

©Linguistica Occitana 5 (setembre de 2007) [www.revistadoc.org]

© Miriam BRAS

Universitat de Tolosa, CLLE-ERSS UMR 5623 CNRS

myriam.bras@univ-tlse2.fr

© Joan THOMAS

Universitat de Tolosa lo Miralh

joanthomas@club-internet.fr

Diccionaris, *corpora*, e basas de donadas textualas

Lo quatren Obrador de Linguistica Occitana (a Montpelhièr lo 7 d'abrial de 2006) foguèt consacrat a la lexicografia occitana. I faguèrem una comunicacion que sa tòca n'èra pas dirèctament un projècte lexicografic particular mas que prepausava d'agachar las donadas linguisticas que los lexicografes utilizan (o an utilizadas) per bastir los diccionaris. L'ensem d'aquelas donadas se ditz lo 'còrpus de trabalh' dels lexicografes, o mai generalament lo 'còrpus'.

Dins aquel article, agacharem primièr qualques òbras lexicograficas occitanas (las de De Sauvages, Rochegude, Honnorat, Couzinié, Vayssier, Mistral, Alibert, Laux) e ensajarem de descriure lor còrpus de trabalh e de far lo bilanç dels metòdes de trabalh (seccion 1). Puèi nos virarem cap a d'òbras lexicograficas d'autres lengas que l'occitan. Causirem dos diccionaris de lenga : un diccionari anglés (lo *Collins Cobuild English Language Dictionary*) e un diccionari francés (lo *Trésor de la Langue Française Informatisé*) qu'an totes dos la particularitat d'èsser ligats a de basas de donadas textualas (seccion 2). Aquò nos menarà a pausar la question de la pertinença de la bastison d'una basa de donadas textualas que poiriá constituir lo còrpus d'un projècte lexicografic occitan aveniror (seccion 3).

1. Agach sus la lexicografia occitana

1.1 Filiacion de las òbras lexicograficas

Fins a aquesta debuta de sègle XXI, totas las òbras lexicograficas qu'avèm son de diccionaris bilingues. Son dins la tradicion dels primièrs trabalhs de lexicografia ont lo francés respond al latin. Lo primièr foguèt lo de Jean Doujat editat en 1638 a Tolosa. I aguèt puèi mai d'una reedicion ont los editors (Noulet, Visner, ...) an copat lo tèxte. Los autres se seguisson. I a una filiacion avoada e evidentada entre los autres diccionaris :

- De Sauvages, 1756, 1781 (5 edicions)
- Rochegude, 1819
- Honorat, 1846
- Couzinié, 1850
- Vayssier, 1879 (post-mortem, 4 edicions)
- Mistral, 1886
- Alibert, 1966 (7 edicions)
- Laux, 2000 (francés-occitan) que coneis doas edicions.

1.2 Metòde e còrpus de trabalh

DE SAUVAGES

Lo primièr diccionari de Boissier de la Croix de Sauvages es de 1756 mas utilizam lo de 1785.

Dictionnaire languedocien-français, contenant un recueil des principales fautes que commettent, dans la diction et dans la prononciation françaises, les Habitans des Provinces Méridionales, connues autrefois sous la dénomination générale de la Langue d'Oc. Ouvrage où l'on donne avec bien des termes de la Langue Romance, ou de l'ancien Languedocien, celle de beaucoup de noms propres, autrefois noms communs de l'ancien langage : et qui est enrichi de plusieurs de ses articles de Remarques critiques, historiques, grammaticales, et d'Observations de Physique et d'Histoire naturelle. Nouvelle édition, corrigée d'un grand nombre de fautes, augmentée d'environ dix mille articles, et en particulier d'une nombreuse Collection de Proverbes Languedociens et Provençaux. Par L.D.S. (Pierre-Augustin Boissier de La Croix de Sauvages), Nismes, Gaude Libraire, 1785.

Aqueste diccionari de mai de 17 400 dintradas es l'òbra d'un intellectual que participèt a la redaccion de mai d'un article de l'*Encyclopédie*. Estúdia tant la geologia coma la fisica o la sericultura. Son objectiu es clar, vòl participar a l'espandiment del francés e participar a la correccion de la lenga francesa ; mas pauc a pauc se pren de passion per la lenga occitana e o confessa :

Je m'applaudis d'avoir écrit sur la langue romane lorsque je vois mon labueur recherché par M. Court de Gibelin, ce savant qui est bien plus en état d'apprécier le mérite de notre idiome que nos franchimans littérateurs, ces ingrats qui méconnaissent leur langue maternelle, qui se donnent des airs de la dédaigner parce qu'ils n'en aperçoivent pas les beautés ou qu'ils n'ont pas les talents de las mettre en oeuvre. Letra al Nimesenc J.-F. Séguier citada per l'abat de Boves (1987).

L'étude de la langue romane m'absorbe et me charme. [...] Mon esprit et mon coeur sont remplis de ses beautés, mais mon amour est si délicat et si paisible que rien ne pouvait être plus bienfaisant. Oh! que cet amour me rend heureux. (Boves 1987).

De Sauvages pren la defensa de l'occitan e ne fa l'illustracion. D'en primièr es per el necessari de clarificar l'apelacion de la lenga e partís en guèrra contra la denominacion de "patois".

[...] différents jargons grossiers et rustiques que parle le bas-peuple dans les provinces ou la Capitale. C'est une langue dégénérée d'une langue plus parfaite. Il n'en est pas de même du gascon, ou du languedocien auquel on a donné la dénomination de patois par suite de l'espèce d'avilissement et d'oubli où il est tombé depuis environ un siècle[...] Il n'a manqué à ce prétendu patois pour devenir la langue dominante du royaume que de s'être trouvé dans les mêmes circonstances qui ont favorisé les progrès de la langue française¹.

Ailàs, çò que l'òm pòt regretar d'aquel diccionari quasi enciclopedic e d'una granda riquesa etnografica, es la pauretat dels articles. Pauques exemples e paucas etimologies, la

¹ Tirat de l'article "patois".

situacion definicionala demòra encara paura. La grafia es pas brica justificada, ni mai la redaccion dels articles o las causidas tipograficas.

ROCHEGUDE

Rochegude es nascut e mòrt a Albi (18-12-1741, 16-03-1834)². Contra-amiral, viatja per explorar las tèrras australas, se consacra a l'estudi dels trobadors (Rochegude 1819a). Dins *l'Essai d'un glossaire occitanien* (Rochegude 1919b), se prepausa d'estudiar la lenga dels trobadors al lum de sos davanciers³, es çò que nos apren primièr dins lo prefaci. Mas lo prefaci d'aquel diccionari que compta, a quicòm prèp, 13 300 dintradas es tanben interessant per la concepcion que Rochegude a de la lenga occitana —que nomena sempre: *occitanien* o *langue occitanienne*— e de la lenga en general, puèi de sas originas e de son istòria (Swiggers 2003).

Le langage de chaque peuple est plus ou moins mêlé d'expressions et de mots empruntés aux autres peuples, avec lesquels il entretient des relations. Toute nation cultivant les sciences a, pour ainsi dire, deux langues, l'une littéraire et l'autre vulgaire ; ou, si l'on veut, l'une à l'usage de la classe éclairée, l'autre à l'usage du peuple. (p. xiii).

Vaquí, tre aquela debuta del sègle XIX, pausada la question del contacte entre las lengas e de la diglossia.

L'interés del prefaci de Rochegude per l'istòria comparativista e per los estudis romanics de son temps es evident. Per Rochegude, l'occitan es pas la lenga maire de las lengas romanicas ; e sa demostracion arriba en contra a la posicion de Raynouard.

Le langage de chaque peuple est toujours plus ou moins mêlé d'expressions empruntées aux autres peuples, avec lesquels il entretient des relations ; par conséquent, aucune nation civilisée ne peut se flatter d'avoir conservé sa langue primitive dans toute sa pureté. S'il existe des langues mères, l'occitanienne n'est pas de ce nombre ; car, malgré l'altération de ses traits, sa physionomie trahit son origine, et son air de famille prouve assez évidemment qu'il est dérivé du latin. Comment cet étranger a-t-il pu chasser l'enfant de la maison ? (p. v-vi).

HONNORAT

L'objectiu del diccionari d'Honorat (1846) es plural. D'un costat vòl una descripcion la mai completa possible de la lenga. Honorat utiliza essencialament Achard, Aubert (parlar de Valençolles, ms.), Avril, Béronie, Castor, Doujat, las *Flors del gay Saber*, Garcin, Ménage, Pellas, Raynouard, Roquefort, Sauvages. Son idèa es de recampar totes los dialèctes ; çò que correspond a una amira "pan-occitanista". Ne serà çò meteís per Piat dins son diccionari *Français-Occitanien* (1893). Mas Honorat, a l'encòp, fa una nomenclatura de la lenga en sincronia tot en plaçant la lenga occitana dins un mitan pus grand, qu'es l'espaci de la Romania. E mai balha d'equivalents alemands. Lo diccionari es doblament extensiu ; i se

² Véser Masson 1941, Olier 1901.

³ Qual son sos davanciers? Al defòra de Raynouard, segurament que Rochegude fa allusion a de Sauvages, ne menciona mai d'un còp lo diccionari que pareguèt a Nimes en 1785.

tròba las variantas dialectalas, per cada radical, i trobam tanben l'etimologia e los derivats. E d'un autre costat, vòl permetre al monde d'aprene lo francés.

COUZINIE

Quand escriu son *Diccionari de la lenga Romano-Castresa*, Couzinié (1850) a entre las mans lo diccionari de De Sauvages. Mas coneis segurament los trabalhs de Rohegude e a accés a la bibliotèca de Noulet, doncas al diccionari de Doujat (1638).

VAYSSIER

Vayssier utiliza los diccionaris de De Sauvages, Honnorat, Couzinié, Béronie (1824). Mas a tanben de sorgas orales : los susvelhants del seminari ont es Vayssier son de correspondents que pòrtan pèira a son obratge. Utiliza tanben sas lectures. Lo diccionari (1879) coneis 4 edicions .

MISTRAL

A la nòstra coneissença, l'òbra lexicografica de Frederic Mistral, *Lou Tresor dóu Felibrige*, coneguèt 9 edicions :

- 1878-1886, Ramondet-Aubin.
- 1886, Roumanille-Champion.
- 1932, Delagrave.
- 1968, Bérenguié.
- 1979, Culture Provençale et Méridionale, Petit⁴.
- 1979, Édisud.
- 1979, Slatkine.
- 1983, Édisud.
- 2003, Princi Negre.

Lo trabalh de lexicografia de Mistral comença per una lectura critica del diccionari d'Honorat. Aquel trabalh es materializat per la creacion d'un *Supplément* a l'òbra d'Honorat. L'objectiu de Mistral es caracterizat pel títol qu'es seguit per una definicion del diccionari meteis :

Lou Tresor dóu Felibrige ou dictionnaire Provençal-Français embrassant les divers dialectes de la langue d'oc moderne et contenant
1° *Tous les mots usités dans le Midi de la France, avec leur signification française, les acceptions au propre et au figuré, les augmentatifs et diminutifs, et un grand nombre d'exemples et de citations d'auteurs ;*
2° *Les variétés dialectales et achaïques à côté de chaque mot, avec les similaires des diverses langues romanes ;*
3° *Les radicaux, les formes bas-latines et les étymologies ;*
4° *La synonymie de tous les mots dans leurs divers sens ;*
5° *Le tableau comparatif des verbes auxiliaires dans les principaux dialectes ;*
6° *Les paradigmes de beaucoup de verbes réguliers, la conjugaison des verbes irréguliers et les emplois grammaticaux de chaque vocable ;*

⁴ Notar que l'edicion de Marcel Petit (Raphèle-lès-Arles) es accessibla en fichièr pdf sul sit de la Bibliothèque Nationale de France : <http://gallica.bnf.fr/>.

- 7° *Les expressions techniques de l'agriculture, de la marine et de tous les arts et métiers ;*
 8° *Les termes populaires de l'histoire naturelle, avec leur traduction scientifique ;*
 9° *La nomenclature géographique des villes, villages, quartiers, rivières et montagnes du Midi, avec les diverses formes anciennes et modernes ;*
 10° *Les dénominations et sobriquets particuliers aux habitants de chaque localité ;*
 11° *Les noms propres historiques et les noms de famille méridionaux ;*
 12° *La collection complète des proverbes, dictons, énigmes, idiotismes, locutions et formules populaires ;*
 13° *Des explications sur les coutumes, usages, moeurs; institutions, traditions et croyances des provinces méridionales ;*
 14° *Des notions biographiques, bibliographiques et historiques sur la plupart des célébrités, des livres ou des faits appartenant au Midi.*
par Frédéric Mistral, 1878-1886, Vve Remondet-Aubin, Aix-en-Provence⁵.

Vòl èsser un diccionari pan-occitan «embrassant tous les dialectes de la langue d'oc moderne»; «et contenant tous les mots usités dans le Midi de la France». L'ambicion de Mistral sul plan de la simpla nomenclatura es granda ; mas los autres punts detalhats del contengut de l'òbra daissan tanben entreveire la dificultat del prètzfach e la volontat d'un "mèstre" de voler balhar un trabalh complet, exhaustiu, al qual i aurà pas cap de repròche de far. Aquela volontat exprimida dins la descripcion del *Tresor dóu Felibrige* revèrta la del diccionari d'Honorat. Ne tornam balhar aici sota lo titol, lo punt primièr de la descripcion interiora :

Dictionnaire provençal-français ou dictionnaire de la langue d'oc, ancienne et moderne, suivi d'un vocabulaire français-provençal, S.-J. Honorat, Repos, Digne, 1846.⁶

Contenant:

1° Tous les mots de ses différents dialectes que l'auteur a pu connaître (près de 107 202) ; leur prononciation figurée, leurs synonymes, leur équivalent italien, espagnol, catalan, allemand, etc., quand ils ont le même radical ; leurs définitions et leurs étymologies.

Es evident que Mistral se plaça dins la continuitat d'Honorat⁷ coma Louis Piat (1893) se plaça qualque temps aprèp dins la seguida de las piadas de Mistral. E dins aquela volontat de donar un diccionari que siague pan-occitan, que Mistral ne ven a prepausar una concepcion de la lenga qu'es ben lo sol a aver presentada d'un biais clar, volèm dire per de definicions,

⁵ Utilizam l'edicion de 1979 prefaciada per Jean-Claude Bouvier, Édisud, Aix-en-Provence.

⁶ *Le vocabulaire porte un titre particulier et une pagination à part. Honorat, médecin, botaniste et philologue (1783-1852), fut originaire d'Allos, ar. de Barcelonnette, mais exerça sa profession à Digne. En 1840, il fit déjà paraître un Projet d'un dictionnaire provençal-français, ou d'un dictionnaire de la langue d'oc ancienne et moderne, dans lequel il expose surtout les problèmes de l'orthographe du provençal et annonce une grammaire provençale qui n'a jamais paru. Son dictionnaire, ouvrage considérable pour son époque et exposant soigneusement la signification des mots, est malheureusement conçu sur un plan trop vaste ; mais la majeure partie des mots furent recueillis par l'auteur et sont donc propres aux parlers des vallées du Verdon et de la Bléone. (Wartburg 1944).*

⁷ Es pas brica question de "plagiat" mas de filiacion o d'eiretatge Honorat-Mistral. Honorat servís de basa de trabalh a Mistral, pas mai. Sus la question véser l'introduccion de J.C. Bouvier dins la reedicion de 1979 per Edisud.

d'articles dins lo Tresor dóu Felibrige. Véser las dintradas « oucitan », « oucitanio » ; « patés ».

Las sorgas del *Tresor dóu Felibrige*

Mantunas fonts van abeurar l'òbra de Mistral :

Ce que prouve la comparaison incontestablement, c'est que Honorat a servi de modèle à Mistral. Qui pourrait le nier? La démarche de Mistral, avec certains abus peut-être, est après tout semblable à celle de la plupart des lexicographes qui ont toujours besoin de prendre appui sur les œuvres de leurs prédécesseurs et ne peuvent pas innover sans cesse dans le choix des définitions ou des exemples. Prefaci a l'edicion de 1979 (Edisud) per J.-C. Bouvier, p. 15.

Plan solide cèrca de modèls e al delà del diccionari d'Honorat devíá èsser en possession de mantunes obratges⁸. Véser l'article « diciounàri » dins Mistral.

Mistral aviá doncas una font pro importanta —partent d'aquela bibliografia de dictionaris⁹— mas utiliza tanben la correspondéncia. Subretot un còp l'afar de Font-Segunha lançat, a de correspondents als quatre caires d'Occitània per li porgir de donadas. Coma o escriu Marcelle d'Herde-Heiliger (1998) podrà illustrar son diccionari de mai de 40 000 citacions tiradas de 805 autors, sens comptar lo nombre de provèrbis¹⁰.

La tresena sorga es la literatura. La celebritat de Mistral fa que recep segurament una part granda de tot çò que se publica, e mai se una part bèla de las citacions son tiradas de son òbra poetica e aiçò ilustra tanben l'ideologia mistralenca que se pòt legir dins la sintèsi que nos porgís Marcelle D'herde-Heiliger (1998)¹¹. Se nos en tenèm a D'Herde-Heiliger (1998) vesèm que las sorgas son posadas¹² :

- En Provença essencialament: 348 autors son citats contra 7 en Auvèrnha, 19 en Bearn, 2 en Catalonha, 24 en Dalfinat, 77 en Gasconha, 223 en Lengadòc, 13 en lemosin, 348 en Provença. L'autor pus sovent citat es Mistral —una mena d'auto-illustracion amb 1950 citacions— Los cinc primièrs autors citats son Mistral (1950), C. Brueys (1738), A. Crousillat (1063), P. Goudelin (1062), J. Roumanille (1024) ; çò que mòstra clarament lo pes provençau dins l'òbra de Mistral. Pr'auquò gascons e lengadocians son pas doblidats mas lo pes dels autres ven pus leugièr quand òm s'aluènha de Provença, del ròdol rodanenc. Los autors dalfinencs, lemosins e auvernhasques son plan mal representats. Çò que fa del *Tresor dóu Felibrige* una òbra essencialament provençala, rodanenca, que la part de representacion dels autres parlars per l'exemple es plan mediòcra. Dins l'inconscient mistralenc lo dialècte d'Auvèrnha o de

⁸ Sabèm pas quines obratges a utilizat mas l'òm pòt ne coneisser una tièra en vesent ont a posat los exemples de sos articles. Un estudi menimós de la redaccion de las definicions permetriá tanben de se sarrar de la realitat bibliografica utilizada per Mistral.

⁹ Notam que doas òbras concernisson lo parlar del Forez, çò que pòt enganar lo lector se de dintradas forezianas son pas localizadas dins l'article.

¹⁰ Podèm considerar lo *TdF* coma una de las fonts paremiologicas mai importantas de Provença.

¹¹ Remarcam çà que la que l'autor dessepara Bearn de Gasconha, çò que nos pareis èsser una concepcion particulara de Gasconha ; puèi balha A. Galhard a la Gasconha —benlèu per çò que moriguèt a Pau ?— alara qu'es un escrivan lengadocian. Es çò meteís per Josèp Daubian-Delisle qu'es un autor Castrés coma Bru (Thomas et. al. 1982). Aquelis autors apareisson dins l'indèx «tri sur la localisation géographique» dins la part gascona mas a l'indèx general lor origina vertadièra es mantenguda.

¹² Las resultas chifradas de Marcelle d'Herde-Heiliger son a acceptar amb aproximacion.

Lemosin a pas tant de plaça —es pas per una manca de documentacion mas per una volontat de politica lingüistica— es per aquò que cap de dialòg se durbís pas quand interven un Josèp Roux, per exemple ; que jamai son trabalh de normativizacion serà pas discutit per Mistral. I a la Lei, lo *Tresor*, e i a lo cap d'òbra, *Mirèio*. Pas de plaça per lo de mai.

Lo quatren biais d'acampar de material lexicografic èra per Mistral d'anar sul terrenh per trapar los bons informators e collectar de faches linguistics (Prefaci a l'edicion de 1979 d'Edisud per J.-C. Bouvier, p. 31).

Resumiguèt totas sas intencions dins un sonet datat de Malhana, lo 7 d'octòbre de l'an 1878 que s'acaba aital:

*"O pople dòu miejour, escouto moun arengo:
Se vos recounquista l'empèri de ta lengo,
Per t'arnesca de nou, pesco en aqueu Tresor."*

ALIBERT

Lo diccionari d'Alibèrt (1966), correntament sonat "l'Alibèrt", es lo diccionari de referéncia d'una part dels renaissentistas, occitanistas. A aquel nivèl, e en defòra de sa valor intrinsèca, es estat sacralizat. Respond a de questions precisas : «Es dins l'Alibèrt o i es pas?», «Dins l'Alibèrt es escrit aital!».

Quand moríguèt Loís Alibèrt, lo diccionari èra manuscrit. Es sa bèla-sòrre que lo dactilografiet. Robèrt Lafont, al nom de l'I.E.O. se carguèt puèi de lo far publicar. Mas calguèt tornar passar lo primièr tractament del manuscrit. Lo trabalh foguèt donat a Ramond Chabbert e Pèire Bèc. Lo trabalh durèt fins a 1966, data de la primièra edicion. Lo diccionari a conegut fins ara 7 edicions.

Es estat bastit a la seguida del *Tresor dóu Felibrige* de Mistral mas tamben per la consultacion e l'utilizacion d'autres diccionaris. Es pas de bon mostrar a quinas sorgas a posat, mas es segur qu'utilizèt Mistral, Honnorat, Couzinié, Vayssier, Azaïs e segurament un grand nombre d'òbras literàrias (Chambon 2000a).

Es dins la revista OC qu'Alibèrt balha l'orientacion de son trabalh¹³. Dins un de sos articles sus l'epuracion de la lenga ditz :

Al mieu vejaire, la reforma linguistica [...] deu èsser la preocupacion majora dels nòstres escrivans. Aquesta se resumís dins los tres punts següents.

1) Seleccion del vocabulari e de la morfologia per l'adopcion de las formas melhor conservadas e las mai en acòrdi amb l'evolucion de nòstra lenga.

2) Enriquiment per manlèu de las paraulas que nos fan sofraita als autres parlars o dialèctes occitans e per l'adaptacion del vocabulari sabent greco-latin a las (sic) en acòrdi amb l'evolucion de nòstra lenga.

3) Epuracion dels mots venguts del francés o contaminat jos son influéncia e lor remplaçament pels tèrmes occitans encara subsistents dins los autres parlars o dialèctes.

Es dins aquel esperit de reviudar l'occitan e dins l'urgéncia d'un material que fa sofracha per l'ensenhament, que se fa lo diccionari d'Alibèrt. Lo diccionari fonciona per familha de "mot-

¹³ OC 188 (1953) : 35-38.

dintrada" coma o veirem amb lo cas de "taula". Per cada dintrada Alibèrt balha una glòsa gramaticala, puèi venon siá un equivalent, siá una definicion — mas la definicion impausa lo passatge per un dictionari del francés.

La particularitat del dictionari d'Alibèrt es lo classament de la nomenclatura que fa paréisser las familhas de mots : la basa servis de 'vedeta', sos equivalents franceses son seguits dels derivats, puèi dels compausats. Per exemple :

taula *f.* Table, planche, ais, banc d'étalage ; planche à porter le pain ; ciseau de maçon ; planche : plate-bande de jardin ; compartiment de saline ; jante de roue. *Dér.* **taular**, être à table ; mettre la nappe ; verser (*voiture*). *Syn.* *ataular*, **taulada**, convives, contenu d'une table ; lit de pierre d'une carrière ; planche de jardinage ; couche. **taulador**, planche pour le pain ; table où l'on fait le pain. **taulatge**, étalage ; droit d'étalage. **taulassa**, grande table. **taulat**, plancher ; jante de roue. **taulejar**, rester longtemps à table ; banqueter. **taulejada**, banquet. **taulejaire**, celui qui aime les plaisirs de la table, qui reste longtemps à table ; convive. **tauleta**, petite table ; table de chevet ; petite armoire à deux battants, *Cév.* **taulièr**, comptoir d'un marchand ; établi d'un artisan ; étal de boucherie ; métier de tisserand. **taulats**, pièce de bois qu'on met au haut d'un cuvier pour en augmenter la contenance. *Ag. Syn.* *ariscle*. **tautilhon**, petit établi, petit carré de jardin. *Éty.* *L. tabula*. *Comp.* **taula dobla**, table pliante.

LAUX

Christian Laux bastís son dictionari francés-occitan (1997) d'un autre biais. Se ten a posita los dictionaris ja citats, trabalha mai que mai amb l'esqueleta del *Diccionari catalan* de Castellanos i Llorenç ; mas sos equivalents son tirats de sas lectures de Cordas, Roqueta, e de fòrces autors del sègle XIX, de sa coneissença personala de la lenga d'Albigès e de Lunhan (Cessenon), e de sos correspondents (Rapin, Lagarda, Carbona, ...). Aicí un exemple d'article:

table *nf* **1** taula | table d'orientation taula d'orientacion | table de chevet, table de nuit tauleta de nuèch / nuèit | table ronde taula redonda | mettre, dresser la table metre la taula | **2** (*pour récapituler*) taula, indèx (*m*), ensenhador (*m*) | **3** loc div faire table rase faire taula rasa | se mettre à table fig/fam voidar lo sac, escutlar.

Conclusion

Aquel estat del luòc mòstra que las donadas que constituïsson lo còrpus de trabalh d'aqueles lexicografes son plan diversas : donadas obtengudas pel metòde de l'introspeccion, donadas lexicograficas anterioras, donadas eissidas d'enquistas, recampadas a partir d'informators, donadas eissidas de tèxtes, de correspondéncias o de conversacions transcrichas. L'estat del luòc fa tanben aparéisser una filiacion dins lo trabalh lexicografic. Dins aquel rapòrt entre los diferents obratges, legissèm a l'encòp de complementaritats al nivèl del còrpus e de mancas al nivèl del còrpus e de la metodologia.

Prepausam de sortir del domèni de la lexicografia occitana per descobrir d'autres òbras lexicograficas e confrontar los biais de bastir lo còrpus de trabalh.

2. Agach cap a d'òbras lexicograficas dins d'autres lengas

Prepausam de nos virar cap a d'òbras lexicograficas d'autres lengas que l'occitan. Anam tanben quitar lo paradigme lexicografic del *diccionari bilingue* per agachar de *diccionaris de lenga*, que son *monolingües* (la metalenga e la lenga objècte i son las meteissas), perque nos sembla qu'es un tal diccionari que manca a l'occitan. Avèm causit un diccionari anglés, lo *Collins Cobuild English Language Dictionary*, e un diccionari francés, lo *Trésor de la Langue Française*.

2.1 Lo diccionari de lenga Collins Cobuild English Language Dictionary

Lo diccionari

Lo diccionari de lenga *Collins Cobuild English Language Dictionary* es un diccionari de referéncia pels aprepreneires. Sa tòca es de lor far descobrir la lenga 'vertadièra' : *Helping learners with real English*. Aquò es ben plan la cèrca d'aquela lenga de la vida vidanta contemporanèa que faguèt espelir una còla de cercaires de l'Universitat de Birmingham e de lexicografes de l'ostal d'edicion Collins per bastir un còrpus que revèrte aquel estat de la lenga. Aquele projècte, amodat en 1980 pel Professor John Sinclair de l'Universitat de Birmingham, es descrich dins [Sinclair 1987](#) e [Renouf 1994](#). Se debanèt sus una temporada de 7 ans. Aquí çò que ne ditz Russon Wooldridge dins son analisi dels ligams entre lexicografia e informatica (Wooldridge 2004) :

The COBUILD (Collins and Birmingham University International Language Database) Project started in 1980, with the intention of creating a corpus of contemporary English for the writing of an entirely new dictionary and grammar. The young discipline of corpus linguistics and the COBUILD Project fed off each other in this innovative lexicographical environment.

Lo còrpus : *The Bank of English*

Lo còrpus bastit per la còla COBUILD se ditz *The Bank of English*. Es una colleccion grandarassa de tèxtes escriches e orals de totas menas que venon de fonts variadas : tèxtes escriches eissits de libres (romans, nòvas, fictionals o non), de manuals, de jornals, de revistas o de sits de la telaranha ; mas tanben de tèxtes orals transcriches : emissions de television, de ràdio, confèrencias, entrevistas, convèrsas, etc. Revèrta la variacion diatopica de la lenga anglesa, al mens al nivèl macroscopic, que los tèxtes venon de fonts britanicas, americanas, australianas e canadianas. Conten ara mai de 500 milions de mots e lo meton a jorn regularament dins l'amira de gardar un « imatge fisèl » de la lenga.

Un còp bastit, lo còrpus es estat analizat. Per aquò far, la còla COBUILD farguèt d'aisinas computacionalas per tractar los tèxtes del còrpus e melhor manipular las donadas linguisticas que clau. aital an lematizat e categorizat lo còrpus, es a dire qu'an etiquetat automaticament cada ocurrencia de mot amb sa categoria gramaticala (nom, vèrbe, adjectiu, etc.) e sa lèma (forma « rasic » neutra del ponch de vista de la flexion). An tanben fach virar un analisador sintactic sus gaireben la mitat del còrpus. Tot aqueles tractaments permeton d'interrogar la basa sus de critèris lexicals, morfosintactics e sintactics : traire de concordanças d'una forma, cercar d'esquèmas de combinasons de mots, de frequéncias de mots, recampar d'atestacions

de totes los usatges d'un mot. Puèi, en se basant sus l'analisi d'aquelas donadas trachas del còrpus, an bastit lo diccionari, amb, per cada lema causit coma entrada del diccionari son (sos) sens lexical(s), los environatges sintactics possibles (amb quina(s) preposicion(s) tal vèrbe o nom s'emplega per exemple), las restriccions semanticas suls arguments (amb quins tipès de subjèctes e d'objèctes tal vèrbe s'emplega per exemple), las constrenchas pragmaticas, las locucions bastidas sus la basa del mot descrich, etc. Enfin, cal apondre que lo còrpus *The Bank of English* a permès d'enriqueir lo diccionari d'exemples atestats e actuals, mai comprenedijs pels aprenejaires que los exemples de la literatura (cf. *Le Petit Robert* per exemple).

Cal senhalar tanben que *The Bank of English* a permès de produire de gramaticas e d'autres diccionaris. E mai es ara a la disposicion dels cercaires en linguistica, dels ensenhaires de lenga, dels traductors e dels estudiants. Coneis doncas una utilizacion en defòra de la lexicografia. Es tanben arremosada amb de *corpora* franceses, alemands e espanhòls dins una basa multilingua que se ditz lo *Collins Word Web*.

Linguistica de còrpus e lexicografia computacionala

Podèm retèner, per l'amira de la lexicografia occitana avenidoira, que lo projècte COBUILD de creacion e d'analisi d'un còrpus a permès la produccion d'un diccionari de lenga novèl pels aprenejaires de la lenga. Nos sembla qu'un diccionari d'aquela mena fa besonh per l'occitan qu'una grand part dels utilizaires de diccionaris occitans son d'aprenejaires, començaires o avançats. Lo trabalh de recèrca menat dempuèi mai de 20 ans a l'Universitat de Birmingham a l'entorn de COBUILD e tanben a l'Universitat d'Oxford amb *lo New Oxford English Dictionary Project* mòstra que d'escampilhs de granda escala son necites per l'analisi linguistica e la produccion lexicografica. Aqueles projèctes an marcat un doble cambiament : una avançada en linguistica, amb la naissença de la 'linguistica de còrpus', un novèl biais d'analisar las lengas que s'apièja sul tractament automatic de donadas textualas recampadas dins de *corpora* ; e una avançada en lexicografia, amb las novèlas aisinas informaticas d'ajuda al trabalh lexicografic.

2.2 Lo Trésor de la Langue Française

Lo TLF

Lo *Tresor de la Langue Française*, ara endavant lo *TLF*, es un diccionari de referéncia de la lenga francesa dels sègles XIX e XX. Foguèt realizat per un laboratòri del CNRS, l'INaLF, jos la direccion del Rector Paul Imbs, sus una temporada d'un quarantenat d'ans gaireben (lo projècte comencèt en 1957, lo darrièr del 16 volums foguèt publicat en 1994). La tòca èra de fargar una òbra nòva dins la posteritat del *Littré* e de donar a la lenga francesa una òbra lexicografica comparabla al *New Oxford English Dictionary*. Un dels aspèctes mai novadors del projècte foguèt la bastison d'un fons textual informatizat, aviada per Imbs entre 1960 e 1964, per ne tirar totes los exemples del *TLF*. Aquela basa textuala recampava de tèxtes literaris e scientifics dels sègles XIX e XX.

Lo TLFi

Lo *TLF* foguèt puèi informatizat entre 1991 e 2001 per donar lo *TLFi* (Dendien et Pierrel 2003). Coma dins lo *TLF*, cadun dels 100 000 mots i es descrich d'un biais complet : definicion, citacions per ilustrar las acepcions, prononciacion, ortografia, etimologia, istòria, estadisticas de frequéncia, bibliografia. Mas evidentament la cèrca d'informacions dins lo diccionari es grandament facilitada e de possibilitats novèlas son ofèrtas. Se pòt legir lo *TLFi* a mai d'un nivèl : legir los articles coma dins lo diccionari papièr, legir los articles en cercant un element particular (definicion o citacion per exemple), legir un article particular, dintrar dins una cèrca d'articles a partir d'un critèri de cèrca : per exemple totes los mots manlevats a l'occitan (al provençal pels lexicografes del *TLFi*), far de requèstas mai complèxas.

Frantext

Mas la tòca nòstra es pas de descriure lo *TLFi*, volèm puslèu tornar a la basa de donadas textualas ligada al *TLFi*. Coma o venèm de dire, bastiguèron una primièra basa textuala per ne tirar los exemples del *TLF* abans son informatizacion. Puèi, un còp lo *TLF* publicat, la basa venguèt la basa textuala *Frantext* qu'existís d'un biais autonòm coma aisina de trabalh pels cercaires e ensenaires en linguistica e en literatura. Recampa 5000 òbras dels sègles XVI a XX : gaireben 4000 tèxtes literaris e scientifics, e un milierat de tèxtes tecnicos per 210 milions de mots en tot. Lo *TLFi* e la basa *Frantext* son manejats pel meteis motor de cèrca e l'ensem forma una granda basa lexicala e textuala, cada basa es ligada a l'autra per de ligams dinamics que permeton per exemple de cercar totes las occurréncias d'un mot dins *Frantext* amb lor contèxte (concordàncias). *Frantext* es ara manteguda per l'ATILF e consultabla sus abonament (pas car, las universitats francesas son abonadas), l'integralitat dels tèxtes liures de dreches son accessibles sul siti del CNRTL.

Se l'integralitat de la basa textuala a constituït lo còrpus del *TLF*, las causas son diferentas per la basa textuala *Frantext* utilizada d'un biais autonòm. Cada cercaire i pòt causir son còrpus de trabalh, es a dire seleccionar tot o partida de la basa. La seleccion se pòt far segon mai d'un critèri : causir un autor o maites, una òbra o maites, un genre textual o maites, una temporada, e se pòt combinar los critèris.

Un còp lo còrpus de trabalh establït, se pòdon far de requèstas suls tèxtes causits. Çò pus simple per començar son las requèstas simplas coma la cèrca d'una grafia (valent a dire una forma grafica), d'un sintagma, d'una frasa, de mai d'una grafia en co-ocurréncia. aital se pòt cercar per exemple la locucion adverbiala *un peu plus tard* en cap de frasa, e recampar totes los extraches que clauson aquelas occurréncias. La basa *Frantext* es lematizada, doncas se pòt cercar totes las formas flexionalas d'un vèrbe, d'una locucion verbala, d'un nom o d'un adjectiu, se pòt tanben cercar totes los derivats d'un mot. Per donar un autre exemple se pòt cercar una frasa al passat simple que comença per '*puis*' e que seguís una altra frasa al passat simple. D'autres menas de requèstas son possiblas a partir de la bastison de tièras de mots : per exemple de bastir una tièra de noms que recampa totes los noms d'un camp lexical particular, permet puèi de cercar totes los contèxtes que clauson aqueles noms. D'autres requèstas permeton de far de causidas multiplas en un còp, per exemple « l'année (dernière|prochaine|précédente|suivante) » cercarà totes las occurréncias de *l'année dernière*, *l'année prochaine*, *l'année précédente*, *l'année suivante*. En fin, d'autres requèstas encara mai poderosas son possiblas sonque sus la partida categorizada de la basa (gaireben la mitat). Aquesta partida conten los tèxtes mai recents de *Frantext* (a partir de 1830) que son escriches amb l'ortografia francesa modèrna e que son estats etiquetats per un categorizador gramatical. Es possible de far virar sus aquela partida de la basa de gramaticas formals que permeton de

fargar de règlas parametrablas. aital se pòt cercar per exemple totes los vèrbes que començan una frasa o mai generalament estudiar l'òrdre dels mots dins la frasa.

Enfin, es aisit, sus l'integralitat de la basa de calcular de frequéncias d'ocurréncias e lor reparticion sul còrpus causit.

Podèm pas anar mai luènh aici sus la descripcion de las capacitats del motor de cerca de Frantext, mas podèm far la tièra de las aplicacions possiblas sus un còrpus : tirar de joslexics, tirar de concordàncias, de co-ocurréncias e de col·locacions, recampar las donadas per menar d'estudis morfològics, sintactics, semantics, estilistics, cercar d'etimologias, e fin finala, bastir un diccionari....

2.3 Conclusion

Ara que coneissèm melhor aqueles dos projèctes lexicografics universitaris, totes dos ligats a la bastison d'una basa de donadas textuala grandarassa : *The Bank of English* per la còla COBUILD e *Frantext* pel *TLFi*, nos sembla que lo camin de seguir es plan traçat. Se volèm amodar un projècte lexicografic seriós per l'occitan, poirem pas estalviar la bastison d'una basa de donadas textualas.

3. Cap a una basa de donadas textuala occitana

Aquela basa de donadas textualas occitanas poiria constituir una partida del còrpus d'un projècte lexicografic occitan avenidor. L'autra partida del còrpus seria constituïda de las donadas lexicograficas occitanas amassadas fins ara e qu'avèm descrichas dins la seccion 1. Calrà informatizar aquelas donadas lexicograficas, dins la dralha del trabalh menat pel GIDILOC pel diccionari d'Alibèrt. Mas nos anam concentrar aici sus la basa textuala que permetrà de recampar de matèria viva, e que deuria clare de tèxtes escriches de totes menas (tèxtes literaris, tecnicos, jornalistics, etc.), tanplan coma de tèxtes orals (entrevistats, emissions de radio, racontes, etc.).

La temporada de cobrir es en prioritat la de l'occitan contemporanèu (sègles XIX, XX e XXI), se volèm bastir un diccionari d'occitan contemporanèu. L'epòca modèrna (XVI-XVII) deurà èsser presa en compte tanben. Per las òbras anterioras, existís ja una basa textuala dins la *Concordance de l'Occitan Médiéval*, menat pel Professor Ricketts, que recampa la totalitat dels tèxtes en lenga occitana dempuèi lo primièr tèxte atestat fins als tèxtes de la fin del sègle XV. Lo trabalh lexicografic es per se complir dins lo projècte del *Dictionnaire de l'Occitan Médiéval* del Professor Stempel (véser Chambon 2000b).

La bastison d'una basa de tèxtes orals es ja aviada dins l'encastre del projècte del Thesaurus Occitan (*THESOC*) manjat per l'UMR 6039 de l'Universitat de Niça e del CNRS. S'agís d'enregistaments sonòrs que son tanben transcriches. Son doncas dins lo domèni de l'occitan contemporanèu. Manca una basa de tèxtes escriches, que poiria èsser interfaciada amb lo modul tèxte del *THESOC* per çò que concernís l'oral-escrich coma los contes e lo teatre. La basa de tèxtes escriches poiria completar las donadas ja amassadas pel GIDILOC dins aquela amira lexicografica.

Prepausèrem doncas, a l'escasença de l'obrador de linguistica occitana d'abrial de 2006, d'aviar la bastison d'una tala basa de tèxtes en lenga occitana (la *Basa Telòc*) que recampariá de tèxtes escriches de totes menas e tanben d'oral-escrich. Lo projècte foguèt puèi descrich

dins Bras 2006. Dempuèi, avèm avançat, e ne sèm a complir lo primièr pretzfach, qu'èra de bastir una basa experimentalada d'un milion de mots. Per aquò far, mercés a l'ajuda de Robèrt Martí d'IDECO, avèm recampat un vintenat d'òbras contemporanèas, doncas ja jos format numeric, e las avèm encodadas dins lo lengatge XML (*eXtensible Markup Language*) en seguissent las nòrmas internacionals d'encodatge de tèxtes de la TEI (*Text Encoding Initiative*) dichas P5. Avèm establìt un partenariat amb l'ATILF a Nancy, qu'es lo laboratòri que manten *Frantext* e lo *TLFI* per bastir la *Basa Telòc* sul modèl de la basa *Frantext*. La basa Telòc serà mesa a la disposicion del public dins l'encastre del CNRTL (Centre National de Ressources Textuelles et Lexicales)¹⁴.

La basa experimentalada e sas aisinas d'interrogacion permetràn a cada cercaire de causir son còrpus de trabalh en combinant de critèris coma lo nom de l'autor, lo títol de l'òbra, la temporada, la data, lo genre textual, lo dialècte de l'occitan, la nòrma grafica. Poirem far de requèstas simplas : cerca de concordàncias, trachas de contèxtes que clauson una forma (un mot-grafic), una partida o una sequéncia de formas ; cerca de co-ocurréncias, calcul de frequéncias.

Puèi farem créisser la basa pauc a cha pauc. I a pro de matèria dins lo domèni occitan per se donar l'amira de passar a mantunes milions de mots. La basa dins sa pagèla reala serà classada d'un biais mai fin que la basa experimentalada. Pensam de l'estructurar per genre e per domèni, per tipe de supòrt, per temporada e data, per dialècte e varianta, e per tipe de grafia causida per l'autor. Volèm que la basa respècte las règlas de bastison d'un còrpus vertadièr que los genres, domènis, dialèctes, tipus de supòrt i sián representats d'un biais equilibrat. Podèm seguir per aquò, non pas l'exemple de *Frantext*, mas los dels *corpora* bèls de la lenga anglesa *The Bank of English* o *The British National Corpus*. Per aquesta rason, volèm pas redusir lo contengut de la basa a la literatura, mas i inclure tanben d'autres menas d'escriches-oralas coma los discorses politics, los sermons, coma las cronicas e los articles d'almanacs o de revistas, los tèxtes tecnicos juridics o ligats a de mestiers, las correspondéncias e los tèxtes orals transcriches.

Tre que la basa reala farà 4 o 5 milions de mots, poirem aviar lo tractament linguistic que permetrà d'enriquir los tèxtes, es a dire d'etiquetar las formas graficas (o las sequéncias de formas), amb d'informacions morfosintacticas per ne far una basa lematizada e categorizada. Aquò permetrà de far de requèstas mai complèxas coma o vegèrem per *Frantext*.

Lo motor de cerca, es a dire l'aisina que permetrà de far de requèstas sus la basa, serà pas de bon far : tanplan coma lo lematizador, deurà integrar lo gras naut de variacion dels tèxtes occitans, que siá al nivèl grafic, morfologic (flexional), lexical. Mas partèm pas de res, las autras lengas an ja permès de capitalizar d'experiéncias e d'aisinas per aquò far, e los collègas que nos van ajudar, los cercaires de la còla « Tractament Automatic de las Lengas » de CLLE-ERSS tanplan coma los de l'ATILF e del CNRTL, coneisson plan aquelas tecnicas.

Per clavar, volèm convidar los collègas a bastir la *Basa Telòc* amb nautres. Lo projècte es organizat en dos cercles concentrics : al còr del projècte sèm qualques linguistas e linguistas-informaticians qu'aurem las mans dins la pasta ; a l'entorn, dins un cercle mai larg, cadun pòt ajudar a far créisser la basa en nos fisant un tèxte, e aital venir contributor de la *Basa Telòc*. Dins aquel cercle larg i convidam de segur totes los que son implicats dins la produccion de tèxtes occitans coma autor o coma editor e tanben los qu'an de tèxtes occitans coma matèria de trabalh o objècte d'estudis, que siá en lexicografia, en literatura, en istòria, en etnologia, en sociologia, etc.

¹⁴ Lo trabalh de codatge es estat menat mercés al sosten financièr del CROM e del CNRTL.

Los tèxtes de la *Basa Telòc* seràn protegits del ponch de vista dels dreches dels autors : seràn pas accessibles al public dins lor integralitat, seràn pas que de colleccions de paragrafes coma resultas de las requèstas que seràn porgits. Nos permetrà de trabalhar amb de tèxtes jos dreches en respectant los dreches dels autors. Nòstre projècte es diferent d'un projècte de bibliotèca virtuala, ont la tòca es de metre a posita l'integralitat dels tèxtes, çò que se pòt far quand los tèxtes son liures de dreches. Es çò que fan los collègas del CIEL d'OC amb l'Universitat de Provença, qu'an bastit aquela aisina remirabla qu'es la *Bibliotèca virtuala de la Tor Manha*. Los dos projèctes son complementaris e lors portaires an decidit de s'associar per far avançar la coneissença de la lenga nòstra e de sas òbras.

Volontam d'autres associacions coma aquela, que nos sembla que, en mai de son utilizacion per la lexicografia occitana e la linguistica, la *Basa Telòc* poirà èsser una aisina preciosa per l'ensenhament de la lenga e sa salvagarda.

Referéncias bibliograficas

BOVES, (Abbat de), (1987), « L'abbé de Sauvages », *Mémoires et comptes-rendus de la Société scientifique et littéraire de la ville d'Alais*, 27.

BRAS, M., (2006), « Le projet TELOC », [Langues et Cité](#), 8, 9.

CHAMBON, J.P., (2000a), « *Alibert* vu d'Auvergne : notes de lexicographie critique en domaine occitan à propos de deux sources aurillacoises du *Dictionnaire Occitan-Français* », *Revue des Langues Romanes*, 104, 169-185.

CHAMBON, J.P., (2000b), « Un événement dans la lexicographie occitane : la publication du DOM », *Revue des langues romanes*, 104, 439-458.

DENDIEN, J., PIERREL, J.M. (2003), « Le Trésor de la Langue Française informatisé : un exemple d'informatisation d'un dictionnaire de langue de référence », *TAL (Traitement Automatique des Langues)*, numéro sur les dictionnaires électroniques, Hermès Sciences Edition, 44-2.

D'HERBE-HEILIGER M., (1998), Frédéric Mistral et les écrivains occitans dans le Trésor dóu Felibrige, A.I.E.O., Pau.

MASSON, P., P., (1941), *Henry Paschal de Rochegude*, Imp. Coop. du Sud-Ouest, Albi.

MARTIN, R., (2000), « Le Trésor de la Langue Française », in G. Antoine & B. Cerquilini (eds.) *Histoire de la Langue Française 1945-2000*, 969-971.

OLIER, J., (1901), *Henry Paschal de Rochegude*, Nouguiès, Albi.

RENOUF, A. (1994), « Corpora and historical dictionaries », in I. Lancashire & R. Wooldridge (eds) *Early Dictionary Databases*, Toronto : Centre for Computing in the Humanities, 219-35.

SINCLAIR, J. M. (Ed) (1987), *Looking Up*. London & Glasgow : Collins.

SWIGGERS, P., (2003), « Une perspective 'occitanienne' sur l'histoire et l'élaboration des langues romanes: les vues de Henri-Pascal de Rochemure (1819) », in: A. VANNESTE - P. DE WILDE - S. KINDT - J. VLEMINGS éd., *Memoire en temps advenir. Hommage à Theo Venckeleer*, pp. 437-452, Peeters, Leuven/Paris.

THOMAS, J., e alii (eds), (1982), *Josèp DAUBIAN, Òbras completas*, I.E.O., Albi.

WARTBURG W. (von), (1944), “Mistral Tresor dóu Felibrige ind die romanische Sprachwissenschaft”, *Zeitschrift für Romanische Philologie*, Tome 69, Niemeyer, Halle.

WARTBURG W. (von), KELLER H.E., GEULJANS R., (1969), *Bibliographie des dictionnaires patois galloromans (1550-1967)*, Droz, Genève.

WOOLDRIDGE, R. (2004), « Lexicography », in *A Companion to Digital Humanities*, Oxford-Maldon-Carlton: Blackwell, 69-78.

Version française (2003) « La lexicographie assistée par ordinateur » en ligne :

http://www.chass.utoronto.ca/~wulfric/articles/lao/lexico_fr.htm

Òbras lexicograficas

Òbras lexicograficas occitanas

ALIBÈRT Louis, (1966), *Dictionnaire occitan-français d'après les parlers languedociens*, I.E.O., Toulouse. 6 reedicions, totas de l'I.E.O. Tolosa, 1977, 1981, 1985, 1988, 1993, 1997.

BÉRONIE, Nicolas, (1824), *Dictionnaire du patois du bas-limousin (Corrèze), et plus particulièrement des environs de Tulle*, Publié par Joseph-Anne Vialle, Imprimerie Drappeau, Tulle.

COUZINIÉ, Jean-Pierre, (1850), *Dictionnaire de la langue Romano-Castraise et des contrées limitrophes*, Imp. Cantié et Rey, Castres.

DOUJAT, Jean, (1638), *Dictionari Moundi*, imprimat à Toulousa chès Jan Boudo.

HONNORAT, Simon-Jude, (1846-1848), *Projet d'un dictionnaire provençal-français ou dictionnaire de la langue d'oc, ancienne et moderne, suivi d'un vocabulaire français-provençal*, Repos, Digne.

LAUX, Christian, (1997), *Dictionnaire français-occitan : languedocien central*, IEO edicions, Pechlaurenç. reedicion 1999, Edicion augmentada 2004.

MISTRAL, Frédéric, (1886), *Lou Tresor dóu Felibrige ou dictionnaire Provençal-Français embrassant les divers dialectes de la langue d'oc moderne[...]*. 9 edicions o reedicions : 1878-1886, Ramondet-Aubin ; 1886, Roumanille-Champion ; 1932, Delagrave ; 1968, Bérenguié ; 1979, Petit ; 1979, Édisud (edicion prefaciada per J.C. Bouvier) ; 1979, Slatkine ; 1983, Édisud ; 2003, Princi Negre.

PIAT, Louis, (1893), *Dictionnaire français-occitanien, donnant l'équivalent des mots français dans tous les dialectes de la langue d'oc moderne*, Imprimerie Centrale du Midi, Montpellier.

ROCHEGUDE, Henri, (PASCHAL de), (1819), *Le Parnasse occitanien / Choix de poésies originales des Troubadours tirées des manuscrits nationaux*, Bénichet cadet, imprimeur-libraire, Toulouse.

ROCHEGUDE, Henri, (PASCHAL de). (1819). *Éssai d'un glossaire occitanien pour servir à l'intelligence des Troubadours*, Toulouse.

SAUVAGES (de) Boissier de La Croix, (1756—1785), *Dictionnaire languedocien-français*, Gaude, Nimes. reedicions :1820, (reedicion de son nebot Louis-Auguste d'Hombres-Firmas) Nimes ; 1971, Slatkine, Genève ; 1972, France-Expansion (reedicion de la de 1820), Paris ; 1993, Lacour, Paris.

STEMPEL, W.D. (ed.), (1996--), *Dictionnaire de l'occitan médiéval (DOM)*. Tübingen.
<http://www.dom.badw-muenchen.de/>

VAYSSIER, Aimé, (1879), *Dictionnaire patois-français du département de l'Aveyron*, Carrère, Rodez. reedicions : 1971, Slatkine, Genève ; 1979, Laffitte, Marseille ; 2002, Laffitte, Marseille.

Òbras lexicograficas francesas e anglesas

Collins COBUILD English Language Dictionary (1987, 2nd edition 1995, 3rd edition 2001, 4th edition 2003, 5th edition 2006) : <http://www.cobuild.collins.co.uk>

Oxford English Dictionary : <http://www.oed.com/>

Trésor de la Langue Française. Dictionnaire de la langue du XIXe et du XXe siècle (1789-1960). (1971-1994). Paris : CNRS.

Trésor de la Langue Française Informatisé (TLFi) : <http://www.atilf.fr/tlfi.htm>

Dictionnaire du Moyen Français (DMF). Base de Lexiques de Moyen Français (DMF1) : <http://www.atilf.fr/blmf>

Autras ressorgas informatizadas

BNC *British National Corpus* : <http://www.natcorp.ox.ac.uk/>

COM *Concordance de l'Occitan Médiéval*. <http://www.brepols.net/>

COM1. Dir. P.T. Ricketts en collab amb A. Reed, F.R.P. Akehurst, J. Hathaway, C. van der Horst, Turnhout: Brepols, 2001 (CD-Rom).

COM 2. *Les Troubadours. Les Textes Narratifs en vers*, Dir. Scientifica P.T. Ricketts, Dir. Tecnica A. Reed, en collab amb F.R.P. Akehurst, J. Hathaway, C. van der Horst, Turnhout: Brepols, 2005 (CD-Rom)

CNRTL : Centre National de Ressources Textuelles et Lexicales : <http://www.cnrtl.fr>

FRANTEXT : <http://www.frantext.fr/> et <http://www.atilf.fr>

THESOC (Thesaurus Occitan): <http://thesaurus.unice.fr/>

ANNÈXAS

ANNÈXA 1 : Una compilacion

Avèm realizada aquela compilacion per tal de mostrar l'ambiguitat de la situacion de la lexicografia que balança entre pauretat e riquesa. Pauretat quand los autors de dictionaris, que son pas de professionals de la lexicografia, practican per represa dels dictionaris ja realizats, tornan copiar sens acceptar totes las donadas. Per exemple. Lo sens de « taula » : per dire una pèira plana coma un dolmen, se trapa pas que dins Alibèrt. E l'expression « copar pas las taulas de marbre » se tròba dins un roman de Bodon e la sabi pas relevada dins los obratges citats. De mai a l'apauriment de las definicions o del nombre de sens diferents s'opausa la riquesa que se pòt legir per la compilacion de la documentacion lexicografica (e encara avèm pas tot passat). La compilacion es doncas lo trabalh que cadun de nosautres fasèm quand trabalham sus un mot ; mas qu'es pas jamai estat menat d'un biais sistematic.

TAULA, nom fem.

I. A. Mòble fach d'una susfàcia plana de fusta, de pèira o de metal sus de pès, per manjar, escriure, trabalhar, etc.

Polida sala, polidas taulas. (J. Bodon, Grands Jorns, 1973, p. 28).

E coma elses refusavan, en s'escusant d'aver dejà donat tròp de rambalh per ausar encara far de fraisses de manjisca, la pagesa los atrapèt pel bras e los assetèt a la taula. (E. Mouly, Bortomieu, 1973, p. 39).

Dins la sala, de las taulas quarradas de formicà blanc e de las cadieras de la meteissa color. (F. Daval, Los Ômes, 1987, p. 67)

Blau / coma lo davantal que la femna / escampa sus la taula / davant de se faire far l'amor, / drecha contra la paret de sa cosina ... (Y. Rouquette, Messa pels pòrcs, 1972, p. 112).

A l'entorn de la granda taula de garric, dins la granda sala ondrada de retraches de soldats, d'images de la Santa Família e de quauquas reclamas mossurencas, totes s'ataulèron amb la jòia al còr. (E. Mouly, Rajòls, 1986 [1930], p. 45).

Pausava sobre la taula çò que portava. Plan coma caliá botava cada causa a la siá plaça : forqueta, cotèl, veire, sièta, plat [...] (P. Gayraud, Una Filha, 1975, p. 28).

— Lexias nominalas complèxas.

Taula dobla loc. nom. Taula plegadissa, redonda o carrada.

Taula d'òste loc. nom. Taula comuna d'un restaurant, d'una albèrga ont las practicas manjan a la meteissa taula.

Taula d'assemblage loc. nom. Pèças de fustas assembladas sens cap de placatge. Presentacion dins lor ensemble de las diferentas fuèlhas d'un cadastre.

Taula d'un endavant de taulissa loc. nom. Pòste que pòrta l'avançada de la teulada.

— Locucions verbalas.

Quilhar la taula. loc. verb. idem.

Un còp èra, aviam / nòstra taula quilhada / a la paret dels cementèris. (I. Roqueta, L'escritura, 1988, p. 7).

Metre la taula. loc. verb. Pausar sus la taula tot çò que cal pel repais.

Quand Alari arribèt, la taula èra mesa amb una tauleta a costat pels menuts. (L. Còrdas, Sèt Pans, 1977, p. 88).

Una pichòta taula èra mesa per elses, sul costat d'una granda sala plena de miralhs, de tablèus, de boquets de flors. (E. Mouly, Bortomieu, 1973, p. 176).

La taula èra metuda. Lo sopar èra preste. L'ola bulhissiá suls andèrs. S'aprestèron de savorar lo repais tot servit. (Z. Bòsc, Racontes, 1988, p. 48).

Trobar taula mesa e lièit fait. loc. verb. Esser l'esperit tranquile, sens cap de pensament.

Levar la taula. loc. verb. Traire tot çò que s'èra pausat sus la taula.

Copar pas las taulas de marbre. loc. verb. Anar pas al cafè.

Se metre a taula. loc. verb. Se sietar a la taula per manjar.

La maire se metèt pas a taula qu'èran ja pro sens ela, mangèt tot fasant lo servici sens per aquò pèdre res de la convèrsa ont apondiá son mot a prepaus. (L. Còrdas, Sèt Pans, 1977, p. 89).

Se metre a avodar.

Èstre a taula. loc. verb. Èstre sietat a la taula per manjar.

Al cap de dètz minutas, totes èran a taula, davant una escudèla comola d'olada que noliá lo cau e lo grais d'auca. (E. Mouly, Bortomieu, 1973, p. 71).

Se téner plan a taula. loc. verb. Aver un brave apêtís.

Se ten melhor a taula qu'a caval. loc. form. Se ditz de qualqu'un que manja plan e qu'es gormandàs.

Valent a taula e pigre al trabalh. loc. form. Se ditz de qualqu'un de pigre, un fug l'òbra.

A taula! Per elipsi, formula per convidar lo monde a se sietar al torn de la taula. ¹⁵

Ardít, a taula! los cantaires! (E. Mouly, Bortomieu 1973, p. 45).

— Dins una comparason. [Idea de quicòm de planièr].

Pepon montèt dins l'autre panièr, e Riquet seguèt, a pèd, duscas sul planòl ont la rota ne va coma sus una taula, e tota drecha. (E. Mouly, Bortomieu, 1973, p.71).

¹⁵ Se vei dins l'exemple del tractament de « taula » cossí mancan d'atestacions d'unes emplecs. Podèm ajustar que per exemple avèm pas trobada l'atestacion dins la literatura escricha de l'expression lengadociana ausida mantunes còps per un occitanofòn nascut en 1898 de « En taula ! » alara que se tròba atestada per S. Palay.

B. [Sens especials]

— **Taula.** Pòste per portar lo mortier.

— **Taula fòrta.** Taula ont los boquièrs e los mazelièrs retalhan la carn.

— **Taula de comunion.** loc. nom. Taula ont lo prèire balha la comunion. — Sin. **Santa taula.**

S'es aprochat de la santa taula. (J.-P. Couzinié, Dict., 1850).

— **Taula de glaç.** loc. nom. Barra de glaç produsida dins una glacièra.

— **Taula traucada.** Pèça de fusta traucada sul comun, al privat, ont l'òm se seta.

— **Taula fermada.** Mòble amb de tiradors ont se recatan los vestits. (Honorat, Mistral).

— **Taula salanta.** Part d'una salina ont se fa la cristalizacion de la sal.

— **Taula.** Enumeracion metodicada d'una tièra d'informacions o dels capitols d'un libre.
A l'escòla lo regent nos ensenhava las taulas de multiplicacion.

TAULA

L'escritura publica o pas	5
Lo fuòc es al cementèri.....	27
Lo Castèl dels cans.....	65

(Ives Roqueta, *L'escritura*, 1988, p. 395).

— **Taula.** Espaci de tèrra mai long que larg, dins un òrt, un jardin, ont se fan de flors o de legumes.

Veni de semenar una taula de carròtas. (J.-P. Couzinié, Dict., 1850).

Una taula de lachugas. (M. D'Hombres / Gr. Charvet, *Dict.*, 1884).

— Per extension : Campet.

— **Taula.** Pèira planièra qu'es levada pel peirièr e que pòt servir de lausa. Grand ròc que la susfàcia planièra es a ras de tèrra.

Doas pèiras levadas pòrtan una taula de granit.

Lo Pont de la Taula se tròba en Arièja.

— **Taula.** Cèrcle de fusta o de fèrre qu'es lo torn de la ròda.

La taula èra cobèrta de pneumatics nòus.

— **Taula.** Pèça de fusta ressegada en long e mai o mens espessa. (Honorat)

— **Taula.** Emplegat dins lo sens de TAULIER o de BANC, taula ont se pausan las mèrças per vendre.

— **Taula**. Doana. Luòc ont se pagan las talhas. (Rohegude)

C. [Per metonimia]

— Çò qu'es servit a taula.

I a bona taula. (J.-P. Couzinié, *Dict.*, 1850).

— **Fèr bona taula**. loc. verb. Manjar plan. [Lagarde].

— **Una taula cofada**. loc. nom. Una taula plan provesida. (Doujat, 1628)

— Plaça que ten un escolan dins una classa. (J.-P. Couzinié, *Dict.*, 1850).

ANNÈXA 2 : un ensag de redaccion d'article de diccionari

Prepausam la redaccion d'una dintrada de diccionari amb « article » causit per talastre. Al delà de la definicion prepausada que demòra d'escriure, es la levada dels exemples posats dins la literatura occitana que permet sovent de trobar d'emplecs pertinents e de permetre una ierarquizacion de l'article. Una tala recèrca demanda d'oradas de lectura e se pòt pas far scientificament a la man, es necite un aplech tecnologic. Es de donar lo mai d'exemples possible dels emplecs d'un mot, emplegat sol o dins d'expressions, que servirà a trobar las diferentas occurréncias e per consequéncia a donar una definicion lo mai larg possible de l'espandi lexicografic de la lenga. La literatura escricha (literària, jornalística, etc) mas tanben oralá permetrà un còp enregistrada de donar aquela matèria al lexicograf.

ARTICLE nom masc.

I.A. Tèxte que forma un tot a despart, mas que fa part d'una publicacion editada o pas.

1. [Dins un jornal, un butletin, una revista].

Monsur lo Curat ne profetèt per escriure sul Roergat un article long, long ont pretendiá que lo vòstre filh atacava la familha, la religion, la morala, la Republica, etc, etc. (J. Bodon, *Letras*, 1986, [1949], p. 118)

Qualques jorns après, sus lo butletin de la paròchia i aviá un polit article : «Un cavalier benevòle» Cl. Majorel, Brassado, 1980, p. 36)

Dins los articles del Midi-Libre i aviá tanben aqueles Norvegians «gigants blonds en cerca d'arma-sòrre». (I. Roqueta, *Made in France*, 1970, p. 54)

2. [Dins un diccionari].

Dins lo prefaci [de J.-L. Fossat, *Dictionnaire occitan-français de l'agriculture*] son citats Alibèrt e Mistral coma exemples, Tanben d'autres autors apareisson dins los articles. (P. Sauzet, *Amiras*, 4, 1983, p. 96)

B. Division d'una lei, d'un contracte, d'un memòri.

II. Cada objècte qu'un mercadièr ten en botiga per la venda a la practica, produch per la venda dins lo comèrci.

Las tantas i avián vendut frofrós e ombrèlas, las nebòdas i vendián shòrts e biquinis a costat de l'article dit "d'usança" qu'èra la granda tradicion de l'ostal. (L. Còrdas, Sèt Pans, 1977, p. 18).

Al fons de la carrièra Drecha, anatz tustar encò del felibre Benazet : lo trobaretz a drecha en davalant, un merchant d'articles de pesca! vos podètz pas trompar. (E. Mouly, Bortomieu, 1973, p. 52)

III. [Gramatica e lingüistica]. Classa de mots que marcan per d'unas lengas la determinacion o l'indeterminacion del substantiu. Mot qu'aparten a aquela classa. L'article pòt èsser : definit, indefinit, contractat o partitiu.

Sovent l'article indefinit, al plural, es remplaçat coma l'article partitiu, per l'article definit ambe la preposicion de. (J. Salvat, Gramatica occitana, 1950, p. 22)

L'article jòga lo meteís ròtle qu'en francés, de mai a conservat melhor la valor demonstrativa dins los usatges del tipe seguent : Ai pas de mòstra, la qu'as vista es de mon paire. Aquel gos es lo de mon cosin. (L. Alibert, Gramatica, 1976, p. 72)

La sinerèsi es pas jamai notada graficament ; l'elision o es per l'article lo, la e los significants personals [...]. (R. Lafont, Ortografia, 1971, p. 61)

Article de comèrci loc. nom. Objècte per la venda sul mercat.

Aqueles obrièrs, obligats de se vendre al detalh, son una mèrça coma tot autre article de comèrci e son doncas expausats a totas las reviradas de la concurréncia e a totas las variacions del mercat. (K. Marx e Fr. Engels, Manifest del partit comunista, trad. J. Blanc, 1976, p. 34).

Far l'article de loc. verb. Lausar sas mèrças per las vendre, per las plaçar.

Lo monsuret fasiá pas l'article de sas accions als òmes qu'èron ataulats aquí. (P. Gayraud, Una filha, 1975, p. 20).

Èsser a l'article de la mòrt loc. verb. Esser per morir, far las telas.

Lo monde esperan pas, coma un còp èra, que la tatà o lo pipí siague a l'article de la mòrt per sonar lo medecin. (P. Gayraud, Libre del Causse, 1968, p. 189).