'air ou de matière lors d'une réaction chimique, physique ou physiologique' > EXPLOSION**2** 'manifestation soudaine et violente', comme dans *explosion de joie*). Dans le premier cas, l'analyse étymologique rejoint la formation des mots, dans le second, elle se fait sémantique historique. Comme le rappelle Durkin (2009 : 94), la formation des mots est tout à fait centrale pour l'étymologie :

> An understanding of word-formation processes is crucial to etymological research. Let us imagine for a moment a future age where only a few scraps of modern English survive. If an etymologist in this future age encounters the word *bitterness* in a scrap of surviving writing, and the word *bitter* is not attested in the surviving records, but *dark*, *darkness*, and *hardness* all are, then comparison among these words will provide just enough evidence to show that *bitterness* should be analysed as *bitter-ness* (not **bit-terness*), that *-ness* is a derivative suffix, and that *bitter* is an adjective.

La sémantique historique occupe une place tout aussi primordiale dans la recherche étymologique (voir Blank 1997), mais plutôt en tant que soubassement théorique,

l'apparition de nouveaux lexèmes étant moins systématiquement étudiée que celle de nouveaux vocables, ou en tout cas de façon plus superficielle.

La sous-classe des créations internes donnant lieu à la formation de nouveaux vocables (dans un premier temps monolexématiques) se subdivise elle-même en une multitude de types de formations, dont les suivants [ZZ chap. 29 et 47] :

Dérivés affixaux : *ânerie* s.f. 'acte d'une personne sotte' < *âne* s.m. 'personne sotte' + *-erie* '(suffixe formant des dérivés nominaux désignant un trait de caractère ou un comportement)' (TLF[i])
Conversions : *pédaler* v.intr. 'actionner les pédales d'un cycle' (dp. 1885 ; *-er* s'analyse comme un suffixe flexionnel) < *pédale* s.f. 'pièce fixée à chaque extrémité des manivelles du pédalier d'un cycle et sur laquelle agit le pied' (dp. 1867, Badiou-Monferran 2015 *in* TLF-Étym)
Dérivés régressifs : *aristocrate* s.m. 'membre de l'aristocratie' (1550 ; dp. 1778) < *aristocratie* s.f. 'forme de gouvernement où le pouvoir appartient aux nobles' (dp. 1370, Roques 1986 *in* FEW 25, 230b, ARISTOKRATÍA 3)
Composés : *malbouffe* s.f. 'ensemble d'habitudes alimentaires caractérisées par la consommation riche et déséquilibrée de produits industriels dans l'élaboration desquels certains excès priment l'intérêt nutritionnel' (dp. 1979) < *mal* adj. 'mauvais' (dp. 881) + *bouffe* s.f. 'nourriture' (dp. 1926, familier, Delorme 2015 *in* TLF-Étym)
Ellipses : *imparfait* s.m. 'temps verbal du passé présentant le procès comme non accompli' (dp. 1596) < *prétérit imparfait* loc.m. 'id.' (dp. début 14^{e} s., Holbecq et Merlo 2008 *in* TLF-Étym)
Confixés : *claustrophobie* s.f. 'angoisse morbide ressentie dans un endroit clos' (dp. 1879) < *claustro-* 'fermeture' + *-phobie* 'angoisse morbide' (Steinfeld 2005 *in* TLF-Étym)
Déonomastiques (dont l'étymon est un nom propre) : *libaniser* v.tr. 'rendre semblable au Liban' (1986) < nom de pays *Liban* + *-iser* '(suffixe formateur de dérivés verbaux véhiculant le sens 'transformer')' (Schweickard 1992 : 105)
Délocutifs (dont l'étymon est un fragment du discours) : *cessez-le-feu* s.m. 'arrêt négocié des combats' (dp. 1949) < *Cessez le feu !* (Büchi 1995 : 148)

45.3.3.2 Critères permettant d'établir qu'une lexie est de création interne

Les critères que l'on peut convoquer pour décider si une lexie est de création interne dépendent du type de formation, morpho-syntaxique (« constructionnelle ») ou sémantique. Pour la sous-classe des créations internes de nouveaux vocables, trois critères sont opératoires : la datation relative, la compositionalité sémantique et, afin d'exclure l'hypothèse d'un calque, la comparaison avec les langues avec lesquelles le français a pu être en contact au moment de la formation du vocable.

Le critère de la datation relative stipule que l'ensemble des constituants génétiques (lexèmes et affixes) du nouveau lexème (qui formera en même temps un nouveau vocable, monolexématique) doivent avoir existé antérieurement à sa formation. Le critère est par exemple rempli pour le dérivé *détachable* adj. 'que l'on peut détacher' (dp. 1879, TLF[i]), dont les éléments constitutifs, *détacher* v.tr. 'dégager (qqch. ou qqn) de ce qui (l')attache' (dp. 1165, TLF[i]) et *-able* '(suffixe exprimant la possibilité)' (dp. 12^{e} s., TLF[i]), sont attestés bien avant qu'on le relève à la fin du 19^{e} s.

S'il s'agit d'un signe complexe (les cas prototypiques étant les dérivés et les composés, tandis que les conversions, par exemple, en sont exclues), le critère déterminant sera en réalité constitué de la compositionalité sémantique. En effet, l'étymologiste part du principe qu'à leur naissance, les créations lexicales plurimorphé-

matiques sont sémantiquement transparentes : l'opacité sémantique qui caractérise certaines d'entre elles est à considérer comme l'effet de leur évolution ultérieure. Dès lors, le sens du néologisme formé par création interne doit être calculable à partir de la combinaison du sens de chacun de ses éléments de formation. Pour pouvoir s'appuyer sur ce critère dans l'analyse d'un dérivé, l'étymologiste se retourne vers les sources qui fournissent des informations fiables sur les propriétés non seulement sémantiques, mais aussi combinatoires (bases sélectionnées, particularités morpho-phonologiques) des affixes au moment où la dérivation est censée être intervenue. Dans le domaine français, on pensera notamment au volume 3 « Formation des mots » de la grammaire historique de Nyrop (1908) et, pour le vocabulaire savant, à l'ouvrage de Cottez (1985[4] [1980[1]]), mais aussi aux entrées affixales du TLF(i), qui sont en général très fournies et contiennent souvent des informations précieuses sur la productivité des préfixes et suffixes au cours des siècles.

Un raté ponctuel épinglé par Chauveau (2005 : 4-6) illustrera l'application de ce critère. L'étymologie de *badinage* s.m. 'plaisanterie légère' (dp. 1541) du TLF(i) se lit ainsi : « dér. de *badin*[1] [s.m. 'personne qui aime plaisanter'] ; suff. *-age* ». L'entrée *-age* du TLF(i) précise que le suffixe est utilisé pour créer des noms d'action sur des bases verbales (*laver → lavage*, *nettoyer → nettoyage*, *tisser → tissage*) et des noms collectifs sur des bases nominales (*branche → branchage*, *outil → outillage*, *rayon → rayonnage*). Or, *badinage* est initialement un nom d'action ('action de plaisanter') et non pas un nom collectif (*'ensemble de plaisantins'), raison pour laquelle il faut corriger l'étymologie de *badinage* du TLF(i) en « dérivé de *badiner* v.intr. 'plaisanter' [dp. 1541], + suffixe *-age* », formule étymologique qui explicite la compositionalité sémantique du dérivé ('action de' + 'plaisanter').

Ces deux critères, la datation relative et la compositionalité sémantique, étant insuffisants pour distinguer les créations internes des calques structurels (comme le terme de phonétique *labialisation*, calqué d'all. *Labialisierung* sur la base de l'adjectif *labial*, Mertens 2012 *in* TLF-Étym), une comparaison avec les principales langues de contact du français au moment de la formation du vocable en question s'impose. Sur cette base, on peut par exemple établir que fr. *mazouter* v.intr. 'faire le plein de mazout' (dp. 1936) représente une conversion de fr. *mazout* s.m. 'fuel' (Buchi 2010 : 295), et donc une création interne, car aucun étymon-candidat crédible ne se présente en dehors du français (Ø russe *мазутить v., Ø anglais **mazut* v. etc.).

Exclure l'hypothèse d'un calque représente une préoccupation encore plus centrale dans le cas des créations internes formées par innovation sémantique. Ainsi, afin de valider l'étymologie de Badiou-Monferran 2015 *in* TLF-Étym, qui analyse PÉDALER**2** v.intr. 'bouger vite' (dp. 1891, dans un article du *Figaro* : « la course [Paris-Brest et retour] aura eu [la conséquence] de modifier certaines locutions de l'argot parisien. *On ne dira plus* : […] Détaler […] *Mais on dira* : […] Pédaler ») comme le résultat d'une évolution sémantique métaphorique à partir de PÉDALER**1** v.intr. 'actionner les pédales d'un cycle' (dp. 1885), il convient de montrer que l'anglais, la langue qui a le plus influencé la terminologie française du cycle (voir ci-dessus 45.3.2.2 et TLF-Étym *s.v. bicyclette* et *cycliste*), ne possède pas d'étymon valable pour PÉDALER**2**, ce qui est le cas : angl. *pedal* v. ne connaît pas le sens 'bouger vite'.

45.3.3.3 Créations internes diatopiquement marquées

Quelquefois, les innovations sémantiques restent confinées à des aires régionales plus ou moins larges, ce qui peut donner lieu à des désaccords dénotatifs entre variétés diatopiques voisines et entre variétés régionales et français standard. Les dénominations de deux légumes, la carotte et la betterave, étudiées par Chambon (2006), en constituent un exemple instructif. À la fin du 16^{e} s., la betterave se propage en France, comme en témoigne Olivier de Serres dans son traité d'agriculture (1603) :

> Une especе de Pastenades [‘carotte’] est la Bette-Raue, laquelle nous est venue d'Italie n'a pas long temps. C'est une racine fort rouge, assés grosse, dont les fueilles sont des Bettes, et tout cela bon à manger, apareillé en cuisine : voire la racine est rengee entre les viandes delicates, dont le jus qu'elle rend en cuisant, semblable à syrop au succre, est tres-beau à voir pour sa vermeille couleur.

Dans la plus grande partie du domaine francophone, c'est le composé *betterave* (< *bette* s.f. ‘poirée’ + *rave* s.f. ‘navet’ : on peut y voir l'idée d'une bette ronde) qui s'est imposé pour désigner ce nouveau légume. En revanche, le français de Lyon a opté pour *carotte*, mettant ainsi fin, dans son régiolecte, à la concurrence synonymique entre *carotte* (qui désigne désormais la betterave) et *pastenade*, qui conserve le sens ‘carotte’. Lyon a de tout temps été un centre directeur linguistique, et on ne s'étonnera pas que la ville a diffusé le lexème *carotte* ‘betterave’ (attesté depuis 1722, mais sûrement plus ancien à l'oral) dans une assez large zone d'influence (Rhône, Loire, Haute-Loire, Drôme, Isère, Ain). Le résultat de cette innovation sémantique, c'est que *carotte* est devenu ambigu en français : en fonction de la région d'origine ou de résidence du locuteur, le terme désignera la carotte ou la betterave.

45.4 Disparition de lexies

Traditionnellement, l'intérêt des chercheurs en étymologie s'est beaucoup plus porté sur l'apparition des lexies que sur leur disparition, même si dans le cadre de l'étymologie-histoire des mots, il devrait paraître normal que la « mort » des unités lexicales fasse partie, tout autant que leur « naissance », du programme visé. Si la disparition de lexies concerne l'ensemble des classes étymologiques (lexique héréditaire, emprunts et créations internes), elle a une conséquence particulièrement tangible pour le lexique héréditaire, qui ne se renouvelle pas : petit à petit, sa part au sein du lexique français est appelée à s'amenuiser. À titre d'exemple, il est significatif que parmi les 142 vocables concernés des 150 premiers articles du DÉRom (huit d'entre eux ne présentant pas de donnée française), 21 lexèmes n'ont pas survécu jusqu'au français d'aujourd'hui :

> AF. *aire* adj. ‘aigre’, AF. MF. *aisil* s.m. ‘vinaigre’, AF. MF. *ascouter* v.tr. ‘écouter’, AF. MF. *avrille* s.m. ‘avril’, AF. MF. *deschevauchier* v.intr. ‘descendre de selle’, AF. *eissir* v.intr. ‘sortir’, AF. MF. *enchevaucher* v.intr. ‘monter en selle’, AF. MF. *espardre* v.tr. ‘disperser’, AF. MF. *esvoler* v.intr. ‘s'envoler’, AF. *fraindre* v.tr. ‘briser’, AF. *ierre* s.m. ‘lierre’, AF. *lor* s.m. ‘laurier’, AF. *maistre* s.f. ‘maîtresse’, AF. MF. *moudre* v.tr. ‘traire’, AF. MF. *nef* s.m. ‘navet’, AF. MF. *noif* s.f. ‘neige’, AF. MF. *oint* s.m. ‘saindoux’, AF. MF. *querre* v.tr. ‘quérir’, AF. MF. *rore* v.tr. ‘ronger’, AF. MF. *saiete* s.f. ‘flèche’, AF. FM. *teil* s.m. ‘tilleul’.

À l'exception de *ierre* et *querre*, supplantés respectivement par *lierre* (variante dotée de l'article agglutiné) et *quérir* (variante flexionnelle), ces lexies ont été évincées par des

concurrents synonymiques, qu'ils soient eux-mêmes héréditaires, empruntés ou formés par création interne. Un exemple concret, étudié par Marchello-Nizia (2006), illustrera l'imbrication de l'apparition et de la disparition d'unités lexicales. On peut légitimement se demander pourquoi *moult*, la réalisation courante du primitif sémantique 'much/many' en ancien français, s'est fait évincer, entre le 13e et le 16e s. (voir ci-dessus 45.3.1.3), par *beaucoup* (essentiellement devant verbe ou nom) et *très* (essentiellement devant adjectif ou adverbe), d'autant que la grande majorité des cognats de *moult* (roumain *mult*, italien *molto*, espagnol *muy*/*mucho*, etc., voir REW[3] *s.v. mŭltus*) se sont maintenus jusqu'à nos jours. Se situant dans une perspective structurale de macro-grammaticalisation, Marchello-Nizia (*ibid.* : 137-179) fournit la clé de cette énigme. L'adverbe *beaucoup* tire de son sémantisme originel ('coup frappé avec force et élégance') un atout déterminant : exprimant originellement la vigueur d'une action humaine, il est prédisposé à la subjectivation, qui constitue un facteur déclencheur de grammaticalisation. Le passage de l'idée d'intensité d'un coup porté par un agent humain à la notion abstraite d'intensité (et de quantification) s'accompagne d'un ensemble de restrictions : morphologique (singulier), combinatoire (perte de déterminant), enfin lexicale (disparition de la variante *grand coup*). On passe ainsi du syntagme libre originel (a) d'abord à une locution (b), puis à un composé adverbial (c) (*ibid.* : 146) :

(a) {(un/Ø/*le[s]*) + (*beau[s]*/*grant[z]*) + (*coup[s]*)} {*de* + SN} ;
(b) {(Ø) (*beau*/*grant*) + (*coup*)} *de* {SN} ;
(c) {*beaucoup* (*de*) N}.

Mais le déclin de *moult* ne peut pas être expliqué par la seule montée en puissance de *beaucoup* ; il est aussi facteur de l'extension des emplois de *très* (< */'tras/), originellement cantonné à la fonction de préfixe verbal, qui a été le premier concurrent de *moult* dans certains de ses emplois. En effet, au moment où *très* commence à pouvoir déterminer des adjectifs et des adverbes, *moult* est pris en tenaille entre ses deux concurrents, qui finissent par assumer de façon complémentaire ses anciennes fonctions. Le processus aboutit en fin de compte à une distinction systémique entre un « intensifieur-quantifieur portant sur les catégories majeures, ou 'têtes', du nom et du verbe (éléments constitutifs de l'énoncé), et [un] intensifieur portant sur les catégories adjointes que sont l'adjectif et l'adverbe » (*ibid.* :178).

Un objectif séduisant que la communauté des étymologistes pourrait se donner consisterait, sur la base d'études de ce type, à tenter de modéliser le phénomène de la disparition d'unités lexicales en tant que « 'processus' dont les étapes rest[ai]ent à circonscrire, et l'éventuelle gradualité (obsolescence, désattestation, perte) à définir » (Badiou-Monferran et Verjans 2015 : 8). De telles études pourraient aussi se poser la question d'éventuelles survivances dans le registre ludique (ainsi *moult* en français contemporain) de lexies par ailleurs disparues (voir Winter-Froemel 2018 : 252-253).

45.5 Perspectives

L'étymologie française, qui s'appuie sur une tradition qui la dépasse en partie, est parvenue à un niveau de maturation méthodologique très avancé : la communauté des chercheurs actifs dans ce domaine, qui entretiennent des liens d'échange et d'entraide

notamment dans le cadre de projets lexicographiques comme l'ALW (qui, en dépit de son orientation atlantographique, contient une forte dimension étymologique), le DEAF, le DÉRom, le FEW ou encore le TLF-Étym, a atteint une grande maîtrise collective des instruments de travail à convoquer et des comportements de recherche à adopter dans le but de retracer l'origine, la formation et l'évolution du lexique français. Les pratiquants de l'étymologie française ont aussi conscience du degré d'autarcie variable de cette dernière, qui, en réalité, ne se suffit à elle-même que quand elle traite le lexique de création interne, mais ne se conçoit qu'au sein de l'étymologie romane quand elle aborde le lexique héréditaire (voir Buchi, Chauveau, Gouvert et Greub 2010) et doit même considérer l'ensemble des langues du monde quand elle porte sur le lexique emprunté.

Cela ne veut pas dire – tant s'en faut ! – que l'ensemble du lexique français dispose d'ores et déjà d'une étymologie solidement établie. Non seulement certaines étymologies erronées, même une fois réfutées, ont la vie étonnamment dure dans la conscience collective (par exemple le prétendu russisme *bistrot*, de création interne française, voir Chauveau 2001 *in* FEW 22/2, 259a-260b), mais, comme en attestent les volumes 21 à 23 du FEW, consacrés aux matériaux d'origine inconnue ou incertaine, une portion non négligeable du lexique français et galloroman résiste (malgré les efforts méritoires de Baldinger 1988–2003) encore à toute étymologisation définitive, dont des vocables fréquents comme CAMION, FOULARD, GAUCHE, MOQUETTE, REGRETTER ou encore RÊVER (Buchi et Dworkin à paraître : 18-19). Il va de soi que ces cas devront être traités avec la même rigueur scientifique que celle qui a prévalu à l'établissement des étymologies que l'on considère désormais comme acquises. L'approche proposée par Dalbera (2006), imaginative du point de vue de la reconstruction sémantique, mais marquée par une non prise en compte de la phonétique historique et de la chronologie des sens, une approche assez laxiste de la morphologie constructionnelle et un manque de discussion des étymologies de ses prédécesseurs (voir Chambon 2008 et Chauveau 2008), ne sera donc pas la panacée ici.

En outre, il ne faut pas croire que la marge de progression est nulle pour les unités déjà dûment pourvues d'un étymon. Au sein du lexique héréditaire, il sera ainsi du plus grand intérêt de départager les lexies qui remontent au protoroman *stricto sensu* (l'ancêtre commun que le français partage avec l'ensemble des parlers romans) de celles qui ont pris naissance au sein de strates plus récentes de la protolangue. Une première tentative tout à fait prometteuse dans ce sens a été réalisée par Mertens (2018 : 278-279) pour les dérivés en */'ur(-a)/ (> fr. *couture*, *fressure*, *mesure*, etc.). Dans le domaine des emprunts, la situation est très inégale en fonction des langues d'origine, et peu d'ensembles bénéficient d'un traitement aussi abouti que celui que Höfler (1982) réserve aux anglicismes ou Arveiller (1999) aux emprunts à l'arabe, au persan et au turc (et à quelques autres langues « orientales »). En outre, la part des calques sémantiques au sein du stock lexical français – et en particulier parmi ses locutions, qui ont moins bénéficié de l'attention des savants que les lexèmes – est certainement encore largement sous-évaluée. Pour ce qui est du lexique de création interne, un *desideratum* criant frappe les lexies formées par innovation sémantique, car bien souvent, les étymologies connues se limitent, à strictement parler, au lexème originel d'un vocable donné, alors que, comme le montrent les travaux pionniers conduits dans le cadre du TLF-Étym, l'établissement de l'origine et du milieu créateur de chaque lexème individuel entrant

dans la formation d'un vocable complexe apporte à l'étymologie de ce dernier une profondeur structurelle qui en décuple l'intérêt.

Références bibliographiques

ALW ; Arveiller 1999 ; Badiou-Monferran et Verjans 2015 ; Baldinger 1959, 1988–2003 ; Baumgartner et Ménard 1996 ; Bertrand 2011[2] [2008[1]] ; Blank 1997 ; Bloch/Wartburg ; Bopp 1866–1874 ; Bourciez et Bourciez 1967 ; Brunot 1906 ; Büchi 1995, 1996 ; Buchi 2003, 2010, 2014a, 2014b ; Buchi, Chauveau, Gouvert et Greub 2010 ; Buchi et Dworkin à paraître ; Buchi, González Martín, Mertens et Schlienger 2015 ; Buchi et Renders 2013 ; Budzinski 2015 ; Campbell 2013[3] [1998[1]] ; Chambon 1989, 1997, 2006, 2007, 2008, 2010 ; Chambon et Büchi 1995 ; Chambon et Chauveau 2004 ; Chambon et Greub 2000, 2002 ; Chauveau 2005, 2008, 2009 ; Clédat 1929[12] [1912[1]] ; Cottez 1985[4] [1980[1]] ; Courthiade 2008 ; Dalbera 2006 ; Dardel 2009 ; DCECH ; DEAF ; DELI[2] ; DÉRom ; DG ; Diderot et D'Alembert 1751–1772 ; Diez 1887[5] [1853[1]] ; Dubois, Mitterand et Dauzat 2001 [1938[1]] ; Dumont-Demaizière 1973 ; Durkin 2009, 2014 ; Felixberger 2003 ; FEW ; Fouché 1952–1961 ; Frantext ; Fryba-Reber 2003 ; Gamillscheg 1969[2] [1928[1]] ; Gebhardt 1974 ; Gleßgen 2008 ; Gossen 1964 ; Gouvert 2007 ; Greub 2014, 2017 ; Guiraud 1978[3] [1968[1]], 1982 ; Höfler 1982 ; Holtus 1990 ; Hope 1971 ; Jänicke 1991 ; Juilland, Brodin et Davidovitch 1970 ; Körting 1907[3] [1891[1]] ; Kremnitz 2007 ; Lambert 2003[2] [1994[1]] ; Leroy-Turcan 1991 ; Maggiore et Buchi 2014 ; Marchello-Nizia 2006 ; Matsumura 2015 ; Meier 1984 ; Mel'čuk 2012/2013 ; Ménage 1694[2] [1650[1]] ; Mertens 2018 ; Möhren 2016 ; Neveu 2004 ; Nyrop 1904–1930 ; OED[1] ; OED[2] ; OED[3] ; Peeters 2015 ; Petit Robert 2007 ; Pfister et Lupis 2001 ; Picoche 2008 [1971[1]] ; Pitz 2006 ; Polguère 2016[3] [2003[1]] ; Popelár 1967 ; Python 2015 ; Queffélec 2008 ; Renders 2015 ; REW ; Rey 2016 [1992[1]] ; Rickard 1989[2] [1974[1]] ; Roques 1990, 1993 ; Saussure 1972 [1906–1911] ; Scheler 1888[3] [1862[1]] ; Schlegel 1823 ; Schweickard 1992 ; Städtler 2012 ; Stefenelli 1981 ; Straka 1956 ; Swiggers 1990, 1997, 1998 ; Tadmor, Haspelmath et Taylor 2010 ; Thibault 2009 ; Thomassen 1997 ; Timberlake 1993 ; TLF(i) ; TLF-Étym ; TLL ; Trotter 2006 ; Walter 1988 ; Wartburg 1958[5] [1934[1]], 1953 ; Winter-Froemel 2015, 2018 ; Wierzbicka 1992 ; Wolf 1991[2] [1979[1]] ; Zumthor 1958.

E.B.

Grande Grammaire Historique du Français

Chapitre 45 : Étymologie

Items à intégrer à la bibliographie générale

ALW = Remacle, Louis, Legros, Élisée, Lechanteur, Jean, Counet, Marie-Thérèse, Boutier, Marie-Guy et Baiwir, Esther (1953–). *Atlas linguistique de la Wallonie. Tableau géographique des parlers de la Belgique romane*. Liège : Vaillant-Carmanne/Université de Liège.

Arveiller, Raymond (1999). *Addenda au FEW XIX (Orientalia)*, édités par Max Pfister. Tübingen : Niemeyer.

Badiou-Monferran, Claire et Verjans, Thomas (2015). Avant-propos. In Cl. Badiou et Th. Verjans (éd.) Disparitions. Contributions à l'étude du changement linguistique. Paris : Champion, 7-25.

Baldinger, Kurt (1959). L'étymologie hier et aujourd'hui. *Cahiers de l'Association internationale des études françaises* 11, 233-264.

Baldinger, Kurt (1988–2003). *Etymologien. Untersuchungen zu FEW 21-23*, 3 vol. Tübingen : Niemeyer.

Baumgartner, Emmanuèle et Ménard, Philippe (1996). *Dictionnaire étymologique et historique de la langue française*. Paris : Librairie Générale Française.

Bertrand, Olivier (2011[2] [2008[1]]). *Histoire du vocabulaire français. Origines, emprunts et création lexicale*. Palaiseau : Les Éditions de l'École Polytechnique.

Blank, Andreas (1997). *Prinzipien des lexikalischen Bedeutungswandels am Beispiel der romanischen Sprachen*. Tübingen : Niemeyer.

Bloch/Wartburg = Bloch, Oscar et Wartburg, Walther von (1968[5] [1932[1]]). *Dictionnaire étymologique de la langue française*. Paris : PUF.

Bopp, François (1866–1874 [original allemand : 1857–1860[2] (1833–1852[1])]). *Grammaire comparée des langues indo-européennes comprenant le sanscrit, le zend, l'arménien, le grec, le latin, le lithuanien, l'ancien slave, le gothique et l'allemand*, 5 vol., trad. M. Bréal. Paris : Imprimerie Impériale.

Bourciez, Édouard et Bourciez, Jean (1967). *Phonétique française. Étude historique*. Paris : Klincksieck.

Brunot, Ferdinand (1906). *Histoire de la langue française des origines à 1900. Tome II : Le Seizième Siècle*. Paris : Colin.

Büchi, Eva (1995). Typologie des délocutifs galloromans. In *Estudis de lingüística i filologia oferts a Antoni M. Badia i Margarit*. Barcelone : Publicacions de l'Abadia de Montserrat, vol. 1, 141-163.

Büchi, Eva (1996). *Les Structures du « Französisches Etymologisches Wörterbuch ». Recherches métalexicographiques et métalexicologiques*. Tübingen : Niemeyer.

Buchi, Éva (éd.) (2003). *Französisches Etymologisches Wörterbuch. Index*, 2 vol. Paris : Champion.

Buchi, Éva (2010). Bolchevik, mazout, toundra *et les autres. Dictionnaire des emprunts au russe dans les langues romanes. Inventaire – histoire – intégration*. Paris : CNRS Éditions.

Buchi, Éva (2014a). Les langues romanes sont-elles des langues comme les autres ? Ce qu'en pense le DÉRom. Avec un excursus sur la notion de déclinaison étymologique. *Bulletin de la Société de linguistique de Paris* 109/1, 257-275.

Buchi, Éva (2014b). What are etymological (and etymographical) units made of : vocables or lexemes ? Communication présentée à la 7th International Conference on Historical Lexicography and Lexicology, Las Palmas de Gran Canaria 9-11 juillet.

Buchi, Éva, Chauveau, Jean-Paul, Gouvert, Xavier et Greub, Yan (2010). Quand la linguistique française ne saurait que se faire romane : du neuf dans le traitement étymologique du lexique héréditaire. In Fr. Neveu, V. Muni Toke, Th. Klingler, J. Durand, L. Mondada et S. Prévost (éd.) *Congrès Mondial de Linguistique Française – CMLF 2010*. Paris : Institut de Linguistique Française, 111-123, ‹http://dx.doi.org/10.1051/cmlf/2010025›.

Buchi, Éva et Dworkin, Steven N. (à paraître). Etymology in Romance. In M. Loporcaro (éd.) *Oxford Research Encyclopedia of Romance Linguistics*. Oxford : Oxford University Press.

Buchi, Éva, González Martín, Carmen, Mertens, Bianca et Schlienger, Claire (2015). L'étymologie de FAIM et de FAMINE revue dans le cadre du DÉRom. *Le français moderne* 83, 248-263.

Buchi, Éva et Renders, Pascale (2013). Gallo-Romance I : Historical and etymological lexicography. In R. H. Gouws, U. Heid, W. Schweickard, H. E. Wiegand (éd.) *Dictionaries. An International Encyclopedia of Lexicography. Supplementary Volume : Recent Developments with Focus on Electronic and Computational Lexicography*. Berlin/Boston : De Gruyter Mouton, 653-662.

Budzinski, Laure (2015). *Étude historique et étymologique de la terminologie linguistique française*. Thèse de doctorat. Université de Lorraine.

Campbell, Lyle (2013[3] [1998[1]]). *Historical Linguistics. An Introduction*. Cambridge/Massachusetts : MIT Press.

Chambon, Jean-Pierre (1989). Tradition et innovation dans la refonte du *FEW*. In D. Kremer (éd.) *Actes du XVIII[e] Congrès International de Linguistique et de Philologie Romanes (Université de Trèves [Trier] 1986)*. Tübingen : Niemeyer, vol. 7, 327-337.

Chambon, Jean-Pierre (1997). Les emprunts du français moderne aux dialectes ou patois. *Lalies. Actes des sessions de linguistique et de littérature* 17, 33-53.

Chambon, Jean-Pierre (2006). Pour l'étude des régionalismes du français moderne : *carotte* 'betterave' et formations connexes. *Revue belge de philologie et d'histoire* 84, 737-770.

Chambon, Jean-Pierre (2007). Remarques sur la grammaire comparée-reconstruction en linguistique romane (situation, perspectives). *Mémoires de la Société de Linguistique de Paris* 15, 57-72.

Chambon, Jean-Pierre (2008). Compte rendu de Jean-Philippe Dalbera, *Des dialectes au langage. Une archéologie du sens*, Paris, 2006. *Bulletin de la Société de linguistique de Paris* 103/2, 274-277.

Chambon, Jean-Pierre (2010). Pratique étymologique en domaine (gallo)roman et grammaire comparée-reconstruction. À propos du traitement des mots héréditaires dans le *TLF* et le *FEW*. In I. Choi-Jonin, M. Duval et O. Soutet (éd.) *Typologie et comparatisme. Hommages offerts à Alain Lemaréchal.* Louvain/Paris/Walpole : Peeters, 61-75.

Chambon, Jean-Pierre et Büchi, Eva (1995). "Un des plus beaux monuments des sciences du langage" : le FEW de Walther von Wartburg (1910–1940). In G. Antoine et R. Martin (éd.) *Histoire de la langue française 1914-1945*. Paris : CNRS-Éditions, 935-963.

Chambon, Jean-Pierre et Chauveau, Jean-Paul (2004). Un cas de dialectologite, ou le français rendu invisible : à propos des vues de Pierre Gardette sur francoprovençal *polailli* et moyen français régional *poulaille* 'poule'. *Bulletin de la Société de linguistique de Paris* 99/1, 155-180.

Chambon, Jean-Pierre et Greub, Yan (2000). Données nouvelles pour la linguistique galloromane : les légendes monétaires mérovingiennes. *Bulletin de la Société de linguistique de Paris* 95/1, 147-181.

Chambon, Jean-Pierre et Greub, Yan (2002). Note sur l'âge du (proto)gascon. *Revue de linguistique romane* 66, 473-495.

Chauveau, Jean-Paul (2005). Remarques sur la dérivation dans les notices historiques et étymologiques du *Trésor de la langue française*. In É. Buchi (éd.) *Actes du Séminaire de méthodologie en étymologie et histoire du lexique (Nancy/ATILF, année universitaire 2005/2006)*. Nancy : ATILF, ‹http://www.atilf.fr/IMG/pdf/seminaires/Seminaire_melh_Chauveau_2005-11-16.pdf›.

Chauveau, Jean-Paul (2008). Compte rendu de Jean-Philippe Dalbera, *Des dialectes au langage. Une archéologie du sens*, Paris, 2006. *Revue de linguistique romane* 72, 244-246.

Chauveau, Jean-Paul (2009). De la nécessité pour l'étymologie de reconstituer l'histoire des sens. *Recherches Linguistiques de Vincennes* 38 : 13-44.

Clédat, Léon (1929[12] [1912[1]]). *Dictionnaire étymologique de la langue française*. Paris : Hachette.

Cottez, Henri (1985[4] [1980[1]]). *Dictionnaire des structures du vocabulaire savant. Éléments et modèles de formation*. Paris : Le Robert.

Courthiade, Marcel (2008). Rom. In E. Graf, H. Rouanet et J. Mayorkas (éd.) *Encyclopaedia universalis*. Paris : Encyclopaedia universalis France, vol. 20, 921-926.

Dalbera, Jean-Philippe (2006). *Des dialectes au langage. Une archéologie du sens*. Paris : Champion.

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

DCECH = Corominas, Joan et Pascual, José Antonio (1980–1991). *Diccionario crítico etimológico castellano e hispánico*, 6 vol. Madrid : Gredos.

DEAF = Baldinger, Kurt *et al*. (1974–). *Dictionnaire Étymologique de l'Ancien Français*. Québec/Tübingen/Paris/Berlin : Presses de l'Université Laval/Niemeyer/Klincksieck/De Gruyter.

DELI[2] = Cortelazzo, Manlio et Zolli, Paolo (1999[2] [1979–1988[1]]). *Dizionario etimologico della lingua italiana*, 5 vol. Bologne : Zanichelli.

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

DG = Hatzfeld, Adolphe et Darmesteter, Arsène, avec le concours de Thomas, Antoine (1890–1900). *Dictionnaire général de la langue française du commencement du XVII[e] siècle jusqu'à nos jours, précédé d'un traité de la formation de la langue*, 2 vol. Paris : Delagrave.

Diderot, Denis et D'Alembert, Jean le Rond (éd.) (1751–1772). *Encyclopédie ou Dictionnaire raisonné des sciences, des arts et des métiers par une société de gens de lettres*, 28 vol. Paris : Briasson/David/Le Breton/Durand (= ‹http://encyclopedie.uchicago.edu›).

Diez, Friedrich (1887[5] [1853[1]]). *Etymologisches Wörterbuch der romanischen Sprachen*. Bonn : Marcus.

Dubois, Jean, Mitterand, Henri et Dauzat, Albert (2001 [1938[1]]). *Grand dictionnaire étymologique & historique du français*. Paris : Larousse.

Dumont-Demaizière, Colette (éd. et trad.) (1973). *Charles de Bovelles, Sur les langues vulgaires et la variété de la langue française. Liber de differentia vulgarium linguarum et Gallici sermonis varietate (1533)*. Paris : Klincksieck.

Durkin, Philip (2009). *The Oxford Guide to Etymology*. Oxford : Oxford University Press.

Durkin, Philip (2014). *Borrowed Words. A History of Loanwords in English*. Oxford : Oxford University Press.

Felixberger, Josef (2003). Sub-, Ad- und Superstrate und ihre Wirkung auf die romanischen Sprachen : Galloromania. In G. Ernst, M.-D. Gleßgen, Chr. Schmitt et W. Schweickard (éd.) *Histoire linguistique de la Romania. Manuel international d'histoire linguistique de la Romania*. Berlin/New York : De Gruyter, vol. 1, 594-607.

FEW = Wartburg, Walther von *et al*. (1922–2002). *Französisches Etymologisches Wörterbuch. Eine darstellung des galloromanischen sprachschatzes*, 25 vol. Bonn/Heidelberg/Leipzig-Berlin/Bâle : Klopp/Winter/Teubner/Zbinden (= ‹https://apps.atilf.fr/lecteurFEW›).

Fouché, Pierre (1952–1961). *Phonétique historique du français*, 3 vol. Paris : Klincksieck.

Frantext = Montémont, Véronique (dir.) (2002–). *Frantext. Outil de consultation de ressources informatisées sur la langue française*. Nancy : ATILF, ‹http://www.frantext.fr›.

Fryba-Reber, Anne-Marguerite (2003). Étude et description étymologique et historique du lexique des langues romanes : le français et l'occitan. In G. Ernst, M.-D. Gleßgen, Chr. Schmitt et W. Schweickard (éd.) *Histoire linguistique de la Romania. Manuel international d'histoire linguistique de la Romania*. Berlin/New York : De Gruyter, vol. 1, 357-368.

Gamillscheg, Ernst (1969[2] [1928[1]]). *Etymologisches Wörterbuch der französischen Sprache*. Heidelberg : Winter.

Gebhardt, Karl (1974). *Das okzitanische Lehngut im Französischen*. Berne/Francfort : Lang.

Gleßgen, Martin-Dietrich (2008). Histoire interne du français (Europe) : lexique et formation des mots. In G. Ernst, M.-D. Gleßgen, Chr. Schmitt et W. Schweickard (éd.) *Histoire linguistique de la Romania. Manuel international d'histoire linguistique de la Romania*. Berlin/New York : De Gruyter, vol. 3, 2947-2974.

Gossen, Charles-Théodore (1964). Sur quelques correspondances entre l'argot français et le rotwelsch. In *Mélanges de linguistique romane et de philologie médiévale offerts à Maurice Delbouille*. Gembloux : Duculot, vol. 1, 257-270.

Gouvert, Xavier (2007). Le traitement étymologique des « francoprovençalismes » dans le *Trésor de la langue française*. Problèmes méthodologiques et étude de cas. In P. Rézeau (éd.) *Richesses du français et géographie linguistique*. Bruxelles : De Boeck & Larcier/Duculot, vol. 1, 361-413.

Greub, Yan (2014). Débat méthodologique. In É. Buchi et W. Schweickard (éd.) *Dictionnaire Étymologique Roman (DÉRom). Genèse, méthodes et résultats*. Berlin/Munich/Boston : De Gruyter, 269-288.

Greub, Yan (2017). [2. Épistémologie : histoire de la discipline, FEW, occitan.] Présentation. In É. Buchi, H. Carles, Y. Greub, P. Rézeau et A. Thibault (éd.) *Jean-Pierre Chambon, Méthodes de recherche en linguistique et en philologie romanes*. Strasbourg : ÉLiPhi, vol. 1, 177-180.

Guiraud, Pierre (1978[3] [1968[1]]). *Patois et dialectes français*. Paris : PUF.

Guiraud, Pierre (1982). *Dictionnaire des étymologies obscures*. Paris : Payot.

Höfler, Manfred (1982). *Dictionnarie des anglicismes*. Paris : Larousse.

Holtus, Günter (1990). Geschichte des Wortschatzes. In G. Holtus, M. Metzeltin et Chr. Schmitt (éd.) *Lexikon der Romanistischen Linguistik (LRL)*. Tübingen : Niemeyer, vol. 5/1, 519-529.

Hope, Thomas E. (1971). *Lexical Borrowing in the Romance Languages. A Critical Study of Italianisms in French and Gallicisms in Italian from 1100 to 1900*, 2 vol. Oxford : Blackwell.

Jänicke, Otto (1991). *Französische Etymologie. Einführung und Überblick*. Tübingen : Niemeyer.

Juilland, Alphonse, Brodin, Dorothy et Davidovitch, Catherine (1970). *Frequency dictionary of French words*. La Haye : Mouton.

Körting, Gustav (1907[3] [1891[1]]). *Lateinisch-romanisches Wörterbuch (Etymologisches Wörterbuch der romanischen Hauptsprachen)*. Paderborn : Schöningh.

Kremnitz, Georg (2007). La langue d'oc : une et plurielle. *Langue et cité. Bulletin de l'observatoire des pratiques linguistiques* 10, 7.

Lambert, Pierre-Yves (2003[2] [1994[1]]). *La langue gauloise. Description linguistique, commentaire d'inscriptions choisies*. Paris : Éditions Errance.

Leroy-Turcan, Isabelle (1991). *Introduction à l'étude du « Dictionnaire étymologique ou Origines de la langue française » de Gilles Ménage (1694). Les étymologies de Ménage : science et fantaisie*. Lyon : Université Lyon III Jean Moulin.

Maggiore, Marco et Buchi, Éva (2014). Le statut du latin écrit de l'Antiquité en étymologie héréditaire française et romane. In Fr. Neveu, P. Blumenthal, L. Hriba, A. Gerstenberg, J. Meinschaefer et S. Prévost (éd.) *Actes du Congrès Mondial de Linguistique Française 2014 (Berlin, 19-23 juillet 2014)*. Paris : Institut de Linguistique Française, 313-325, ‹http://dx.doi.org/10.1051/shsconf/20140801161›.

Marchello-Nizia, Christiane (2006). *Grammaticalisation et changement linguistique*. Bruxelles : De Boeck/Duculot.

Matsumura, Takeshi (2015). *Dictionnaire du français médiéval*. Paris : Les Belles Lettres.

Meier, Harri (1984). Anmerkungen zu Pierre Guirauds *Dictionnaire des étymologies obscures*. *Romanische Forschungen* 96, 53-77.

Mel'čuk, Igor A. (2012/2013). *Semantics. From meaning to text*, 2 vol., édités par David Beck et Alain Polguère. Amsterdam/Philadelphie : Benjamins.

Ménage, Gilles (1694[2] [1650[1]]). *Dictionnaire étymologique, ou Origines de la langue françoise*. Paris : Anisson.

Mertens, Bianca (2018). *Le suffixe */'ur-a/ : recherches sur la morphologie dérivationnelle du protoroman*. Thèse de doctorat. Université de Liège/Université de Lorraine.

Möhren, Frankwalt (2016). *Dictionnaire Étymologique de l'Ancien Français. Complément bibliographique 2016*. Berlin/Boston : De Gruyter.

Neveu, Franck (2004). *Dictionnaire des sciences du langage*. Paris : Armand Colin.

Nyrop, Kristoffer (1904–1930). *Grammaire historique de la langue française*, 6 vol. Copenhague : Gyldendal.

OED[1] = Murray, James A. H., Bradley, Henry, Craigie, William A. et Onions, Charles T. (dir.) (1884–1928). *A New English Dictionary on Historical Principles*, 9 vol. Oxford : Clarendon.

OED[2] = Simpson, John Andrew et Weiner, Edmund S. C. (dir.) (1989[2] [1884–1928[1]]). *The Oxford English Dictionary*, 20 vol. + 3 suppl. Oxford : Clarendon.

OED[3] = Simpson, John, puis Proffitt, Michael (dir.) (2000–[3] [1884–1928[1]]). *Oxford English Dictionary. The definitive record of the English language*. Oxford : Clarendon, ‹http://www.oed.com›.

Peeters, Bert (2015). French Semantic Primes, with English equivalents. Brisbane : Griffith University, ‹https://intranet.secure.griffith.edu.au/schools-departments/natural-semantic-metalanguage/downloads› (sous « French »).

Petit Robert 2007 = Rey-Debove, Josette et Rey, Alain (éd.) (2007). *Le nouveau petit Robert. Dictionnaire alphabétique et analogique de la langue française*. Paris : Le Robert.

Pfister, Max et Lupis, Antonio (2001). *Introduzione all'etimologia romanza*. Soveria Mannelli : Rubbettino.

Picoche, Jacqueline (2008 [1971[1]]). *Dictionnaire étymologique du français*. Paris : Le Robert.

Pitz, Martina (2006). Pour une mise à jour des notices historiques consacrées aux emprunts à l'ancien francique dans le *Trésor de la langue française informatisé*. In É. Buchi (éd.) *Actes du Séminaire de méthodologie en étymologie et histoire du lexique (Nancy/ATILF, année universitaire 2005/2006)*. Nancy : ATILF, ‹http://www.atilf.fr/IMG/pdf/seminaires/Seminaire_melh_Pitz_2006-04-05.pdf›.

Polguère, Alain (2016[3] [2003[1]]). *Lexicologie et sémantique lexicale. Notions fondamentales*. Montréal : Les Presses de l'Université de Montréal.

Popelár, Inge (1967). Die Etymologien Ménages im Lichte der modernen Wortforschung. *Beiträge zur romanischen Philologie* 6, 347-357.

Python, Fabien (2015). *La duplicité étymologique du lexique. Étude des doublets relevant du clivage héréditaire/savant en langue française*. Thèse de doctorat. Université de Fribourg/Université de Lorraine.

Queffélec, Ambroise (2008). Histoire interne du français (hors d'Europe) : lexique et formation des mots. In G. Ernst, M.-D. Gleßgen, Chr. Schmitt et W. Schweickard (éd.) *Histoire linguistique de la Romania. Manuel international d'histoire linguistique de la Romania*. Berlin/New York : De Gruyter, vol. 3, 2974-2982.

Renders, Pascale (2015). *L'informatisation du* Französisches Etymologisches Wörterbuch. *Modélisation d'un discours étymologique*. Strasbourg : ÉLiPhi.

REW = Meyer-Lübke, Wilhelm (1930–1935[3] [1911–1920[1]]). *Romanisches Etymologisches Wörterbuch*. Heidelberg : Winter.

Rey, Alain (dir.) (2016 [1992[1]]). *Dictionnaire historique de la langue française*. Paris : Le Robert.

Rickard, Peter (1989[2] [1974[1]]). *A History of the French Language*. Londres/New York : Routledge.

Roques, Gilles (1990). Étymologie et histoire du lexique. In G. Holtus, M. Metzeltin et Chr. Schmitt (éd.) *Lexikon der Romanistischen Linguistik (LRL)*. Tübingen : Niemeyer, vol. 5/1, 507-518.

Roques, Gilles (1993). Compte rendu d'Alain Rey (dir.), *Dictionnaire historique de la langue française*, Paris, 1992[1]. *Revue de linguistique romane* 57, 235-240.

Saussure, Ferdinand de (1972 [1906–1911]). *Cours de linguistique générale*, publié par Charles Bally et Albert Sechehaye avec la collaboration d'Albert Riedlinger, éd. Tullio De Mauro. Paris : Payot.

Scheler, Auguste (1888[3] [1862[1]]). *Dictionnaire d'étymologie française d'après les résultats de la science moderne*. Bruxelles : Muquardt.

Schlegel, August-Wilhelm (1823). De studio etymologico. *Indische Bibliothek* 1, 277-294.

Schweickard, Wolfgang (1992). *« Deonomastik ». Ableitungen auf der Basis von Eigennamen im Französischen (unter vergleichender Berücksichtigung des Italienischen, Rumänischen und Spanischen)*. Tübingen : Niemeyer.

Städtler, Thomas (2012). Von der Unmöglichkeit, ein Wörterbuch des Altfranzösischen zu schreiben. In St. Dörr et Th. Städtler (éd.) *Ki bien voldreit raisun entendre. Mélanges en l'honneur du 70e anniversaire de Frankwalt Möhren*. Strasbourg : ÉLiPhi, 247-258.

Stefenelli, Arnulf (1981). *Geschichte des französischen Kernwortschatzes*. Berlin : Schmidt.

Straka, Georges (1956). La dislocation linguistique de la Romania et la formation des langues romanes à la lumière de la chronologie relative des changements phonétiques. *Revue de linguistique romane* 20, 249-267.

Swiggers, Pierre (1990). Lumières épistolaires sur l'histoire du F.E.W. : lettres de Walther von Wartburg à Hugo Schuchardt. *Revue de linguistique romane* 54, 347-358.

Swiggers, Pierre (1997). *Histoire de la pensée linguistique. Analyse du langage et réflexion linguistique dans la culture occidentale, de l'Antiquité au XIXe siècle*. Paris : PUF.

Swiggers, Pierre (1998). L'occitan et les langues romanes. In G. Holtus, M. Metzeltin et Chr. Schmitt (éd.) *Lexikon der Romanistischen Linguistik (LRL)*. Tübingen : Niemeyer, vol. 7, 67-82.

Tadmor, Uri, Haspelmath, Martin et Taylor, Bradley (2010). Borrowability and the notion of basic vocabulary. *Diachronica* 27, 226-246.

Thibault, André (2009). Présentation. In A. Thibault (éd.) *Gallicismes et théorie de l'emprunt linguistique*. Paris : L'Harmattan, 11-16.

Thomassen, Helga (1997). Zur Behandlung der Datierungen im *Dictionnaire historique de la langue française*. *Zeitschrift für romanische Philologie* 113, 39-50.

Timberlake, Alan (1993). Russian. In B. Comrie et G. G. Corbett (éd.), *The Slavonic Languages*. Londres/New York : Routledge, 827-886.

TLF(i) = Imbs, Paul et Quemada, Bernard (éd.) (1971–1994). *Trésor de la langue française. Dictionnaire de la langue du XIX^e^ et du XX^e^ siècle (1789–1960)*, 16 vol. Paris : Éditions du CNRS/Gallimard (= ‹http://atilf.atilf.fr/tlf.htm›).

TLF-Étym = Steinfeld, Nadine (dir.) (2005–). *Trésor de la langue française étymologique*. Nancy : ATILF, ‹http://www.atilf.fr/tlf-etym›.

TLL = (1900–). *Thesaurus Linguae Latinae*. Leipzig/Stuttgart/Berlin/New York : Teubner/Saur/De Gruyter.

Trotter, David (2006). Contacts linguistiques intraromans : roman et français/occitan. In G. Ernst, M.-D. Gleßgen, Chr. Schmitt et W. Schweickard (éd.) *Histoire linguistique de la Romania*. *Manuel international d'histoire linguistique de la Romania*. Berlin/New York : De Gruyter, vol. 2, 1776-1785.

Walter, Henriette (1988). *Le français dans tous les sens*. Paris : Laffont.

Wartburg, Walther von (1958[5] [1934[1]]). *Évolution et structure de la langue française*. Berne : Francke.

Wartburg, Walther von (1953). Organisation et état actuel des travaux relatifs au Französisches Etymologisches Wörterbuch. In *Essais de philologie moderne. Communications présentées au Congrès international* ***de philologie moderne****, réuni à Liège du 10 au 13 septembre 1951*. Paris : Les Belles Lettres, 97-114.

Winter-Froemel, Esme (2015). Le français en contact avec d'autres langues. In Cl. Polzin-Haumann et W. Schweickard (éd.) *Manuel de linguistique française*. Berlin/Boston : De Gruyter, 401-431.

Winter-Froemel, Esme (2018). Ludicity in lexical innovation (I) – French. In S. Arndt-Lappe, A. Braun, C. Moulin et E. Winter-Froemel (éd.) *Expanding the Lexicon. Linguistic Innovation, Morphological Productivity, and Ludicity.* Berlin/Boston: De Gruyter, 229–259.

Wierzbicka, Anna (1992). Back to definitions : Cognition, semantics, and lexicography. *Lexicographica*. *International Annual for Lexicography* 8, 146-174.

Wolf, Heinz Jürgen (1991[2] [1979[1]]). *Französische Sprachgeschichte*. Heidelberg/Wiesbaden : Quelle & Meyer.

Zumthor, Paul (1958). Fr. *étymologie* (essai d'histoire sémantique). In *Etymologica. Walther von Wartburg zum siebzigsten Geburtstag (18. Mai 1958)*. Tübingen : Niemeyer, 873-893.