

André Thibault

Exposé présenté dans le cadre des séminaires de l'ATILF
CNRS-Nancy, vendredi le 7 septembre 2001

Résumé

La lexicographie de langue espagnole, qui accuse un fort retard face à celle du monde francophone, dispose depuis quelques années de nouveaux instruments de recherche qui permettront à ceux qui veulent bien s'en donner la peine d'améliorer substantiellement l'état de nos connaissances sur l'histoire du lexique castillan. De nombreux cédéroms mettent aujourd'hui à notre disposition des masses considérables de textes, littéraires ou non, interrogeables grâce à des logiciels de recherche plus ou moins performants; en outre, le site internet de la Real Academia Española met gratuitement à la disposition des chercheurs du monde entier deux banques de textes, l'une diachronique (des débuts de la langue à 1975 - CORDE) et l'autre synchronique (de 1975 à nos jours - CREA), ainsi que la totalité du texte des 21 éditions de son dictionnaire. Nous essaierons de présenter en temps réel le fonctionnement de ces nouveaux instruments informatiques, en illustrant la démarche à partir d'exemples tirés de notre futur "Dictionnaire des gallicismes de l'espagnol".

Plan de l'exposé

0. Introduction

1. Les banques de données textuelles pour l'étude de l'espagnol

- 1.1. ADMYTE
- 1.2. Alphonse X le Sage
- 1.3. CORDE
- 1.4. CREA
- 1.5. Teatro Español del Siglo de Oro
- 1.6. *Obras completas* de Menéndez Pelayo
- 1.7. Les dictionnaires de l'Académie Espagnole en ligne
- 1.8. Les dictionnaires sur CD-ROM
 - 1.8.1. Real Academia Española, *Diccionario de la lengua española* en CD-ROM multiplateforma. 21^a edición.
 - 1.8.2. María Moliner, *Diccionario de uso del español*. CD-ROM de la primera edición papel.

2. Etudes de cas

- 2.1. *(Sauce) béchamel*
- 2.2. *Marron*
- 2.3. *Aromathérapie*

0. Introduction

Les instruments informatiques dont je vais vous parler aujourd'hui sont tous relativement récents et ouvrent de nouvelles perspectives à la lexicologie historique de l'espagnol. Je m'y intéresse car j'ai entrepris un projet de recherche sur les emprunts de cette langue au français, projet qui devrait déboucher sur l'élaboration d'un dictionnaire historique, philologique et critique des gallicismes de l'espagnol. Historique, par l'importance qu'il compte accorder

aux premières attestations et aux canaux de transmission des mots ; philologique, car j'aimerais consacrer une place importante aux citations dans ce futur dictionnaire, ce qui est malheureusement encore assez inhabituel en lexicographie espagnole ; critique, car il s'agira également de revoir le traitement des gallicismes dans le discours scientifique (dictionnaires, mais aussi monographies, etc.), de façon à suggérer à l'Académie espagnole, par exemple, d'améliorer cet aspect de son dictionnaire (il se trouve qu'il en a bien besoin).

J'ai donc été amené à identifier et à utiliser un certain nombre d'outils informatiques qui permettent d'accélérer énormément le travail de recherche et de dépouillements, et dont on serait bien fou de se passer aujourd'hui, une grande part du boulot le plus ingrat ayant déjà été fait. Pour qu'il n'ait pas été fait en vain, lançons-nous donc dans l'exploitation de ces nombreuses richesses.

1. Les banques de données textuelles pour l'étude de l'espagnol

Les banques de données textuelles informatisées pour l'étude de l'espagnol se présentent tantôt sous forme de CD-ROM, tantôt en accès libre sur Internet. Leur exploitation a déjà commencé à porter ses fruits au sein de la communauté scientifique (je pense par exemple à un ouvrage de Joel Rini, *Exploring the Role of Morphology in the Evolution of Spanish*, dont j'ai rendu compte dans RLiR 66, 2002, pp. 270-276), ouvrage qui tire un grand profit des ressources d'ADMYTE et du CD-ROM d'Alphonse le Sage).

1.1. ADMYTE

Ce sigle signifie *Archivo Digital de Manuscritos y Textos Españoles* (*Archives numériques de manuscrits et textes espagnols*). Edité à Madrid par la compagnie Micronet. Il comporte deux volumes, numérotés 0 et 1, qui couvrent tout le moyen âge. Les plus grands monuments de la langue (comme le *Poema de Mio Cid*) y sont inclus, mais on y trouve un sage équilibre entre ouvrages littéraires et non-littéraires.

Une grande particularité de ces CD-ROMs est qu'ils présentent non seulement la transcription du manuscrit en mode texte*, ce qui permet bien sûr divers types de recherche avec opérateurs booléens, mais également le manuscrit original numérisé en mode image. N'ayant pas encore pu utiliser ce programme, je ne peux malheureusement pas vous dire si la résolution est assez bonne pour en permettre une véritable utilisation critique, mais si tel est le cas il s'agit alors vraiment d'un instrument formidable, qui permet de vérifier sur fac-similé le bien-fondé d'une transcription qui, pour une raison ou pour une autre, pourrait nous paraître douteuse. Le tout peut bien sûr être imprimé ou copié-collé.

* Un groupe d'experts d'Espagne et d'autres pays réalisent les transcriptions en suivant les normes du Hispanic Seminary of Medieval Studies de l'Université du Wisconsin à Madison, pour obtenir des fichiers en format ASCII.

1.2. Alphonse X le Sage

Une réalisation du célèbre « Hispanic Seminary of Medieval Studies » de l'Université du Wisconsin à Madison, *The Electronic Texts and Concordances of the Prose Works of Alfonso X, El Sabio* ont été élaborés par Lloyd Kasten, John Nitti et Wilhelmina Jonxis-Henkemans (1997). Source essentielle pour l'histoire du vocabulaire technique et scientifique naissant (*Sachprosa*) de l'espagnol du XIIIe siècle (Tolède 1221-Séville 1284).

1.3. CORDE

Nous abordons maintenant des corpus électroniques accessibles en ligne sur le site de la Real Academia Española (www.rae.es/).

CORDE est une banque de données textuelles de l'espagnol de toutes les époques et de toutes les régions de l'hispanophonie, de la naissance de la langue jusqu'en 1975, date à partir de laquelle CREA (v. ci-dessous) prend la relève. CORDE recueille un ensemble de textes sélectionnés et ordonnés selon des critères précis, et permet l'extraction d'informations lexicales et grammaticales. Le sigle correspond à *Corpus Diacrónico del Español* (corpus diachronique de l'espagnol).

L'entreprise a démarré en 1994, à une époque où l'Académie (la *Real Academia Española*) a décidé d'exploiter les nouvelles techniques informatiques pour constituer une banque de données qui puisse être consultée par tous les chercheurs, où qu'ils soient, sans nécessité de se déplacer à l'Académie. Elle compte actuellement environ 125 millions de mots-occurrences, et regroupe des textes appartenant à différents genres (narratifs, lyriques, dramatiques, scientifiques-techniques, historiques, juridiques, religieux, journalistiques, etc. Le but est de représenter toutes les variétés géographiques, historiques et génériques, afin que l'ensemble soit suffisamment représentatif.

Comme c'est aussi le cas pour CREA (v. ci-dessous), les ouvrages s'obtiennent de deux manières : en format électronique et en format papier. Dans le second cas, les ouvrages doivent être numérisés grâce à un programme de reconnaissance optique de caractères (OCR). Dans un cas comme dans l'autre, une équipe se charge de corriger les erreurs possibles et de codifier postérieurement les textes.

CORDE est régulièrement utilisé pour attester des mots et rendre compte de leur première apparition. Les lexicographes de la *Real Academia* répondent aux consultations internes et externes (tout comme ceux de l'ATILF avec FRANTEXT), lorsqu'il s'agit par exemple de savoir si un mot ou une expression ont été en usage dans tel pays et/ou à une telle époque. L'une des principales fonctions de ce corpus diachronique sera de fournir les matériaux de base pour la confection du *Diccionario histórico de la lengua española*, que prépare la *Real Academia*.

1.4. CREA

CREA est une banque de données textuelles d'espagnol contemporain, c'est-à-dire un ensemble de textes d'origines diverses numérisés et emmagasinés sur support informatique. Le sigle correspond à *Corpus de Referencia del Español Actual*. CREA a été lancé en 1994, conjointement avec CORDE, avec la collaboration du Secrétariat d'Etat aux Universités, à la Recherche et au Développement, dont le soutien financier a servi à subventionner l'équipe de linguistes de l'Institut de Lexicographie qui s'y consacrent.

CREA compte aujourd'hui au moins 125 millions de mots-occurrences, et sa taille est en augmentation croissante. Il est composé d'une grande variété de textes écrits et oraux, produits dans toute l'Hispanophonie, de 1975 à nos jours. Les textes écrits, qui proviennent autant de livres que de revues et de journaux, se rattachent à plus d'une centaine de thématiques. La langue parlée est représentée par des transcriptions d'enregistrements relevant d'une large palette de domaines : conversations, nouvelles, reportages, documentaires, interviews, magazines télé.

Ces matériaux ont reçu un ensemble de marques textuelles, établies selon des standards internationaux, qui permettront de multiples possibilités de récupération de l'information. En outre, les matériaux sont soumis à un travail de lemmatisation des formes, ce qui devrait rendre possible une interrogation beaucoup plus fine des richesses de la base.

CREA est actualisé en permanence. Les textes de presse sont reçus sous format électronique pour être ensuite soumis à un processus de standardisation. Les livres sont numérisés, tout comme pour CORDE, avec des programmes de reconnaissance optique des caractères ; en outre, une équipe se charge de transcrire les enregistrements inclus dans la partie orale du corpus. La mise à jour respecte la diversité des types de sources.

Tout comme à l'ATILF, les lexicographes et les linguistes qui travaillent à l'Institut de Lexicographie pour les divers dictionnaires académiques ont des stations de travail à leur table qui leur donnent directement accès à la base ; cela dit, à travers Internet, le monde entier a accès, à toute heure et en toute gratuité, aux mêmes richesses documentaires (ce qui n'est malheureusement pas le cas avec FRANTEXT). Il est permis de croire (ou à tout le moins, d'espérer) que la qualité des études de lexicologie espagnole va décupler dans les années à venir, grâce à ce formidable instrument.

1.5. Teatro Español del Siglo de Oro

Cette base de données regroupe environ 850 œuvres dramatiques en texte intégral de 16 auteurs du Siècle d'Or (16^e siècle et première moitié du 17^e siècle). Etant donné l'importance de cette époque dans l'histoire de l'espagnol, et le caractère particulier de la langue du théâtre, il s'agit d'une source très prometteuse pour les recherches lexicologiques.

1.6. *Obras completas* de Menéndez Pelayo

Marcelino Menéndez Pelayo (1856-1912), historien et érudit espagnol, grand polygraphe devant l'Eternel. Le CD-ROM réunit, en format texte et avec un moteur de recherche, plus de 60.000 pages de texte, répartis en deux grands sous-ensembles : les *Obras completas*, numérisées à partir de l'*Edición Nacional* (Madrid, Consejo Superior de Investigaciones Científicas, 67 volumes), et la correspondance (*Epistolario*, à partir de l'édition réalisée par la *Fundación Universitaria Española*, qui réunit 23 volumes).

1.7. Les dictionnaires de l'Académie Espagnole en ligne

La « Bibliothèque virtuelle de dictionnaires académiques » consiste en une base de données d'images numérisées, associée à un programme de gestion documentaire qui contient les lemmes répertoriés dans la totalité des éditions des dictionnaires de la *Real Academia Española*.

Sa fiabilité est entière, dans la mesure où elle nous permet de consulter, non pas une transcription du texte de ces dictionnaires, mais bien une photographie de chacune de leurs pages, sans intervention de quelque sorte que ce soit de la part des éditeurs. Cette qualité implique toutefois un grand défaut, à savoir l'impossibilité de faire des recherches en texte intégral, dans les définitions. Le projet est né des études faites pour l'informatisation des fonds lexicographiques de l'Académie, qui se présentent traditionnellement sous forme de fiches-papier. Ce fichier comporte deux grandes subdivisions : les fiches d'information

lexicale (il s'agit surtout de citations d'auteurs) et les fiches d'information lexicographique (résultant du dépouillement des dictionnaires). Pour le traitement des fiches lexicographiques, l'Académie a opté pour la création d'une base de données faites d'images numérisées, associée à un programme de gestion. Sur Internet, l'utilisateur a accès à la totalité des éditions des dictionnaires de l'Académie ; dans le commerce, on trouve en outre un double DVD (*Nuevo Tesoro Lexicográfico de la Lengua Española*) qui inclut de nombreux autres dictionnaires, de Nebrija jusqu'à la fin du XIXe siècle. Voici les éditions des dictionnaires de l'Académie qu'on peut consulter librement sur Internet, gratuitement et sans mot de passe :

Diccionario de Autoridades:

1.^a ed.

Tomo I (1726), 723 págs.

Tomo II (1729), 714 págs.

Tomo III (1732), 816 págs.

Tomo IV (1739), 696 págs.

Tomo V (1737), 656 págs.

Tomo VI (1739), 578 págs.

2.^a ed.

Tomo I (1770), 552 págs.

Diccionario de la Real Academia:

1.^a ed. (1780), 965 págs.

2.^a ed. (1783), 974 págs.

3.^a ed. (1791), 867 págs.

4.^a ed. (1803), 927 págs.

5.^a ed. (1817), 918 págs.

6.^a ed. (1822), 869 págs.

7.^a ed. (1832), 788 págs.

8.^a ed. (1837), 791 págs.

9.^a ed. (1843), 762 págs.

10.^a ed. (1852), 731 págs.

11.^a ed. (1869), 813 págs.

12.^a ed. (1884), 1118 págs.

13.^a ed. (1899), 1050 págs.

14.^a ed. (1914), 1073 págs.

15.^a ed. (1925), 1270 págs.

16.^a ed. (1936-39), 1326 págs.

17.^a ed. (1947), 1337 págs.

18.^a ed. (1956), 1366 págs.

19.^a ed. (1970), 1422 págs.

20.^a ed., (1984), 1416 págs.

21.^a ed., (1992), 1513 págs.

Diccionario Manual:

1.^a ed. (1927), 2011 págs.

2.^a ed. (1950), 1572 págs.

3.^a ed. (1983-85), 2319 págs.

4.^a ed. (1989), 1666 págs.

Diccionario Histórico 1933-36: 2141 págs.

Tomo I (1933), 1108 págs.

Tomo II (1936), 1033 págs.

1.8. Les dictionnaires sur CD-ROM

Il existe un certain nombre de dictionnaires espagnols disponibles sur CD-ROM. En ce qui concerne leur taille, ils ne représentent pas grand-chose en comparaison avec les banques de données textuelles que nous avons vues jusqu'ici, mais leurs logiciels de recherche raffinés permettent d'accéder à des informations très précises (comme, par exemple, la liste de tous les mots donnés dans la rubrique étymologique comme issus du français).

1.8.1. REAL ACADEMIA ESPAÑOLA, *Diccionario de la lengua española* en CD-ROM multiplataforma. 21^a edición. Madrid, Espasa-Calpe, 1998.

La version numérisée, avec de nombreuses possibilités de recherche en texte intégral, de la toute dernière édition du célèbre dictionnaire de la *Real Academia Española* (elle date de 1992, c'est la 21^e édition).

1.8.2. MARÍA MOLINER, DICCIONARIO DE USO DEL ESPAÑOL. CD-ROM. Versión 1.1 (De la 1^a ed. en papel).

Malheureusement, le CD-ROM du fameux dictionnaire de María Moliner a été fait à partir de la première édition papier (1966), alors que celle-ci est tout-à-fait dépassée depuis qu'est parue la toute nouvelle édition, celle de 1998, riche en néologismes souvent absents de la dernière édition du DRAE. On attend avec impatience la parution d'un CD-ROM de la seconde édition.

2. Etudes de cas

Pour illustrer concrètement le fonctionnement des bases de la Real Academia Española, nous allons mener quelques petites recherches sur des gallicismes relativement récents, et très mal représentés jusqu'à maintenant dans la lexicographie espagnole.

2.1. (*Sauce*) *béchamel*

Les termes de gastronomie constituent l'un des plus forts contingents de gallicismes dans la plupart des langues occidentales. Penchons-nous sur l'un d'entre eux, (*sauce*) *béchamel*. Voici d'abord ce que l'on peut apprendre sur ce mot dans la lexicographie française :

« 1742. Tiré de (Louis de) *Béchamel*, marquis de Nointel, nom d'un gourmet renommé de la fin du XVII^e s. » BlochWartburg (8^e éd.).

« **Étymol. et Hist.** 1735 art culin. *béchamelle* (*Cuis. mod.*, t. 4, p. 75 ds *Fr. mod.*, t. 23, p. 306 : Turbot à la **Béchamelle**); 1742 *béchamel* (**F. Marin**, *Suite des Dons de Comus*, 3, 100-101 ds **Quem.**). Du nom de Louis de *Béchamel* (1630-1703), marquis de Nointel, célèbre gourmet et maître d'hôtel de Louis XIV. » TLF.

On remarquera que deux orthographes co-existent depuis les plus anciennes attestations :

« L'orth. de *Ac.* est *béchamel* (cf. aussi **Besch**. 1845, *Lar. 19e*, **Litttré**, **Guérin** 1892, *DG*, *Pt*

Lar. 1906, **Rob.** et **Dub.**). **Land.** 1834 et *Nouv. Lar. ill.* écrivent *béchamelle*. **Besch.**, s.v. *béchamelle*, signale : "On écrit aussi *béchamel*." *Lar.* 19e fait, au sujet de la graph. du mot, la rem. suiv. : "Quelques lexicographes, entre autres Legoarant, critiquent fort cette forme masculine *béchamel* donnée à un mot qui, avec l'article *la* qui précède, a toutes les apparences du féminin, et ils voudraient qu'on écrivît *béchamelle*. La première forme orthographique prouve tout simplement que l'on tient à mettre ce mot en rapport avec son origine, ce qui s'explique d'autant mieux que cette manière d'orthographier n'offre rien d'illogique, puisque *une sauce à la Béchamel* signifie *une sauce à la manière du financier Béchamel*." *Lar.* 20e, *Lar. encyclop.* et **Quillet** 1965 admettent *béchamel* ou *béchamelle*. » TLF.

Cela dit, dans Frantext, la graphie *béchamelle* est très rare : une seule attestation, très récente : B. Blier, *Les valseuses*, 1972, p. 146 : « Qu'est-ce que ça voulait dire cette béchamelle? »

En ce qui concerne la syntagmatique, une recherche dans Frantext nous confirme que la plus ancienne attestation de *sauce béchamel* est bien celle qui se trouve dans l'article du TLF, à savoir un passage de G. Duhamel daté de 1945 : « Et lui aussi, le vieux Peuch, il a tout son menu sur son gilet. Mais, lui, ce n'est pas de l'œuf. Il n'en mange plus, à cause du foie. C'est de la sauce béchamel ou quelque autre saleté de cette espèce. » G. Duhamel, *Chronique des Pasquier, La Passion de Joseph Pasquier*, 1945, p. 21.

On trouve 21 exemples de *béchamel* dans Frantext, de 1869 (Flaubert) à 1995 (Bianciotti). En résumé, nous avons donc deux graphies, dont l'une, graphiquement féminisée, est en fait très rare dans l'usage malgré son ancienneté ; nous avons aussi deux lexies, l'une simple (*béchamel-le*), l'autre composée (*sauce béchamel-le*), elle aussi beaucoup plus rare et surtout d'apparition plus tardive. C'est sur ces bases que nous allons orienter nos recherches dans les bases textuelles espagnoles. Une recherche dans les dictionnaires les plus récents permet de relever les données suivantes :

(Petite parenthèse avant de commencer : aucune des formes suivantes ne se trouve dans le DCECH de Corominas, qui ne s'intéresse pas du tout à ce genre de mot, ce qui revient à dire qu'on ne peut obtenir aucune information – ou presque – de nature étymologique ou historique dans la lexicographie espagnole, et que tout le travail est à faire ; heureusement, on dispose des instruments nécessaires pour y arriver.)

Voyons d'abord la situation dans les deux dernières éditions du dictionnaire de l'Académie (1992₂₁, 2001₂₂) :

besamel (mot-vedette dans DRAE 1992 ; comm. historique : « Del fr. *béchamel*. »)

besamela (var. de *besamel* dans DRAE 1992)

bechamel (renvoi à *besamel* dans DRAE 1992)

bechamela (Ø DRAE 1992)

salsa besamel, bechamel, besamela, bechamela (Ø DRAE 1992)

Voyons maintenant la chronologie de ces formes dans les éditions successives du DRAE :

besamela (1931, v. DRAE Hist. 1936 ; DRAE Hist. 1936—DRAE 2001; « Del fr. *bechamel* [sic] » DRAE Hist. 1936)

besamel (DRAE Supl. 1970—DRAE 2001)

bechamel (DRAE Supl. 1970—DRAE 2001)

bechamela (Ø)

salsa besamel, bechamel, besamela, bechamela (Ø)

Si l'on fait maintenant une recherche dans CORDE, voici ce que l'on trouve :

besamel (2 att. ; 1948, 1967)

besamela (un seul ex, métaling., dans *Historia de la lengua española* de R. Lapesa, 1942)

bechamel (665 att. dans seulement 5 doc. : 1940, 1942, 1952, 1961, 1972)

bechamela (19 att. dans 3 doc. : 1820/1823, 1891/1894, 1961)

salsa besamel (Ø)

salsa bechamel (121 att. dans 2 doc. : 1940, 1972)

salsa besamela (Ø)

salsa bechamela (6 att. dans 2 doc. : 1891/1894, 1961)

Voici maintenant les résultats dans CREA :

besamel (1 att. : 1986)

besamela (1 att. : 1983, dans un contexte métalinguistique : « Si la bechamel o besamela, como se llamó un tiempo en España, es buena, se le añaden unas rodajas de trufa, auténtica, no de plástico, y ya me dirán luego qué piensan de ello. » Carlos Fisas, *Historias de la Historia*, 1983)

bechamel (135 att. dans 19 doc. : 1977–1999)

bechamela (Ø)

salsa besamel (Ø)

salsa bechamel (26 att. dans 9 doc. : 1983—1997)

salsa besamela (Ø)

salsa bechamela (Ø)

Il y a donc, parmi toutes ces données, un fait central qui ressort dans l'histoire de ce type lexical : la variante la plus fréquente au XIX^e siècle est *bechamela* (à rattacher à la var. fr. *béchamelle*), alors qu'au XX^e siècle la forme qui s'est définitivement imposée (d'ailleurs dans toute l'Hispanophonie) est *bechamel*, emprunt pratiquement dépourvu d'adaptation graphique (si ce n'est de l'absence d'accent aigu) et dont l'adaptation phonétique respecte les correspondances graphème-phonème de l'espagnol (*ch* [tʃ]). Si l'on confronte ces données à ce que l'on trouve dans les dictionnaires de l'Académie, on peut d'abord proposer une première attestation (1820/1823) beaucoup plus ancienne que le 1931 du Dict. Hist. de 1936 ; ensuite, on pourrait suggérer à la *Real Academia Española* de retenir la forme *bechamel* (canal écrit) comme mot-vedette, et non *besamel* (canal oral) qui s'avère en fait extrêmement rare. Ensuite, il faudrait ajouter à la micro-structure de l'article la lexie *salsa bechamel*, bien attestée dans CORDE tout comme dans CREA. Enfin, on remarquera que la plus ancienne attestation de la lexie composée, sous la forme *salsa bechamela* (1891/1894), est antérieure de plusieurs décennies à la première attestation de *sauce béchamel* en français (Duhamel, 1945), ce qui nous amène à remettre en cause son statut de gallicisme et à proposer, sur le mode conditionnel, l'hypothèse d'une création autochtone (très banale et s'inscrivant dans un paradigme bien fourni : cf. *salsa bearnesa, mayordoma, rosa, rubia, tártara, verde*, etc., v. DRAE 1992 s.v. *salsa* – où l'on pourrait ajouter, bien sûr, *salsa bechamel*). Cela dit, des recherches plus approfondies du côté français nous permettraient peut-être de trouver de nouvelles attestations de *sauce béchamel* au XIX^e siècle.

2.2. Marron.

Ce mot, qui désigne en français le fruit comestible du châtaignier cultivé, mais aussi la couleur typique de ce fruit, a été emprunté par la langue espagnole. On connaît aussi la lexie composée *marrons glacés*, qui désigne ce même fruit confit dans du sucre. Voyons d'abord comment ce mot est traité dans la lexicographie française (TLF), puis ensuite espagnole.

Étymol. et Hist. A. 1. 1526 «fruit du marronnier» (C. Gruget, *Les Diverses leçons de Pierre Messie*, p.888); 1640 loc. *tirer les Marrons du feu avec la patte du chat* (Oudin *Curiositez*); 2. 1718 *Marrons d'inde* (Ac.); 3. 1706 *couleur de maron* ([D. A. de Brueys], *L'Avocat Pathelin*, 1715, sc. 3, p.13); 1750 *Maron nom de couleur* (J. Hellot, *L'Art de la teinture des laines*, p.485). B. 1. 1680 coiffure (Rich.); 2. 1752 pyrotechnie (Trév.); 3. id. «jeton que les personnes chargées de faire une ronde devaient déposer dans une boîte pour marquer leur passage» (ibid.); 4. 1764 *maron roti* «sorte de limaçon de mer» (Valmont de Bomare, s.v. *limaçon*); 5. 1777 «noyau non calciné d'une pierre passée au four à chaux» (*Encyclop. Suppl.*); 1782 «grumeau dans la pâte à pain» (*Encyclop. méthod. Arts et métiers* t.1, p.281b). C. 1821 arg. «des coups, de la bagarre» emploi partitif (Ansiaume, *Arg. Bagne Brest*, fo 6 vo, 67: Prens toutes tes Baioffes [armes à feu], car il pourra bien y avoir du marron); 1881 «coup au visage» (Rigaud, *Dict. arg. mod.*, p.244). Empr. à l'ital. *marrone* «grosse châtaigne comestible» (dep. début XIV^e s., Cenne de La Chitarra ds *Batt.*; cf. lat. médiév. *marro*, *-onis*, 1176, doc. de Côme ds *Nov. gloss.*), prob. dér. d'un rad. prerom. *marr-* «pierre, rocher», att. de l'Italie au Portugal, particulièrement dans les Alpes et les Pyrénées (v. J. Hubschmid ds *Romanica Helvetica* t.41, pp.52-57). Le mot a prob. pénétré en fr. par la région lyonnaise (v. K. Baldinger ds *Mél. Gardette* (P.). 1966, p.61). Fréq. abs. littér.: 149. Bbg. Dauzat (A.). Notes lexicol. *Fr. mod.* 1954, t.22, pp.84-88. – Hope 1971, pp.209-210. – Mack. t.2 1939, p.201. – Quem. DDL t.17. – Terracini (B.). Problemi di etimologia preromana. Archivio glottologico italiano. 1954, t.39, pp.120-141.

Donc, le mot est attesté depuis 1526 pour désigner le fruit, et depuis 1706 dans *couleur de maron* ; depuis 1750, il est employé seul comme nom de couleur. La lexie *marrons glacés* étant absente du commentaire étymologique et historique, j'ai dû compléter avec une recherche dans FRANTEXT :

FLAUBERT.G / CORRESPONDANCE / 1845 page 90 / 1841 T 1 : « Le port étincellera de rouennais et de rouennaises avec leurs petits qu' on bourrera de marrons glacés, et dont on collera les entrailles avec du sucre de pomme. »

Maintenant, la recherche dans les dictionnaires de l'Académie espagnole nous donne ceci :

DRAE M1927, DRAE M 1950 : « **MARRÓN**. Adj. Galicismo por **castaño** de color de castaña. U. t. c. s. || Galicismo por castaña confitada. »

DRAE 1936, 1956 : Ø

DRAE 1970 : « **marrón**². Adj. Dicho de las telas y prendas de vestir, de color castaño. No se aplica al cabello de las personas ni al pelo de los animales. »

DRAE 1984 : « **marrón**². (Del fr. *marron*, y este del ital. y bizantino, castaña comestible, de color castaño.) adj. Dícese del color castaño, o de matices parecidos. No se aplica al cabello de las personas ni al pelo de los animales. U. t. c. s. m.

DRAE M 1984, 1989 : « **marrón**². Adj. Dicho de las telas y prendas de vestir, de color castaño. No se aplica al cabello de las personas ni al pelo de los animales. U. t. c. s. m. || m. Castaña confitada.

DRAE 1992 : « **marrón**². (Del fr. *marron*, castaña comestible, de color castaño.) adj. Dícese del color castaño, o de matices parecidos. U. t. c. s. m. || 2. Dícese de lo que tiene este color. No se aplica al cabello de las personas ni al pelo de los animales.

DRAE 2001: « **marrón**². (Del fr. *marron*, castaña comestible, de color castaño.) adj. Dicho de un color : Castaño, o de matices parecidos. U. t. c. s. m. || 2. De color **marrón**. || 3. m. coloq. Contratiempo u obligación ingrata. || **comerse**, o **tragarse**, alguien **el**, o **un**, ~. Frs. coloqs. Cargar con la culpa de un delito, se haya cometido o no. || 2. **apechugar** (|| cargar con alguna obligación ingrata).

Si l'on récapitule, la plus ancienne att. du mot dans les dict. de l'*Academia* est de 1927 pour l'adjectif de couleur, ainsi que pour désigner la châtaigne confite (mais le syntagme *marrón glacé* n'apparaît dans aucune édition).

Ne cherchez pas le mot (avec cette acception) dans le Corominas ; il ne s'y trouve pas. Encore une fois, ce genre de mot n'intéresse pas le grand étymologiste catalan. Regardons alors, encore une fois, ce que l'on peut tirer des bases de données pour dresser l'historique de cet emprunt.

CORDE : 181 att. dans 66 doc. Quelques att. sont à rejeter (il y a deux homonymes *marrón*, mais ils s'avèrent en fait très rares dans l'usage). L'att. la plus ancienne est de 1884/1885 (donc quelques dizaines d'années avant le 1927 académique) et représente l'adjectif de couleur ; comme substantif désignant un nom de couleur, la plus ancienne att. est de 1912 (*marrón obscurísimo*) ; cf. encore 1917 *no le sienta el marrón* ; quant à la lexie *marron glacé*, elle remonte à 1914 ; cf. encore *bombones de marrón glacé* (1918). On trouve même *color marrón glacé* dans une source datée 1929-1933.

franja marrón foncé 1884/1885 (en parlant d'une tunique)
tela de seda marrón 1905
marrón obscurísimo 1912 (le nom de la couleur)
marrón glacé 1914 (le fruit confit)

Si l'on veut relever la nature des objets qui peuvent être *marrons* en espagnol, il est facile d'en faire la liste ; à titre d'exemples :

pantalón de grandes cuadros en blanco y marrón 1916 ; *terciopelo marrón* 1922 ; *cara de color marrón* 1928 ; *pañó marrón* 1935 ; *seda marrón* 1935 ; *media marrón* 1935 ; *algodón marrón* 1935 ; *lana marrón* 1935 ; *papel marrón* 1935 ; *bufanda marrón* 1935 ; *ojos marrones* 1951 ; *yegua marrón* 1962 ; *perro marrón* 1972 ; *visón marrón oscuro* 1972 ; *brazos marrones* 1972 ; *piernas marrones* 1972 ; *piel marrón* 1972 ; etc.

On remarque que si, effectivement, les cheveux et le poil ne sont jamais *marrons*, en revanche les bras, les jambes et la peau peuvent l'être, et même le vison, les chiens et les juments.

Dans CREA, *marrón* est extrêmement fréquent (plusieurs centaines d'attestations, de la couleur et du fruit) ; on trouve en outre trois att. de *marron glacé* (1986, 1994, 1996 ; le pluriel quant à lui est attesté 6 fois, sous la forme non adaptée *marrons glacés*). Il conviendrait donc d'inclure cette lexie à la micro-structure de l'article (comme l'a fait, d'ailleurs, Manuel Seco dans son admirable *Diccionario del español actual*, 1999).

2.3. Aromathérapie (esp. *aromaterapia*).

Voyons pour terminer un mot dont on ne sait même pas, a priori, s'il s'agit bien d'un gallicisme. Il aurait pu être créé dans plusieurs langues. Il faut donc approfondir un peu la recherche du côté des premières attestations. La lexicographie espagnole ne relève le mot que dans des sources récentes de néologismes (deux ex. de 1993 dans le *Diccionario de voces de uso actual*, 1994 ; le mot figure également dans le *Diccionario de neologismos de la lengua española* de Larousse (paru en 1998). Dans CREA, on trouve 13 attestations dans 9 documents, de 1994 à 1999, en Espagne comme en Amérique.

En anglais, *aromatherapy* est aujourd'hui très fréquent (comme on peut le constater en faisant une recherche sur Internet), mais il doit être relativement récent car il est complètement absent du CD-ROM de l'*Encyclopaedia Britannica* (1999) ; quant à l'OED², il ne lui a pas consacré d'entrée à part, mais en fournit une seule attestation, à l'article *biorhythm*, att. datée de 1985. J'ai donc cherché dans le catalogue en ligne de la Bibliothèque du Congrès à Washington, qui passe pour être l'une des plus grandes au monde, et j'ai trouvé le mot dans le titre de 55 ouvrages, dont le plus ancien a été publié en 1984, et dont la grande majorité date de la seconde moitié des années 90. Pour l'instant, je n'ai donc pu reculer davantage la première attestation du mot en anglais.

OED²: **1985** *Observer* (Colour Suppl.) 21 Apr. 7/3, I have done yoga, reflexology, biorhythmics, aromatherapy and EST.

Catalogue du Congrès (<http://catalog.loc.gov/>) : Ryman, D., *The aromatherapy handbook : the secret healing power of essential oils*, London : Century Publ., 1984.

En français maintenant, une vérification dans le FEW (t. 25, p. 292b s.v. AROMA 2 b) nous apprend que le mot est incorporé à la lexicographie française depuis Larousse 1960. C'est déjà beaucoup plus ancien que le 1984 de l'anglais. Mais une recherche en texte intégral dans le CD-ROM de l'*Encyclopaedia Universalis* m'a permis d'améliorer substantiellement cette datation. J'y ai en effet relevé l'information suivante :

« En 1926, l'ingénieur chimiste R. M. Gatefossé publie *Les Antiseptiques essentiels* puis, en 1928, un livre intitulé *Aromathérapie* (éd. Girardot), attirant l'attention sur les pouvoirs insoupçonnés des essences aromatiques. » (s.v. *aromathérapie*).

Une vérification dans le catalogue de la Bibliothèque Nationale (<http://catalogue.bnf.fr>) donne 61 titres lorsqu'on cherche « aromathérapie » comme mot du titre, ce qui est déjà plus que la recherche dans le catalogue du Congrès à Washington ; en outre, les dates de publication sont mieux réparties dans le temps. L'ouvrage évoqué dans le CD-ROM de l'*Encyclopaedia Universalis* s'y trouve, mais dans une édition de 1937 ; en voici la notice :

Auteur(s) : Gattefossé, René Maurice
Titre(s) : R. M. Gatefossé,... Aromathérapie, les huiles essentielles, hormones végétales. Préface du Docteur Foveau de Courmelles,... [Texte imprimé]
Publication : Paris, Girardot 1937. In-16, 188 p. [2993]
Description matérielle : 187 p.
Autre(s) auteur(s) : Foveau de Courmelles, F.

Il faudrait donc quand même continuer de faire des recherches dans différentes bibliothèques pour confirmer la valeur du 1928 relevé dans l'*Encyclopaedia Universalis*.

Par acquis de conscience, j'ai aussi fait quelques vérifications pour l'allemand. Le catalogue de la Bibliothèque Universitaire de Bâle nous donne 15 titres avec *Aromatherapie*, dont les dates de publications s'échelonnent de 1981 à 1997. Quant à la Bibliothèque de l'Université de Freiburg-in-Breisgau, elle nous donne 6 titres avec *Aromatherapie*, dont 4 en allemand, de 1990 à 1996, et deux en français, de 1955 et 1964.

En conclusion, il semble donc assez clair que le mot a été créé en français, et ne s'est vraiment répandu dans les autres langues qu'à partir des années 80-90 du XX^e siècle. Des contextes comme celui-ci confirment d'ailleurs qu'encore aujourd'hui, le mot semble se diffuser à partir de la France :

« [título] *Aromaterapia* / Los aceites esenciales relajantes del Taller de Alquimia (3.710 pta./17 ml.) se pueden aplicar en inhalaciones, en hidroterapia o sauna y equilibran la polaridad entre el yin y el yang. Los aceites esenciales de la firma francesa Decleor se adaptan a los tratamientos faciales de los más exquisitos salones de belleza.» (*El País semanal*, 21 de febrero de 1999, p. 78).

Cela dit, lorsque l'on trouve une traduction en espagnol d'un livre écrit en anglais, comme c'est le cas ci-dessous, il faut bien admettre que l'anglais peut aussi être une voie de pénétration indépendante :

Berwick, Ann, 1953-
Aromaterapia / Ann Berwick ; traducción al Español por Alberto Rodríguez y Edgar Rojas.
ed. (*Holistic aromatherapy*. Spanish. St. Paul, Minn., U.S.A. : Llewellyn Español, 1999. xii,250 p. : ill. ; 23 cm.

Ce petit exemple illustre bien l'importance que la *quantité* d'informations disponibles sur un mot donné influence directement la *qualité* du traitement explicatif qu'on peut lui consacrer.